

# **Beurteilung der Prüfprozesseignung bei Fahrzeugversuchen mit mobilen Arbeitsmaschinen**

Zur Erlangung des akademischen Grades eines  
**Doktors der Ingenieurwissenschaften (Dr.-Ing.)**  
von der KIT-Fakultät für Maschinenbau des  
Karlsruher Instituts für Technologie (KIT)  
angenommene

**Dissertation**

von

Dipl.-Ing. Tristan Reich

Tag der mündlichen Prüfung: 15. Mai 2018

Hauptreferent: Prof. Dr.-Ing Marcus Geimer

Korreferent: Prof. Dr.-Ing Kai Furmans



# Kurzfassung

Der Nutzen einer Entwicklung kann am besten mit Ergebnissen aus der Praxis belegt werden. Im Bereich der mobilen Arbeitsmaschinen existieren allerdings nur wenige standardisierte und allgemein anerkannte Prüfprozesse. Dennoch fehlen in den meisten Publikationen mit Ergebnissen aus Fahrzeuguntersuchungen Informationen zum zugrunde liegenden Prüfprozess sowie dessen Präzision. Eine objektive Beurteilung der Ergebnisse ist somit nicht möglich.

Zur Qualitätssicherung bei Fahrzeugversuchen mit mobilen Arbeitsmaschinen wird deshalb in dieser Arbeit eine Methode zur Beurteilung der Prüfprozesseignung vorgestellt. Dieses Verfahren sieht eine Analyse des Prüfprozesses hinsichtlich Aussagekraft, Genauigkeit und Präzision vor. Die Abschätzung der Präzision basiert auf der Varianzanalyse (ANOVA) von Referenzversuchen und bildet den Schwerpunkt der Arbeit. Mithilfe verschiedener Variationen der Varianzanalyse können sowohl Einzelergebnisse als auch Messdatenverläufe ausgewertet werden. Weiter kann eine unmittelbare Gegenüberstellung der Vergleichsfahrzeuge die Präzision zusätzlich verbessern. In Anwendungsbeispielen wird der beträchtliche Einfluss des Untergrundes auf die Präzision der Zielgrößen Drehmoment, Drehzahl und Leistung im Fahrantrieb demonstriert.

Die Analyse der Prüfprozesseignung unterstützt bei der Gestaltung des Prüfprozesses und stellt bei angemessener Dokumentation Transparenz sowie Zuverlässigkeit der ermittelten Ergebnisse sicher.

**Schlüsselworte:** Prüfprozess, Versuch, Varianzanalyse, ANOVA  
Genauigkeit, Reproduzierbarkeit, Präzision



# Abstract

Best proof for the benefit of a technical development are the results of practical experience. However there are rarely standardised or common accepted test processes in the field of mobile machines. Though the majority of publications with results determined by vehicle tests is lacking regarding informations about the test processes as well as its precision. Hence an objective evaluation of these results is impossible.

This paper introduces a methodology for evaluating the adequacy of a test processes on purpose to establish quality assurance for vehicle tests with mobile machines. Analysis criteria of this method are validity, accuracy and precision. The estimation of precision is based on the analysis of variance (ANOVA) of reference tests. This is the focus of this paper. It is demonstrated how single values as well as dynamic courses can be analysed using different variations of ANOVA. Besides the precision can be risen by a direct comparison of the test vehicles respectively the technical development. Sample applications present the huge influence to the precision by the surface, examining torque, speed and power of the traction drive.

The analysis of the adequacy of a test process supports both designing the test procedure as well as ensuring transparency and reliability of the determined results, assuming an appropriate documentation.

**Keywords:** test process, trail, experiment, analysis of variance, ANOVA, accuracy, reproducibility, precision



# Danksagung

Die vorliegende Arbeit basiert auf meinen Untersuchungen von und mit dem Akustik-Allradrollenprüfstand des Institut für Fahrzeugsystemtechnik (FAST) am Karlsruher Institut für Technologie (KIT) während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Teilinstitut Mobile Arbeitsmaschinen (Mobima).

Mein besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr.-Ing. Marcus Geimer für die Übernahme des Hauptreferats und die wissenschaftliche Anleitung bei der Ausarbeitung meiner Dissertation. Herrn Prof. Dr.-Ing. Kai Furmans (Institut für Fördertechnik und Logistiksysteme, KIT) danke ich für die Übernahme des Korreferats. Ebenso bedanke ich mich bei Frau Prof. Dr.-Ing. Bettina Frohnafel (Institut für Strömungsmechanik, KIT) für die Übernahme des Prüfungsvorsitzes.

Außerdem danke ich meinen Kollegen, insbesondere Rinaldo Arnold, Bernhard Jahnke und Kilian Berthold sowie meinen Studenten mit denen ich zahlreiche anregende wissenschaftliche Diskussionen führen und eine großartige Zeit am Institut verbringen konnte.

Besonderer Dank gilt meinen Eltern und meinen Freunden für ihre Geduld, ihr Verständnis sowie ihre moralische Unterstützung über die gesamte Promotion.

Obersulm, im Mai 2018

*Tristan Reich*





# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Mobile Arbeitsmaschinen als Versuchsfahrzeuge</b>	<b>5</b>
2.1	Antriebsstrangkonfigurationen mobiler Arbeitsmaschinen	8
2.2	Arbeitsprozesse mobiler Arbeitsmaschinen	10
2.3	Arbeitsumgebungen und Umweltbedingungen	13
2.4	Fahrwiderstände mobiler Arbeitsmaschinen	15
2.4.1	Radwiderstand	16
2.4.2	Kräfte aus dem Arbeitsprozess	20
2.4.3	Gesamtfahrwiderstand	23
2.5	Bestehende Prüfprozesse für Gesamtfahrzeuguntersuchungen	24
2.5.1	Fahrzeugversuche mit Landmaschinen	25
2.5.2	Fahrzeugversuche mit Baumaschinen	27
2.5.3	Fahrzeugversuche mit Flurförderzeugen	34
2.5.4	Bedienereinflüsse bei Gesamtfahrzeuguntersuchungen	36
2.6	Forschungsbedarf	38
<b>3</b>	<b>Festlegen von Rahmenbedingungen und Versuchsablauf</b>	<b>43</b>
3.1	Versuchsablauf: stationär oder transient	44
3.2	Systemgrenzen für die Untersuchung	45
3.3	Anpassung der Arbeitsumgebung	47
3.4	Randomisierung der Versuchsreihenfolge	49
<b>4</b>	<b>Beurteilung der Prüfprozesseignung</b>	<b>51</b>
4.1	Aussagekraft des Prüfprozesses	54
4.2	Unsicherheit von Messergebnissen	56
4.2.1	Genauigkeit des Prüfprozesses	58
4.2.2	Präzision des Prüfprozesses	61

4.3	Statistische Methoden und Kennzahlen . . . . .	63
4.3.1	Lagemaße . . . . .	65
4.3.2	Streuungsmaße . . . . .	68
4.3.3	Visualisierung statistischer Kennzahlen . . . . .	69
4.3.4	Einfaktorielle Varianzanalyse . . . . .	71
4.3.5	Darstellung von Untersuchungsergebnissen . . . . .	81
4.3.6	Mehrfaktorielle Varianzanalyse . . . . .	82
4.3.7	Varianzanalyse für abhängige Stichproben . . . . .	85
4.3.8	Kovarianzanalyse . . . . .	88
4.3.9	Kruskal-Walis-Test . . . . .	90
4.4	Quantifizierung der Präzision . . . . .	92
4.4.1	Datenbasis für die Prüfprozessbeurteilung . . . . .	93
4.4.2	Varianz der Beobachtungen eines Referenzexperiments	97
4.4.3	Erzielbare Schärfe . . . . .	99
4.4.4	Auswerteverfahren für Direktvergleiche . . . . .	103
4.4.5	Erzielbare Schärfe bei Direktvergleichen . . . . .	113
4.4.6	Auswerteverfahren für transiente Versuche . . . . .	116
4.4.7	Erzielbare Schärfe bei transienten Versuchen . . . . .	121
4.4.8	Direktvergleiche mit transienten Versuchen . . . . .	122
<b>5</b>	<b>Die Prüfprozessbeurteilung in der Anwendung . . . . .</b>	<b>127</b>
5.1	Wiederholpräzision auf verschiedenen Untergründen . . . . .	127
5.1.1	Wiederholpräzision stationärer Versuche . . . . .	133
5.1.2	Wiederholpräzision transienter Versuche . . . . .	142
5.2	Präzision bei Direktvergleichen auf verschiedenen Untergründen . . . . .	150
5.2.1	Direktvergleiche stationärer Versuche . . . . .	150
5.2.2	Direktvergleiche transienter Versuche . . . . .	152
5.3	Grenzen der Präzisionsabschätzung . . . . .	154
<b>6</b>	<b>Zusammenfassung . . . . .</b>	<b>159</b>
<b>7</b>	<b>Ausblick . . . . .</b>	<b>163</b>
<b>A</b>	<b>Anhang . . . . .</b>	<b>165</b>

## Symbole

$\alpha$	°	Neigungswinkel des Untergrundes
$\alpha$	-	Wahrscheinlichkeit für Fehler 1. Ordnung
$\beta$	-	Wahrscheinlichkeit für Fehler 2. Ordnung
$\varepsilon$	-	Faktor zur Korrektur der Freiheitsgrade
$\mu$	-	Erwartungswert der Grundgesamtheit
$\rho_L$	$\frac{kg}{m^3}$	Luftdichte
$\sigma$	-	Standardabweichung der Grundgesamtheit
$\sigma^2$	-	Varianz der Grundgesamtheit
$A_L$	$m^2$	angeströmte Stirnfläche des Fahrzeuges
$a$	$\frac{m}{s^2}$	Beschleunigung
$ Bi $	-	systematische Messabweichung vom Normal
$b_G$	$m$	Arbeitsbreite des Arbeitsgerätes
$c_W$	-	Luftwiderstandsbeiwert
$df_{Indizes}$	-	Anzahl der Freiheitsgrade
$F$	$N$	Kraft
$F$	-	Prüfgröße der Varianzanalyse
$F_A$	$N$	Zugkraftbedarf aus dem Arbeitsprozess
$F_B$	$N$	Beschleunigungswiderstand
$F_{FW}$	$N$	Gesamtfahrwiderstand
$F_L$	$N$	Luftwiderstand
$F_{N,R}$	$N$	Radlast
$F_R$	$N$	Radwiderstand
$F_{R,\alpha}$	$N$	Schräglaufwiderstand
$F_{R,fr}$	$N$	Lagerreibung und Restbremsmomente
$F_{R,T}$	$N$	Radwiderstand des Reifens
$F_{R,Tr}$	$N$	Radwiderstand der Fahrbahn
$F_{R,TrO}$	$N$	Radwiderstand der Fahrbahn offroad
$F_S$	$N$	Steigungswiderstand
$f_B$	-	Faktor zur Beschreibung der Bodenbeschaffenheit
$f_R$	-	Rollwiderstandsbeiwert
$f_{R,pl}$	-	Rollwiderstandsbeiwert der plastischen Fahrbahn
$f_{R,T}$	-	Rollwiderstandsbeiwert des Reifens

$f_S$	$N$	Schlupfwiderstandsbeiwert
$Gr$	-	Gruppenzuordnungen
$g$	$\frac{m}{s^2}$	Erdbeschleunigung
$gf$	-	Gewichtungsfaktor
$h$	-	Testentscheidung
$I$	-	Anzahl der Experimente
$IQR$	-	Interquartilabstand
$J$	-	Anzahl der Versuchsreihen in einem Experiment
$K$	-	Anzahl der Versuche in einer Versuchsreihe
$L$	-	Anzahl der Stufen des Einflussfaktors $A$
$M$	-	Anzahl der Stufen des Einflussfaktors $B$
$M$	$Nm$	Drehmoment
$MQS_{Indizes}$	-	mittlere Abweichungsquadratsumme
$m_F$	$kg$	Fahrzeugmasse
$N$	-	Anzahl der Beobachtungswerte in jeder Faktorstufenkombination
$P$	$kW$	Leistung
$P$	-	Anzahl der Messzeitpunkte
$p$	-	Fehlerwahrscheinlichkeit
$Q$	-	Anzahl der Versuchspersonen
$QS_{Indizes}$	-	Abweichungsquadratsumme
$R$	-	Spannweite der Beobachtungswerte
$RS$	-	Rangsumme
$r$	-	Exponent des Gewichtungsfaktors $gf$
$S$	-	Schlupf
$S$	-	empirische Standardabweichung einer Stichprobe
$S^2$	-	empirische Varianz einer Stichprobe
$t_{Indizes}$	$s$	Messzeitpunkt
$t_G$	$m$	Arbeitstiefe des Arbeitsgerätes
$U$	-	erweiterte Messunsicherheit
$u$	-	Messunsicherheit
$V$	-	Variationskoeffizient
$v$	$\frac{m}{s}$	Geschwindigkeit
$X_{Indizes}$	-	Messwert der Zielgröße

$Y_{Indices}$	-	Beobachtungswert der Zielgröße
$\bar{Y}_{Indices}$	-	arithmetisches Mittel der Beobachtungswerte, Erwartungswert einer Stichprobe
$\tilde{Y}_{Indices}$	-	Median der Beobachtungswerte
$\tilde{Y}_p$	-	p-Quantil der Beobachtungswerte
$\check{Y}_{Indices}$	-	gewichteter Durchschnitt der Beobachtungswerte
$Z_{Indices}$	-	Transformierter Beobachtungswert der Zielgröße

## Indizes

$A$	bezüglich Einflussfaktor $A$
$A \times B$	bezüglich der Interaktion zwischen den Einflussfaktoren $A$ und $B$
$B$	bezüglich Einflussfaktor $B$
$C$	bezüglich Einflussfaktor $C$
$HA$	Hinterachse
$in$	innerhalb der Gruppen
$i$	Nummer des Experiments
$j$	Nummer der Versuchsreihe
$K$	korrigiert
$k$	Nummer des Versuchs
$l$	Variationsstufe des Einflussfaktors $A$
$m$	Variationsstufe des Einflussfaktors $B$
$n$	Wiederholungsnummer innerhalb einer Faktorkombination
$p$	Nummer des Messzeitpunktes
$PS$	Produktsumme bei der Kovarianzanalyse
$q$	Nummer der Versuchsperson
$R$	residualer / unsystematischer Streuungsanteil
$Ref$	Referenz
$s$	Beispielhafter Einzelindex eines Symbols
$T$	total / gesamt
$t$	Beispielhafter Einzelindex eines Symbols
$VA$	Vorderachse
$Vpn$	bezüglich der Versuchsperson
$zw$	zwischen den Gruppen

## Abkürzungen

AMT	automatisiertes Schaltgetriebe
ANOVA	analysis of variance, Varianzanalyse
ANCOVA	analysis of covariance, Kovarianzanalyse
ASAE	American Society of Association Executives™
AS-Reifen	Ackerschlepperreifen
DEM	Diskrete-Elemente-Methode
DMS	Dehnungsmessstreifen
DIN	Deutsches Institut für Normung
DLG	Deutsche Landwirtschafts-Gesellschaft e.V.
DoE	Design of Experiments, statistische Versuchsplanung
ELR	European Load Response Test
ESC	European Steady State Cycle
ETC	European Transient Cycle
FEM	Finite-Elemente-Methode
HSD	Honest Significant Difference, tatsächlicher signifikanter Unterschied; Test nach Tukey
GNSS	globales Navigationssatellitensystem
GUM	Guide to the expression of Uncertainty in Measurement
ISO	International Organization for Standardization
KST	Kolmogorow-Smirnow-Test
NEFZ	Neuer Europäischer Fahrzyklus
NRSC	Non-Road Steady Cycle
NRTC	Non-Road Transient Cycle
RMC	Ramped Mode Cycle
RTK	Real Time Kinematic, Echtzeitkinematik
SWT	Shapiro-Wilk-Test
TEAM	Projekt zur Entwicklung von Technologien für energie-sparende Antriebe mobiler Arbeitsmaschinen
VDA	Verband der Automobilindustrie e.V.
VDI	Verein Deutscher Ingenieure
WHSC	World Harmonized Stationary Cycle
WHTC	World Harmonized Transient Cycle

# 1 Einleitung

Mobile Arbeitsmaschinen erfuhren in den letzten hundert Jahren unzählige Neu- und Variantenentwicklungen und haben zwischenzeitlich einen sehr hohen Entwicklungsstand erreicht. Nichtsdestotrotz wird es auch in den kommenden Jahrzehnten noch zahlreiche Verbesserungen sowie so manche Neuentwicklung geben. Die entsprechenden Forschungen und Entwicklungen werden in Konferenzen, Fachzeitschriften und Produktvorstellungen Fachleuten, der Branche und schließlich den Kunden vorgestellt. Um die Bedeutung der vorgestellten Entwicklung zu untermauern, wird dabei oftmals die erzielte Verbesserung in Relation zum Vorgängermodell, zur Konkurrenz oder dem Klassenmittel angegeben. Durch diese Quantisierung wird dem Publikum ein unumstößlicher Beweis zur Wirkung der Entwicklung suggeriert.

Essentielle Informationen wie Rahmenbedingungen und Prüfprozess zur Entstehung dieser Zahlen werden nur vereinzelt erläutert. Ohne diese Informationen bleibt allerdings völlig unklar, wie sich diese Entwicklung im kundenindividuellen Alltagseinsatz des Fahrzeuges auswirkt. Ebenso fehlen sehr häufig Angaben zur Versuchsdurchführung und zur Ergebnisstreuung. Dies sind jedoch entscheidende Hinweise zur Zuverlässigkeit der propagierten Verbesserung. Somit sind Quantisierungen ohne entsprechende Angaben zu deren Entstehung praktisch wertlos. Die vorgestellten Unterschiede könnten ebenso gut durch die Streuung innerhalb des Prüfprozesses hervorgerufen werden. Die Entwicklung hätte damit keine signifikanten Auswirkungen auf die betrachtete Zielgröße.

Folglich stellt sich bei Untersuchungen mit mobilen Arbeitsmaschinen die Frage, wie die Eignung des verwendeten Prüfprozesses beurteilt und nachgewiesen werden kann.

Die detaillierte Erläuterung des zugrunde liegenden Prüfprozesses sowie dessen Präzision ist insbesondere im Bereich der mobilen Arbeitsmaschinen unerlässlich. Aufgrund der Vielfalt an Maschinentypen, dem mitunter beträchtlichen Größenspektrum innerhalb eines Typs sowie deren umfangreiche Bandbreite an Einsatzszenarien, gibt es nur vereinzelt standardisierte und allgemein anerkannten Prüfprozesse. Demzufolge muss für eine Bewertung der Betriebseigenschaften häufig vorab eine Vergleichsbasis, ein Prüfprozess, festgelegt werden. Dies ist im Spannungsfeld von Praxisrelevanz und Reproduzierbarkeit kein leichtes Unterfangen.

Unabhängig davon, ob ein Prüfprozess intuitiv oder nach einem methodischen Vorgehen entwickelt wurde, sollte geklärt werden, in wie weit sich der entworfene Prüfprozess für die Prüfaufgabe eignet. Während diese qualitätssichernde Maßnahme bei den Messprozessen im Bereich der industriellen Fertigung selbstverständlich ist, findet dieser Schritt bei Fahrzeuguntersuchungen nur selten Erwähnung und wird in zahlreichen Fällen übergangen oder schlicht vergessen. Demzufolge ist nur für einen Bruchteil der publizierten Ergebnisse auch die Qualität bekannt und abgesichert.

Um die Qualitätssicherung auch bei Fahrzeuguntersuchungen mit mobilen Arbeitsmaschinen zu etablieren, wird in dieser Arbeit eine Methode zur Beurteilung eines entsprechenden Prüfprozesses vorgestellt. Dabei liegt der Schwerpunkt auf der Ermittlung und Abschätzung der Präzision des Prüfprozesses. Schließlich ist von zentraler Bedeutung, welche Unterschiede bei der untersuchten Größe mit dem Prüfprozess erfasst werden können. Schlimmstenfalls streuen die Versuchsergebnisse so stark, dass der untersuchte Effekt mit dem gewählten Prüfprozess gar nicht differenziert werden kann. Die Beurteilung fällt unter diesem Umstand negativ aus.

## **Struktur**

Zum Einstieg in die Thematik werden in Kapitel 2 dieser Arbeit die Besonderheiten von mobilen Arbeitsmaschinen gegenüber sonstigen Fahrzeugen beschrieben. Zur Erfüllung ihrer Arbeitsaufgaben besitzen diese Maschinen spezielle Antriebsstrangkonfigurationen und müssen unter verschiedensten Bedingungen funktionieren. Da mobile Arbeitsmaschinen dabei häufig abseits befestigter Fahrbahnen agieren, liegt ein besonderes Augenmerk auf



den Fahrwiderständen. Die Beschreibung der Fahrzeugeigenschaften und Einsatzbedingungen verdeutlicht die unzähligen Störeinflüsse im realen Betrieb. In Kapitel 2.5 sind die Prüfprozesse diverser Veröffentlichungen zu mobilen Arbeitsmaschinen zusammengestellt. Aus diesen Beispielen lassen sich einige Ideen für die Gestaltung eines zweckmäßigen Prüfprozesses ableiten. In keiner der Veröffentlichungen werden Qualitätsmerkmale, wie die erreichte Trennschärfe des Prüfprozesses, vollständig benannt. Hieraus wird der Forschungsbedarf abgeleitet.

Zur Unterstützung bei der Ausarbeitung eines Prüfprozesses werden in Kapitel 3 einige grundlegende Entscheidungsalternativen diskutiert. Diese sind als anwendungsspezifische Ergänzungen zum methodischen Entwicklungsprozess anderer Autoren zu sehen.

In Kapitel 4 wird die Methode zur Beurteilung der Prüfprozesseignung anhand der drei Kriterien

- Aussagekraft
- Genauigkeit
- Präzision

beschrieben. Die Aussagekraft ist höchst nutzerindividuell und insofern am besten von diesem selbst anhand einer detaillierten Beschreibung zu beurteilen. Eine Quantisierung wird in dieser Arbeit nicht vorgenommen. Im Vergleich dazu ist die Genauigkeit mithilfe von Herstellerangaben zur Messtechnik relativ einfach zu beurteilen. Hierzu wird das Gauß'sche Fehlerfortpflanzungsgesetz sowie die lineare Fehlerfortpflanzung in Erinnerung gerufen. Die Bestimmung der Präzision bildet den Schwerpunkt der Arbeit. Hier werden zunächst die erforderlichen statistischen Methoden und Kennzahlen erläutert. Je nach Versuchsdurchführung finden unterschiedliche Varianten der Varianzanalyse Anwendung. So können sowohl stationäre als auch transiente Beobachtungswerte in unabhängigen Versuchen wie auch in Direktvergleichen ausgewertet werden. Schließlich erfolgt mithilfe der adäquaten Varianzanalyse eine Abschätzung, welche Messschärfe unter Annahme eines bestimmten Versuchsaufwands erreicht werden kann.

Die Anwendung auf reale Prüfprozesse wird in Kapitel 5 geschildert. Bei stationären und dynamischen Prüfprozessen auf verschiedenen Untergrün-

den wird der große Einfluss der Arbeitsumgebung auf die erzielbare Messschärfe aufgezeigt. Ebenso werden die Grenzen der Präzisionsabschätzung reflektiert. Zusammenfassung und Ausblick bilden wie üblich den Abschluss der Arbeit.

## 2 Mobile Arbeitsmaschinen als Versuchsfahrzeuge

In dem Buch *Grundlagen mobiler Arbeitsmaschinen* [GP14, S.I-3ff.] betrachtet GEIMER verschiedene Definitionen mobiler Arbeitsmaschinen und stellt seinerseits eine sehr weit gefasste Begriffsbestimmung auf:

Mobile Arbeitsmaschinen haben die Aufgabe der Verrichtung eines Arbeitsprozesses; sie verfügen dazu über einen Fahr- und Arbeitsantrieb mit signifikanten Energieanteilen in beiden Antrieben.

Mit dieser Definition werden mobile Arbeitsmaschinen sowohl von Fahrzeugen zur Personen- und Güterbeförderung als auch von stationären Maschinen abgegrenzt. Weiter wird von GEIMER eine branchenspezifische Untergliederung in folgende Bereiche vorgeschlagen:

- Landmaschinen
- Forstmaschinen
- Baumaschinen
- Kommunalmaschinen
- Hebe- und Flurfördermaschinen
- Sondermaschinen

Wie der Definition von GEIMER zu entnehmen ist, besitzen mobile Arbeitsmaschinen mindestens zwei, teilweise getrennte Antriebsstränge. Bei Untersuchungen mit diesen Maschinen muss deshalb zunächst geklärt werden, ob die zu untersuchende Größe, die Zielgrößen, von beiden oder nur von

einem der Antriebe beeinflusst wird. Bei Maschinen, die ihre Arbeitsaufgabe während der Fahrt ausführen, wirken sich Lastschwankungen in einem der Antriebe gewöhnlich auf den gesamten Antriebsstrang aus. Ebenso können Temperaturänderungen oder Schwingungen, die von einer Einzel-funktion verursacht werden, einen Effekt auf das Gesamtsystem haben. Je nach Untersuchungsziel muss dementsprechend entschieden werden, ob im Experiment beide Antriebsstränge aktiv sind oder einer davon bewusst konstant gehalten respektive deaktiviert werden kann.

Neben dem Antriebsstrang existieren zahlreiche weitere Faktoren, die das Messergebnis des Prüfprozesses beeinflussen. Im Ishikawa-Diagramm in Abbildung 2.1 sind wichtige Einfluss- und Störgrößen mit Wirkung auf die Messergebnisse gesammelt und gruppiert.

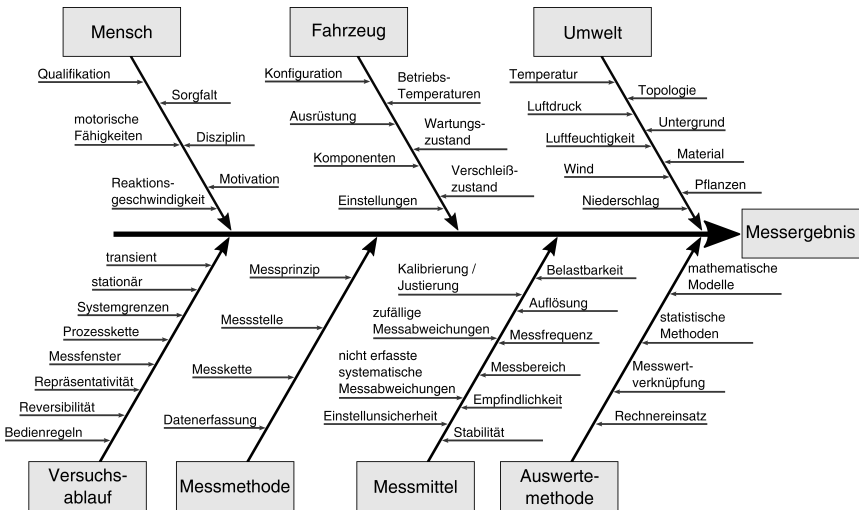


Bild 2.1: Wichtige Einflüsse auf die Messergebnisse (in Anlehnung an VDA 5 [VDA11])

Während Einflussgrößen gezielt für eine Untersuchung angepasst werden können, sind unter dem Begriff *Störgrößen* zufällige, unbeabsichtigte Einflüsse zusammengefasst. Bei Wiederholungen eines Versuchs unter gleichen Bedingungen führen diese Störungen zu einer unerwünschten Streuung der Untersuchungsgröße. Die Abweichungen zwischen derartigen Versuchswiederholungen wird mit der *Präzision* beschrieben. Je kleiner diese Abwei-

chungen ausfallen, desto höher ist die Stabilität des Prüfprozesses und damit die Präzision. Weiter wird zwischen *Vergleichspräzision* und *Wiederholpräzision* unterschieden. Letztere beschreibt die Abweichungen von Versuchswiederholungen unter vermeintlich identischen Bedingungen. Demgegenüber wird mit der Vergleichspräzision die Reproduzierbarkeit der Untersuchungswerte in einer unabhängigen Prüfung beschrieben. Zur Quantisierung der Präzision wird, wie in Kapitel 4.4 beschrieben, auf Streuungsmaße, beispielsweise die Standardabweichung, zurückgegriffen. Im Gegensatz zur *Genauigkeit* liefert die Präzision keine Informationen dazu, wie gut ein ermittelter Wert mit dem tatsächlichen Wert übereinstimmt. Gleichzeitig kann eine gute Genauigkeit nur mit einer hohen Wiederholpräzision erreicht werden und eine hohe Vergleichspräzision wiederum nur mithilfe einer guten Genauigkeit.

Unter der Zielsetzung Unterschiede zu ermitteln, interessiert insbesondere, mit welcher *Schärfe* zwei Ergebniswerte oder Ergebnisgruppen zuverlässig voneinander differenziert werden können. Diese minimale Differenz, die Schärfe, wird mithilfe statistischer Methoden auf Basis der Präzision, der Wiederholungsanzahl und einer Fehlerwahrscheinlichkeit berechnet. Detailliertere Erläuterungen sind in den Kapiteln 4.4.3, 4.4.5 und 4.4.7 für verschiedene Auswerteverfahren zu finden.

Um bei einer Untersuchung nicht nur einen kleinen Ausschnitt, sondern einen Großteil des üblichen Maschineneinsatzes abzuprüfen, ist eine eingehende Auseinandersetzung mit den Arbeitsprozessen der Maschine sinnvoll. Im Abschnitt 2.2 werden zwei Vorgehensweisen zur Ermittlung eines repräsentativen Prüfprozesses skizziert. Mit dem Arbeitsprozess sind häufig die typischen Umgebungsbedingungen des Einsatzortes verknüpft. Deshalb beschäftigt sich das Unterkapitel 2.3 mit den Einflüssen verschiedener Versuchsumgebungen. Diese wirken sich beispielsweise bei Offroadanwendungen wiederum auf die Fahrwiderstände im Fahrantrieb aus. Kapitel 2.4 gibt einen Überblick über die zusätzlichen Einflüsse beim Fahrwiderstand einer mobilen Arbeitsmaschine.

## 2.1 Antriebsstrangkonfigurationen mobiler Arbeitsmaschinen

Um sowohl der Fahrfunktion als auch der Arbeitsfunktion die erforderliche Leistung zur Verfügung zu stellen, verzweigt sich der Antriebsstrang einer mobilen Arbeitsmaschine üblicherweise in mindestens zwei Teile.

Als Antriebseinheit für mobile Arbeitsmaschinen hat der Dieselmotor die größte Verbreitung. Der entscheidende Vorteil für Mobilanwendungen liegt in der hohen Energiedichte und der guten Handhabbarkeit des Dieselmotorkraftstoffes. Im Bereich der Flurförderzeuge sind gleichermaßen batterieelektrische Antriebe üblich. Bei diesen Systemen wird die Leistungsverzweigung direkt nach dem Speicher, durch mehrere parallele Wechselrichter und Elektromotoren realisiert.

Zur Übertragung der Antriebsleistung innerhalb der Maschine werden die verschiedensten Arten von Getrieben verwendet. Zudem wird dabei die geringe Drehzahlstreuung des Verbrennungsmotors an die Anforderungen des Abtriebs angepasst. Prinzipbedingte Nachteile der Verbrennungskraftmaschine, wie das geringe Drehmoment bei kleinen Drehzahlen oder die drehzahlteste Verfügbarkeit der Maximalleistung (vgl. Abbildung 2.5) können so ausgeglichen werden. Neben mechanischen und elektrischen Lösungen haben sich bei mobilen Arbeitsmaschinen aufgrund der hohen Leistungsdichte, der aufgelösten Bauweise und der guten Regelbarkeit vor allem hydrostatische Systeme etabliert. Speziell lineare Bewegungen lassen sich mit verhältnismäßig einfachen Bauteilen sehr effizient umsetzen.

Im Bereich der **Arbeitshydraulik** wurden diverse Steuerkonzepte entwickelt, um beispielsweise mit variablem Druckniveau oder veränderbarem Volumenstrom die Verlustleistungen zu reduzieren. Gleichzeitig sollen dabei kurze Reaktionszeiten, eine gute Regelbarkeit und die zuverlässige Versorgung aller Verbraucher sichergestellt werden. Die Grundlagen der Hydraulik sowie die verschiedenen Steuerkonzepte sind in den Büchern *Hydraulik* [Mur12] und *Fluidtechnik für mobile Anwendungen* [ME14] ausführlich beschrieben.

Für den **Fahrtrieb** fasst NAUNHEIMER Getriebe und Anfaehrelement unter dem Begriff Fahrzeuggetriebe mit den folgenden vier Hauptfunktionen zusammen [NBL07, S.51].

- Anfahren ermöglichen
- Leistungsanpassung durch Wandlung von Drehmoment und Drehzahl sowie Wechsel der Fahrtrichtung
- permanente Leistungsübertragung durch verlustarme Übertragung ermöglichen
- Steuern der Leistungsanpassung

Je nach Prozessanforderungen und Nutzungsanteil unterscheidet sich die technische Realisation dieser Funktionen. So existieren sehr einfache Lösungen mit Reibkupplung und mechanischem Stufengetriebe aber auch hoch komplexe Systeme mit kupplungs- und stufenloser Verstellung des Übersetzungsverhältnisses.

Mobile Arbeitsmaschinen mit großen Fahranteilen auf befestigten Straßen werden angesichts der ähnlichen Anforderungen häufig auf LKW-Fahrgestellen aufgebaut. In diesem Segment haben in Europa zwischenzeitlich **automatisierte Schaltgetriebe** (AMT) die größte Verbreitung. Verlangt die Anwendung der Maschine eine Anpassung der Getriebeübersetzung ohne Unterbrechung der Zugkraft, kann diese Forderung mit einem **Lastschaltgetriebe** erfüllt werden. Alternativ lassen sich hohe Anfahrmomente und eine Anpassung des Übersetzungsverhältnisses unter Last mit einem **hydrostatischen Getriebe** umsetzen. Vorteile sind der im Vergleich zum Wandler höhere Wirkungsgrad und die stufenlose Verstellbarkeit. Des Weiteren kann auch die Fahrtrichtung unter Vollast gewechselt werden.

Um den hohen Wirkungsgrad eines mechanischen Getriebes mit einer stufenlosen Verstellung des Übersetzungsverhältnisses zu kombinieren, wurden **leistungsverzweigte Getriebe** entwickelt. Bei diesen Getrieben wird die Leistung parallel auf zwei Wegen, üblicherweise mit einem mechanischen und einem hydrostatischen Getriebe, übertragen. Zentrales

Element ist ein Planetengetriebe, das die Leistung in einem festen Momentenverhältnis summiert oder verzweigt. An Stelle eines hydrostatischen Zweiges könnte auch ein hydrodynamischer Wandler, ein mechanisches oder ein elektrisches Getriebe mit stufenloser Verstellung eingesetzt werden. Derartige Lösungen sind bei leistungsverzweigten Getrieben jedoch Exoten.

Rein elektrische Getriebe sind bei Flurförderzeugen und großen Muldenkippern im Einsatz. In diesem Fall erfolgt die Leistungswandlung mithilfe von stromrichterbasierten Elektroantrieben. Auch hier steht zum Anfahren bereits das maximale Antriebsmoment zur Verfügung. Eine separate Anfahrkupplung ist nicht erforderlich.

Viele der Arbeiten in der Erdbewegung oder der Land- und Forstwirtschaft werden auf unbefestigtem Untergrund durchgeführt. Dementsprechend sind Getriebe, Fahrwerk und Bereifung an den Einsatz abseits der Straße angepasst. Zur Übertragung der Antriebsleistung auf den Boden ist eine luftbefüllte Gummibereifung üblich. Über den Luftdruck können die Aufstandsfläche, Latsch genannt, und die Tragfähigkeit beeinflusst werden. In der Land- und Forstwirtschaft haben sich zur Reduzierung von Flurschäden besonders voluminöse Reifen mit niedrigem Luftdruck und damit geringer Flächenpressung etabliert. Daneben werden bei mobilen Arbeitsmaschinen Bandlaufwerke mit Ketten oder Gummibändern sowie Vollgummireifen oder Schreitbeine verwendet.

## **2.2 Arbeitsprozesse mobiler Arbeitsmaschinen**

Mobile Arbeitsmaschinen werden zur wirtschaftlichen Durchführung von Arbeitsprozessen entwickelt und eingesetzt. Typische Aufgaben sind Aufnahme, Sammeln oder Verteilen und Bewegen großer Lasten wie beispielsweise in der Logistik, bei der Holzwerbung, bei Feldarbeiten oder bei der Erdbewegung. Häufig muss ein Arbeitsprozess an verschiedenen Orten oder während einer Fahrbewegung durchgeführt werden. Beispiele hierfür sind Mäharbeiten, die Ernte mit einem Mähdrescher oder Feldhäcksler sowie die Straßenreinigung mit einer Kehrmachine. Ebenso gibt es Aufgaben, bei denen Materialien aus einem Speicher verteilt werden müssen, wie beim



Fertigen einer Asphaltsschicht oder dem Ausbringen von Saatgut, Dünger und Pflanzenschutzmitteln. Je nach Komplexität und Arbeitsanfall werden für diese Prozesse hoch spezialisierte Maschinen verwendet, die genau eine Arbeit erfüllen können, oder aber Universalmaschinen wie Mobilbagger oder Traktoren genutzt, die sich erst in Kombination mit einem geeigneten Anbaugerät für die geforderte Aufgabe spezialisieren.

Zur Erfüllung der Arbeitsaufgabe interagiert die Maschine mit ihrer Umgebung, wobei zusätzliche Kräfte auf das Gesamtfahrzeug und den Antriebsstrang wirken und diese beeinflussen. Da die meisten Arbeitsprozesse an natürlichen Produkten und Mineralien stattfinden, kann die Höhe dieser Kräfte in Abhängigkeit der Umweltbedingungen und der Durchführung der Arbeit in einem relativ breiten Bereich variieren.

Um einen Vergleich zwischen verschiedenen Maschinen oder Maschinenkonfigurationen anzustellen, muss nach Festlegung der Prüfprozessanforderungen die Vergleichsbasis, eine repräsentative Nutzung der Maschine, ermittelt werden. Während es bei hoch spezialisierten Maschinen noch verhältnismäßig einfach ist, einen typischen Prozess abzuleiten, ist das bei den vielseitigen Einsatzmöglichkeiten einer Universalmaschine kaum möglich. Aus diesen Gründen gibt es für mobile Arbeitsmaschinen fast keine allgemein anerkannten Vergleichszyklen. Ausnahmen sind die Richtlinie VDI 2198 [VDI02] für Flurförderzeuge und der DLG PowerMix [FD05] für Traktoren. Diese Zyklen und weitere Untersuchungen mit Gesamtfahrzeugen werden in Kapitel 2.5 näher erläutert.

Ein Konzept zur Bestimmung typischer Verfahrenszyklen wird im Themenschwerpunkt „*Ermittlung der Energieeffizienz mobiler Arbeitsmaschinen*“ des TEAM Projektes erarbeitet [SG13]. Eine Dokumentation dieser Methode findet sich bei SCHERER [Sch17a] [Sch17b]. Dabei ist die Grundidee, die Energieflüsse entlang des gesamten Antriebsstrangs im normalen Betrieb messtechnisch zu erfassen, um bei der Auswertung einzelne, sich wiederholende Arbeitsschritte sowie deren Zeitanteile zu identifizieren. Weiter ist in dieser Methode vorgesehen, die Betriebsmessdaten zu analysieren und daraus Last- und Bewegungskollektive als Vorschläge für Testzyklen zu ermitteln. Eine Expertengruppe entscheidet anhand von fünf Kriterien, welche Art von Prüfverfahren sich für die Untersuchung am besten eignet.

Schließlich folgt die Detaillierung und Konkretisierung des gewählten Prüfverfahrens. Die Entscheidungskriterien für die Auswahl des Prüfverfahrens sind [WSK<sup>+</sup>15, S.10]:

- Wiedergabe realistischer Einsatzfälle
- Abbilden von Einsatzfällen mit einer großen Häufigkeit bzw. einer hohen Energieintensität
- hohe Reproduzierbarkeit des Prüfverfahrens
- hohe Wirtschaftlichkeit
- hohe Realitätsnähe des Prüfverfahrens

Im Rahmen einer Diplomarbeit hat sich der Autor selbst bereits mit der Generierung von Verfahrenszyklen beschäftigt: Aus Beobachtungen und Messungen eines Forstspezialschleppers im normalen Arbeitseinsatz wurden über die Zeitanteile die drei wesentlichen Arbeiten des Fahrzeuges identifiziert. Des Weiteren wurden daraus die Arbeitsanweisungen für drei synthetische Vergleichszyklen auf einer befestigten Versuchsfläche abgeleitet. Über die gewichtete Summe der Einzelverbräuche können die Besonderheiten anderer Gelände und Arbeitsweisen berücksichtigt werden. Die Repräsentativität für andere Forstspezialschlepper, Fahrer und Gelände wurde nicht überprüft. [Rei10]

Die Wahl des repräsentativen Verfahrenszyklus hat einen großen Einfluss auf das Ergebnis, weil darin bestimmte Systemcharakteristika gezielt angesprochen oder umgangen werden können. Die Untersuchungsergebnisse gelten folglich immer nur in Bezug auf den Verfahrenszyklus. Allgemeingültigkeit besteht nur, wenn tatsächlich alle Aspekte des realen Fahrzeuges im Verfahrenszyklus repräsentiert sind.

Sollen die Lasten des Arbeitsprozesses durch einen Prüfstand aufgeprägt werden, so sind die Lastverläufe innerhalb dieses Vorgangs zu ermitteln. Ebenso ist für simulationsgestützte Untersuchungen häufig die Bestimmung eines repräsentativen Einzelzyklus aus einer realen Messreihe eines Verfahrenszyklus erforderlich. DEITERS synchronisiert dazu die Abschnitte seiner gemessenen Zykluswiederholungen und erstellt einen synthetischen

Einzelzyklus durch Mittelung [Dei09]. Die Autoren KUNZE, MIETH und VOIGHT verfolgen hingegen den Ansatz mittels statistischer Auswertung diejenige Einzelmessung auszuwählen, die am ehesten der Häufigkeitsverteilung der gesamten Messreihe entspricht [KMV11].

Anstelle eines festen Messdatenverlaufs setzen sowohl WEISSENBÖCK als auch STURM auf die Modellierung der Arbeitsaufgabe (vgl. Abschnitt 2.4.2). In beiden Arbeiten wird ein Grabkraftmodell aufgestellt, welches die Prozesskräfte als Funktion von Schaufelgeometrie, Eindringgeschwindigkeit und weiteren Einflussgrößen errechnet [Wei13, S.21-40] [Stu15, S.88-94]. Vorteile der Modellierung gegenüber einer festen Messwertvorgabe ist die Reaktion auf geänderte Eingangsparameter wie die Geschwindigkeit bei gleichzeitig hoher Reproduzierbarkeit.

Dieses Vorgehen ist in Gesamtfahrzeuguntersuchungen mit PKWs längst etabliert. Die Fahrwiderstände auf der Straße lassen sich verhältnismäßig einfach modellieren und mit einem Rollenprüfstand auf den Prüfling aufprägen. Auf dem Prüfstand werden mit den simulierten Widerständen Geschwindigkeitsprofile wie der *Neue europäische Fahrzyklus* (NEFZ) oder auch beliebige virtuelle Teststrecken abgefahren. Durch die Simulation des Fahrerhaltens und den Einsatz eines Fahrroboters können dabei gegebenenfalls Bedienerinflüsse minimiert werden. Unterschiede in der Modellierung der Fahrwiderstände von mobilen Arbeitsmaschinen sind in Kapitel 2.4 erläutert.

## 2.3 Arbeitsumgebungen und Umweltbedingungen

Die Arbeitsumgebung ist meistens eng mit dem Arbeitsprozess verknüpft. Eine Straßenkehrmaschine wird vor allem auf befestigten Wegen wie Asphalt eingesetzt, wohingegen Feldarbeiten in der Regel auf einem Acker durchgeführt werden. Folglich sind mit der Festlegung eines repräsentativen Arbeitsprozesses in der Regel auch einige Eigenschaften der Arbeitsumgebung bestimmt. Die Umweltbedingungen, die ebenfalls einen erheblichen Einfluss auf den Arbeitsprozess haben, müssen zusätzlich abgesteckt werden. Hierbei handelt es sich um Störgrößen die nicht gezielt eingestellt werden können.

Abseits befestigter Straßen besitzt der Boden einen großen Einfluss auf die Leistungsanforderungen an die mobile Arbeitsmaschine. Wie in Kapitel 2.4 beschrieben, hat die Deformation des Bodens einen wesentlichen Anteil am Fahrwiderstand des Fahrzeuges. Die Ausprägung dieser Verluste ist sowohl von Umweltbedingungen wie Wassergehalt, mineralischer Zusammensetzung, Korngröße, und Bewuchs als auch von Fahrzeugparametern wie Aufstandsfläche und Radlast bestimmt. Außerdem wird durch die Eigenschaften des Bodens die übertragbare Fahrantriebsleistung limitiert. Ebenso hängen die Prozesskräfte bei der Bodenbearbeitung beispielsweise bei einem Materialaushub oder Feldarbeiten von der Bodencharakteristik ab. Hier zeigen sich mitunter gegenläufige Tendenzen. Beim Pflügen eines weichen, nachgiebigen Bodens sind zwar geringere Prozesskräfte erforderlich, gleichzeitig sind jedoch hohe Fahrwiderstände zu überwinden, wodurch die Zugkraft stark reduziert wird oder sogar nicht mehr ausreicht. Umgekehrt nützt auch eine hervorragende Traktion nichts, wenn gleichzeitig die Prozesskräfte zum Beispiel bei der Füllung einer Radladerschaufel nicht mehr aufgebracht werden können. Hinsichtlich Reproduzierbarkeit sind die stark inhomogenen Eigenschaften des Bodens durch Zusammensetzung und Bewuchs sowie der ausgeprägte Einfluss vom wetterabhängigen Wassergehalt kritische Faktoren die kaum zu beherrschen sind.

Ähnlich komplex verhält es sich bei Arbeitsprozessen bei denen Pflanzen eine Rolle spielen. Je nach Bodenqualität, Nährstoffgehalt und Wetterverlauf während des Pflanzenwachstums unterscheiden sich die Ausbildung der Pflanze, Widerstandsfähigkeit und Ertrag. Des Weiteren können Pilz- oder Schädlingsbefall vorliegen. Ebenso ist der Erntezeitpunkt von großer Bedeutung. Gleichzeitig gibt es unzählige verschiedene Sorten einer Pflanzenart. Auch innerhalb eines Feldes sollte nicht von einem homogenen Pflanzenbewuchs ausgegangen werden. So sind beispielsweise in Maisfeldern vielfach Bereiche mit kleineren Pflanzen aufgrund von schlechteren Wachstumsbedingungen zu finden. Ebenso sind bei Kornfeldern häufig Flächenabschnitte zu sehen, in denen die Pflanzen durch die Belastungen von Wind und Ähren umgelegt wurden. Gleichermaßen werden sich im Forstbereich keine zwei gleich gewachsene Bäume finden.

Neben den Einflüssen auf Boden und Pflanzen hat das Wetter auch eine direkte Wirkung auf die Maschine. Lufttemperatur, Druck, Feuchtigkeit und Wind beeinflussen beispielsweise die Funktion von Sensoren, die Sicht des Fahrers, die Fahrzeugkühlung und das Betriebsverhalten des Verbrennungsmotors.

Ferner verändert die mobile Arbeitsmaschine während des Arbeitsprozesses selbst die Umweltbedingungen. Der Boden wird bei der Überfahrt verfestigt oder gegebenenfalls auch aufgewühlt. Pflanzen können unbeabsichtigt beschädigt werden und auch der erwünschte Arbeitsprozess führt im Allgemeinen zu irreversiblen Veränderungen. Selbst wenn ein gegrabenes Loch mit demselben Aushub wieder gefüllt wird, ergibt sich ein anderer Verfestigungszustand und eine neue Materialdurchmischung. Genauso lässt sich ein Feld nicht unmittelbar mehrfach abernten. Überall, wo die Maschine mit der Umwelt in Kontakt kommt, ist folglich mit unwiederbringlichen Veränderungen zu rechnen.

Natürlich gibt es auch einige Beispiele, auf die diese Aussage nicht zutrifft. Konstruktive Arbeitsprozesse, die einen synthetischen Zustand schaffen, beispielsweise das Auftragen einer Asphaltdeckschicht können unter weitestgehend ähnlichen Bedingungen durchgeführt werden. Gleichfalls führt die Arbeit mit Flurförderzeugen und Hubarbeitsbühnen im Regelfall zu Sachlagen, die sich wieder in den Ausgangszustand zurückführen lassen.

In Summe ergeben sich bei der Nutzung mobiler Arbeitsmaschinen zahlreiche Einflussfaktoren, die sich nur schwer gezielt einstellen lassen. Die Schaffung von übereinstimmenden Rahmenbedingungen für einen reproduzierbaren Prüfprozess gestaltet sich somit höchst anspruchsvoll. In Kapitel 3 werden aus diesem Grund Möglichkeiten aufgezeigt, die Ausprägung dieser unerwünschten Einflussfaktoren zu reduzieren.

## **2.4 Fahrwiderstände mobiler Arbeitsmaschinen**

Der Gesamtfahrwiderstand ist die Summe der Kräfte, welche der Bewegung des Fahrzeuges im Allgemeinen entgegenwirken. Während beim PKW einige Widerstandsanteile sehr klein ausfallen und deshalb vernachlässigt werden können, müssen diese Teilwiderstände bei mobilen Arbeitsmaschi-

nen berücksichtigt werden. Diese Einflüsse bedingen zusätzliche Störgrößen und werden im Folgenden umrissen. Gleichung 2.1 zeigt, wie sich der Fahrwiderstand  $F_{FW_{PKW}}$  als Summe der verschiedenen physikalischen Ursachen beim PKW zusammensetzt.

$$F_{FW_{PKW}} = F_R + F_L + F_S + F_B \quad (2.1)$$

Diese Aufteilung in

- Radwiderstand  $F_R$
- Steigungswiderstand  $F_S$
- Luftwiderstand  $F_L$
- Beschleunigungswiderstand  $F_B$

hat sich über Jahrzehnte in der Automobilbranche bewährt. Für mobile Arbeitsmaschinen muss jedoch die Zusammensetzung des Radwiderstandes  $F_R$  zur Abbildung von Fahrten im Gelände erweitert werden. Darüber hinaus ist ein zusätzlicher Term  $F_A$  zur Berücksichtigung der Kräfte aus dem Arbeitsprozess zu ergänzen.

### 2.4.1 Radwiderstand

Bei PKWs wird der Radwiderstand meistens als Produkt aus Rollwiderstandsbeiwert und Radlast dargestellt (Gleichung 2.2). Darin sind der Walkwiderstand, der Reibwiderstand im Latsch sowie Strömungsverluste zusammengefasst. Außerdem ist im Rollwiderstandsbeiwert eine Geschwindigkeitsabhängigkeit hinterlegt, mit der die Deformationswelle hinter dem Reifenlatsch berücksichtigt wird. Da diese Verluste innerhalb des Reifens auftreten werden sie als innerer Rollwiderstand bezeichnet. Vereinzelt werden zusätzlich äußere Rollwiderstände wie Unebenheiten und plastische Deformation der Fahrbahn, Schwallwiderstand sowie Schräglaufrkräfte berücksichtigt.

$$F_{R,T} = f_{R,T} \cdot F_{N,R} \quad (2.2)$$

Bei mobilen Arbeitsmaschinen, die in vielen Fällen zu einem beträchtlichen Anteil auf unbefestigten Böden fahren, ist der äußere Rollwiderstand dagegen obligatorisch. Durch die **Verluste in der Fahrbahn** kann sich der Radwiderstand mehr als verzehnfachen [Sch06, S.144ff.]. Bestandteile sind die Verformungsarbeit am Boden, der Bulldozingwiderstand und die Rei-

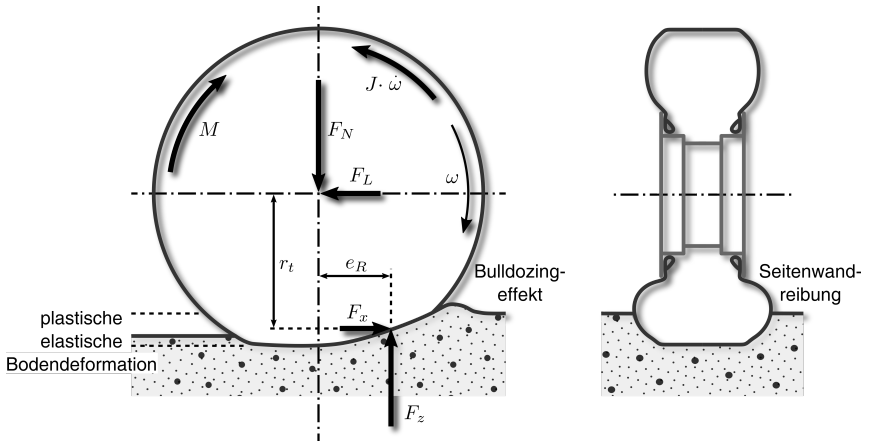


Bild 2.2: Kräfte und Momente am beschleunigten Rad auf nachgiebigen Untergründen (In Anlehnung an [Bar96, S.23])

bung an den Reifenseitenwänden. Die Skizze in Abbildung 2.2 verdeutlicht diese Effekte des äußeren Rollwiderstandes.

Zusätzlich kommt es bei mehrfacher Überfahrt von nachgiebigen Böden zur Reduzierung des Rollwiderstandes bei gleichzeitiger Erhöhung der Traktion, dem sogenannten Multipasseffekt (Abbildung 2.3). Weitere konstante, fahrbahnspezifische Einflussfaktoren auf den äußeren Rollwiderstand sind Bewuchs, Oberboden- und Unterbodenfestigkeit, die jeweils zu einem festeren Boden und geringerer plastischer Deformation beitragen.

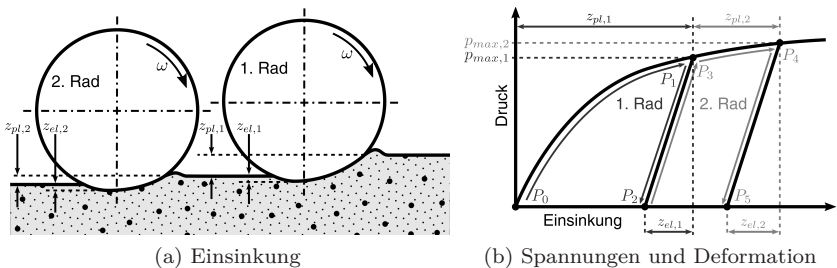


Bild 2.3: Be- und Entlastungsvorgang von elastisch-plastischen Böden beim Multipasseffekt (Nach [Har02, S.105,113])

Wie der innere Rollwiderstand auf ideal steifer Fahrbahn steigt auch der äußere Rollwiderstand annähernd linear mit der Radlast. Folglich kann ein gemeinsamer Rollwiderstandsbeiwert  $f_R$  als Summe aus den Widerstandsbeiwerten des Reifens  $f_{R,T}$  und der plastischen Fahrbahn  $f_{R,pl}$  gebildet werden. In Tabelle A.2 sind einige Werte für den Rollwiderstandsbeiwert  $f_R$  auf verschiedenen Untergründen aus der Literatur zusammengetragen.

$$\begin{aligned} F_{R,T} + F_{R,Tr} &= f_{R,T} \cdot F_{N,R} + f_{R,pl} \cdot F_{N,R} \\ &= f_R \cdot F_{N,R} \end{aligned} \quad (2.3)$$

Im Gegensatz zu der üblichen Annahme einer reinen Gewichtsproportionalität der Rollwiderstände stellen GÜNTER und FREY bei PKW Reifen eine **Abhängigkeit der Rollverluste von den Umfangskräften** fest [Gün94, S.76ff.] [FGG95, S.103-109]. GÜNTER teilt die am Reifen auftretende Verlustleistung in Roll- und Schlupfverluste, wobei lediglich letztere eine Abhängigkeit von der Umfangskraft aufweisen.

Auch BARRELMAYER leitet aus seinen Messdaten eine Verbindung zu den Kräften am Rad her, wählt jedoch den Schlupf  $S$  als Bezugsgröße [Bar96, S.152-156]. Ebenso wird im ASAE Standard 487.4 eine Gleichung für den Rollwiderstand von Diagonalreifen mit einem linear abhängigen Schlupfanteil angegeben [ASA99, S.350]. Da bei Straßenfahrten nur kurzzeitig größere Antriebsschlupfwerte auftreten, ist der Einfluss hier weniger ausgeprägt. Bei Fahrten unter anhaltend hohen Zugkräften, wie beispielsweise bei Feldarbeiten tritt dagegen erheblich Schlupf auf. Weil SCHREIBER bei seinen Daten keinen direkten Einfluss von Reifen- und Bodenparameter feststellen kann, verwendet dieser für die Schlupfabhängigkeit einen konstanten Wert  $f_S = 0,013$  N [Sch06, S.72]. Für AS-Reifen im Feldeinsatz ergibt sich dementsprechend der Rollwiderstand als Funktion von Radlast und Schlupf nach Gleichung 2.4.

$$\begin{aligned} F_{R,T} + F_{R,TrO} &= f_{R,T} \cdot F_{N,R} + f_{R,pl} \cdot F_{N,R} + f_S \cdot S \\ &= f_R \cdot F_{N,R} + f_S \cdot S \end{aligned} \quad (2.4)$$

Ein weiterer Fahrbahnwiderstand, der bei Offroadfahrten deutlich stärker ausgeprägt ist, entsteht durch die **Dämpfung von Bodenunebenheiten**



in Reifen und Fahrwerk [MW14, S.30ff.]. Unrundheiten sowie die Anregung durch die Fahrbahnstruktur führen zu radialen Einfederungen des Reifens. Da sowohl der Reifen als auch das Fahrwerk ein Feder-Dämpfer System darstellen, wird ein Teil dieser kinetischen Energie in Wärme umgesetzt [Wal07, S.36f.]. PLESSER beschäftigt sich in seiner Arbeit intensiver mit der Vertikaldämpfung und stellt einen degressiven Verlauf bei der Abhängigkeit von der Geschwindigkeit fest [PK96]. Im Anhang seiner Arbeit liefert er die Feder- und Dämpferparameter in Vertikal- und Längsrichtung für AS-Reifen in drei Größen unter Variation von Reifeninnendruck, Last und Temperatur [PK96, S.154-157].

Ist die Fahrbahn von einem Wasserfilm bedeckt, muss zusätzlich der **Schwallwiderstand** überwunden werden. Die bei der Verdrängung des Wassers zwischen Reifen und Fahrbahn entstehenden Verluste steigen mit dem verdrängten Wasservolumen pro Zeit. Ein Aufschwimmen der Reifen auf dem Wasser, sogenanntes Aquaplaning, tritt im Bereich der mobilen Arbeitsmaschinen praktisch nicht auf. Allerdings reduziert sich der Kraftschluss, wodurch der Schlupf zunimmt.

Wirken auf das Fahrzeug Querkräfte, bauen sich an den Rädern Seitenführungskräfte auf, die aufgrund des Schräglaufwinkels einen Anteil entgegen der Fahrtrichtung besitzen. Da bei Fahrzeugen, die auch im Gelände fahren, nur ein kleiner Vorspurwinkel von 0 bis  $0,2^\circ$  eingestellt wird, spielen **Schräglaufräfte**  $F_{R,\alpha}$  bei mobilen Arbeitsmaschinen nur während der Kurvenfahrt oder in Kombination mit dem Arbeitsprozess eine Rolle. Im Beispiel eines Graders müssen häufig sehr große Schräglaufwinkel gefahren und entsprechend berücksichtigt werden. Bei der Kurvenfahrt mobiler Arbeitsmaschinen können Schräglaufräfte, Antriebsstrangverspannungen sowie die Reduzierung des Multipasseffektes zu einem erheblichen Anstieg des Fahrwiderstandes führen.

Unabhängig von der Fahrbahn führen **Lagerreibung und Restbremsmomente**  $F_{R,fr}$  zu einer Verringerung der zur Verfügung stehenden Antriebsleistung. Solche Restbremsmomente entstehen, wenn sich die Bremsbeläge bei offener Bremse nicht vollständig von der Brems Scheibe lösen. Ebenso gibt es bei nassen Lamellenbremsen Restkontakte oder Strömungsverluste. Konkret kann bei einem Traktor eine höhere Brems Scheiben-

toleranz die Leistungsverluste bei 50 km/h um bis zu 6 kW reduzieren [Fuc11]. Bei einem 6500 kg schweren Traktor entspricht das einer Bremskraft von 430 N beziehungsweise etwa einem Viertel des Rollwiderstandes auf Asphalt.

## 2.4.2 Kräfte aus dem Arbeitsprozess

Durch den Arbeitsprozess einer mobilen Arbeitsmaschine können Kräfte in allen drei Raumrichtungen auf das Fahrzeug wirken. Zusätzlich kann sich die Beladung sowie die Position von Fahrzeugkomponenten ändern. Diese Belastungen führen gegebenenfalls zu einer Schwerpunktverlagerung sowie zu Änderungen der Aufstandskräfte, der Fahrwiderstände und der Seitenführungskräfte im Fahrwerk. Zudem erfährt das Antriebsaggregat des Fahrzeuges eine zusätzliche Belastung. Durch die Interaktion mit der Umwelt unterliegen diese Kräfte analog zur Inhomogenität von Boden, Pflanzen und Umschlagsmaterialien starken Schwankungen. Die wesentlichen Einflussgrößen auf diese Prozesskräfte wurden in diversen Untersuchungen für Modellierungszwecke herausgearbeitet. Eine Auswahl dieser Modelle ist im Folgenden zusammengefasst.

Insbesondere im Bereich der landwirtschaftlichen Bodenbearbeitung gibt es eine Reihe von Untersuchungen zur Abschätzung der erforderlichen Zugkräfte. Aus diesen Forschungen geht hervor, dass wendende Bodenbearbeitungsgeräte wie beispielsweise ein Pflug eine quadratische Abhängigkeit von der Geschwindigkeit  $v$  aufweisen [BH72, S.132], während nicht wendende Geräte wie ein Grubber einen linearen Anstieg der Zugkraft über die Geschwindigkeit zeigen. Arbeitsbreite  $b_G$  und Arbeitstiefe  $t_G$  gehen ebenfalls linear ein. Da sich jedoch in tieferen Bodenschichten häufig auch Zusammensetzung und Dichte des Bodens ändern, gilt der lineare Einfluss der Arbeitstiefe nur innerhalb einer näherungsweise homogenen Bodenschicht [Sch06, S.50-56]. Im ASAE Standard 487.4 wird zur allgemeinen Berechnung des Zugkraftbedarfs  $F_A$  von landwirtschaftlichen Geräten die Gleichung 2.5 angegeben. Eine Tabelle mit Werten für die dimensionslose Größe  $f_B$  zur Beschreibung der Bodenbeschaffenheit und die gerätespe-

zifischen Parameter  $A$ ,  $B$  und  $C$  sind ebenfalls Teil des ASAE Standards [ASA99, S.351].

$$F_A = f_B \cdot (A + B \cdot v + C \cdot v^2) \cdot b_G \cdot t_G \quad (2.5)$$

Des Weiteren ist eine Tabelle mit den erforderlichen Zapfwellenleistungen für verschiedene Anbaugeräte im Leerlauf, pro Gerätebreite und pro Gutfluss beinhaltet. Je nach Geräteausführung, -zustand und Pflanzeigenschaften werden Abweichungen von  $\pm 50\%$  erwartet [ASA99, S.353ff.]. Die dynamischen Anteile durch Inhomogenitäten im Boden oder im Bewuchs bleiben bei diesen Abschätzungen unberücksichtigt.

Da die Materialeigenschaften im Erdbewegungs- und Baubetrieb insbesondere aufgrund der Gesteinsanteile noch stärker variieren, scheint es hier unmöglich ein allgemeingültiges Modell aufzustellen. Bei GEHBAUER ist zu lesen, dass vor allem die Scherfestigkeit des Bodens den Widerstand gegen äußere Belastungen bestimmt. Bei der Bestimmung der Scherfestigkeit zeigt sich in verschiedensten Verfahren ein deutlicher Anstieg der Bruchscherfestigkeit mit der Abschergeschwindigkeit. GEHBAUERS Analyse der Scherfestigkeitsparameter ergab, dass der starke Festigkeitsanstieg vor allem durch den Kohäsionsanteil bewirkt wird, wohingegen der Reibungswinkel nahezu unabhängig von der Geschwindigkeit ist [GS94]. Kohäsion und Reibungswinkel werden wiederum stark vom Wassergehalt beeinflusst [Kno75].

Sowohl HEUMANN als auch KETTERER haben zur Ermittlung der Schnittkräfte beim Durchdringen verschiedener Böden Versuche mit Schneidbügeln durchgeführt [Heu75]. Letzterer erklärt den nichtlinearen Zusammenhang von Schneidkraft und Geschwindigkeit mit Hilfe dreier verschiedener Widerstandsanteile. Hier nennt er den Lösewiderstand, den Trägheitswiderstand zur Beschleunigung des Materials vor der Schneide und den Haufwerkwiderstand, der durch die Bewegung eines Bodenkeils vor und neben dem Werkzeug entsteht. Mit steigender Geschwindigkeit können sich die Materialkeile nicht so stark ausbilden, weshalb dieser Widerstand abnimmt. Die Trägheitskräfte steigen dagegen an, da mit zunehmender Geschwindigkeit das Material stärker beschleunigt wird. Der Lösewiderstand

hängt im Wesentlichen von den Bodenparametern ab und ist geschwindigkeitsunabhängig. In Summe ergibt sich in rolligen sowie in bindigen Böden ein degressiver Anstieg der Schneidkraft über die Geschwindigkeit, der von KETTERER mit einer Wurzelfunktion und einem Konstantanteil angenähert wird [Ket81, S.49-51,59].

Im Gegensatz zur Breite des Schneidbügels, deren Einfluss KETTERER mit einer reinen Potenzfunktion abbildet, wirken sich Schnitttiefe und Materialdichte linear auf die Schnittkräfte aus [Ket81, S.68f.]. Am Beispiel eines Steinbruchbaggers legt KÜHN dar, dass sich gerade bei schwierigem, grobkörnigem Material die Grabkräfte mit einer größeren Schaufel sogar reduzieren können. Gründe hierfür waren im konkreten Fall: Weniger Materialstau an den gut ausgerundeten Seitenwangen, besseres Eindringen der stark gepfeilten Vorderwand und die geringere Tendenz zur Brückenbildung durch die breitere Schaufel [KW83].

Die vielen Einflüsse von Material und Geometrie des Grabgefäßes erlauben es höchstens ein Modell für einen spezifischen Anwendungsfall aufzustellen. So versucht WEISSENBÖCK die Längskräfte beim Einstech- und Ladevorgang eines Radladers in einen Kieshaufen zu modellieren. Wie bei den Fahrwiderständen wird auch hier der Ansatz verfolgt, die Widerstandskräfte entsprechend ihrer Entstehung zu gliedern. Diese lassen sich in Reib-, Verformungs- und Beschleunigungskräfte kategorisieren [Rei14]. Der Einstechwiderstand wird in diesem Modell in Abhängigkeit von Eindringtiefe, Geschwindigkeit, Beschleunigung, Kippwinkel, Schaufelgeometrie und Schüttwinkel berechnet. Zur Abbildung der Materialinhomogenitäten wird den berechneten Schneidkräften eine Schwingung überlagert [Wei13, S.21-40]. Ziel war in diesem Fall eine qualitative Abbildung der Prozesskräfte in Abhängigkeit der Dynamik. Zur Bestätigung der quantitativen Aussagekraft fehlt eine umfangreiche Validierung.

Im Gegensatz dazu werden bei STURM die dynamischen Kraftanteile vernachlässigt. Er stützt sich bei dieser Entscheidung auf die Untersuchungen von BERGMANN, der den Widerstandsanteil durch Beschleunigungskräfte beim Bagger auf 1,6 % bestimmt hat ([Ber80] nach [Stu15, S.92]). Grundsätzlich verwendet STURM in seiner Simulation ein modifiziertes Grabkraftmodell von PARK [Par02], welches eine Erweiterung des MCKYES Modell

[McK85] darstellt [Stu15, S.28-34]. Abweichend von REICH und WEISSEN-BÖCK werden hier die Prozesskräfte in die drei Anteile Separationskraft, Schneidekraft und Verschiebekraft untergliedert. Erstere beschreibt die Kraft, die erforderlich ist, um die Erde vor dem Löffel zum Bruch zu zwingen und zu verdrängen. Die Schneidekraft wirkt, wie zuvor beschrieben, als Widerstand gegen das Eindringen der Zahn- und Seitenschneiden in den Boden. Gegen Ende des Grabvorgangs ist eine Verschiebekraft erforderlich, um Erde, die sich vor dem Löffel angesammelt hat, entlang der Erdoberfläche zu verschieben [Stu15, 88-94]. Die Validierung des Prozessmodells erfolgte indirekt über einen Vergleich zwischen simulierten und gemessenen Zylinderkräften. Qualitativ zeigen die Zylinderdruckverläufe eine gute Übereinstimmung, wenngleich die Grabkraftschwankungen aus dem realen Grabvorgang nur teilweise abgebildet werden. Für die angestrebte Effizienzbewertung sei die erreichte Genauigkeit jedoch ausreichend [Stu15, 93-101].

### 2.4.3 Gesamtfahrwiderstand

Entsprechend der hier erläuterten Erweiterungen kann der Gesamtfahrwiderstand  $F_{FW_{mA}}$  einer mobilen Arbeitsmaschine mithilfe von Gleichung 2.6 modelliert werden.

$$\begin{aligned}
 F_{FW_{mA}} &= F_R + F_L + F_S + F_B + F_A \\
 &= (f_{R,T} + f_{R,pl}) \cdot F_{N,R} + f_S \cdot S + F_{R,\alpha} + F_{R,fr} \\
 &\quad + c_W \cdot A_L \cdot \frac{\rho L}{2} \cdot v^2 \\
 &\quad + m_F \cdot g \cdot \sin(\alpha) \\
 &\quad + m_{res} \cdot a \\
 &\quad + f_B \cdot (A + B \cdot v + C \cdot v^2) \cdot b_G \cdot t_G
 \end{aligned} \tag{2.6}$$

Je nach Fahrsituation dominiert ein anderer Widerstandsanteil. Der Radwiderstand  $F_R$  ist stets vorhanden und erreicht bei großer Beladung und auf unbefestigtem Untergrund die höchsten Werte. Laut WUESTHOF liegt der Wirkungsgrad bei der Leistungsübertragung vom Reifen auf den Boden bei Traktoren zwischen 40 und 95 % [Wüs71]. Bei hohen Geschwindigkeiten

gewinnt der Luftwiderstand  $F_L$  zunehmend an Bedeutung. Der Steigungswiderstand  $F_S$  kann bei stärkerer Fahrbahnneigung ebenfalls die anderen Widerstandsanteile übertreffen. Während des Anfahrvorgangs oder sonstigen Geschwindigkeitsänderungen überwiegt häufig der Beschleunigungswiderstand  $F_B$ . Insbesondere bei Arbeitsprozessen zur Bodenbearbeitung erreichen die Prozesskräfte  $F_A$  ein Vielfaches der anderen Widerstände. Neben den in Gleichung 2.6 zusammengefassten Modellvorstellungen gibt es insbesondere für den Radwiderstand alternative Ansätze zur Berechnung. Eine Übersicht verschieden komplexer Voigt-Kelvin-Modelle zur Modellierung der Fahrdynamik ist bei HOLST sowie bei SCHUMACHER zu finden [Hol01, S.38-55] [Sch09, S.22f.]. Hierunter zählt auch das Hohenheimer Reifenmodell [Fer09, S.30-54]. Andere Autoren simulieren die Eigenschaften der Reifen mit einem FEM Modell. SMITH gelingt durch die Kombination von DEM Modell für den Untergrund und FEM Modell für den Reifen eine besonders detaillierte Simulation der Rad-Boden-Interaktion [SP13, S.277-286].

## 2.5 Bestehende Prüfprozesse für Gesamtfahrzeuguntersuchungen

Bei Forschungs- und Entwicklungsprojekten am Fahrzeug ergibt sich häufig der Bedarf einer Untersuchung am Gesamtfahrzeug. Die folgenden Ausführungen geben einen Einblick in die Literatur zu Gesamtfahrzeugversuchen mit mobilen Arbeitsmaschinen.

Aus dem Bereich der Abgaszertifizierung existieren aktuell für mobile Arbeitsmaschinen keine Vorgaben für Gesamtfahrzeugversuche. Stattdessen werden die Abgasemissionen am Motorenprüfstand in stationären Betriebspunkten (WHSC, NRSC/ISO 8178, ECE R49, ESC, 13-Mode Cycle) und in transienten Zyklen (WHTC, NRTC, RMC, ELR, ETC) gemessen [Rei10, S.16-20]. Der Bezug zu den tatsächlichen Betriebsbereichen des Verbrennungsmotors fehlt vollständig. Die Abgasemissionen im realen Betrieb der Maschinen bleibt unbekannt.

Aufgrund der früheren Verfügbarkeit für den Entwicklungsprozess und der Problematik bezüglich der Reproduzierbarkeit werden bei Forschungs- und

Entwicklungsprojekten Fahrzeugversuche vielfach in die Simulation übertragen. Parametrierung und Bewegungsabläufe dieser Simulationsmodelle basieren wiederum üblicherweise auf realen Versuchen mit einer vergleichbaren Maschine, beispielsweise dem Vorgängermodell. Folglich wurden auch in diesen Fällen mindestens ein repräsentativer Arbeitsprozess ausgewählt und praktische Versuche mit einer mobilen Arbeitsmaschine durchgeführt. Nicht selten beschränken sich die Untersuchungen dabei auf lediglich eine Maschine und einen Fahrer. Somit bleibt unklar, inwieweit die Messergebnisse für die Arbeitsaufgabe und nicht für die Spezifika der untersuchten Maschine, des Fahrers und der Umweltbedingungen statistisch repräsentativ sind. Mit aus diesem Grund werden aus derartigen Messungen fast nie allgemein akzeptierte Prüfprozesse abgeleitet.

### **2.5.1 Fahrzeugversuche mit Landmaschinen**

Eine der wenigen Ausnahmen für allgemein akzeptierte Prüfprozesse bildet der DLG PowerMix. Hier hat sich das Prüfzentrum der Deutschen Landwirtschaftsgesellschaft (DLG) mit Fahrzeugherstellern, Anwendern und Wissenschaftlern zusammengesetzt, um einen praxisnahen Vergleichstest für Traktoren zu entwickeln. Diesem Test liegt die Idee zu Grund relevante Einsätze in der Praxis zu erfassen, welche anschließend im Labor reproduziert werden [DF03].

Dazu wurden mit verschiedenen Traktor-Gerätekombinationen Feldmessungen durchgeführt und daraus zwölf Leistungsprofile für Zugarbeiten, die Zapfwelle und die Hydraulik ermittelt. Um die Reproduzierbarkeit sicher zu stellen wurden die Messungen vom Feld auf eine Betonbahn übertragen. Anstelle von Anbaugeräten werden die Lasten mit dem Bremswagen in Abbildung 2.4 aufgeprägt. Zusätzlich wird auf öffentlichen Straßen eine Transportfahrt mit Anhänger durchgeführt.

Anders als bei PKWs werden die Leistungsprofile für die Versuche auf die maximale Zapfwellenleistung des Prüflings skaliert. Als Zielgröße wird der spezifische Kraftstoffverbrauch ermittelt. Fahrgeschwindigkeit und Zapfwelldrehzahl werden am Traktor voreingestellt. Dementsprechend werden vom Messwagen vorab festgelegte Zugkraft-, Momenten- und Druckverläufe über die Zeit abgefahren. Je nach Ansprechverhalten und Drü-

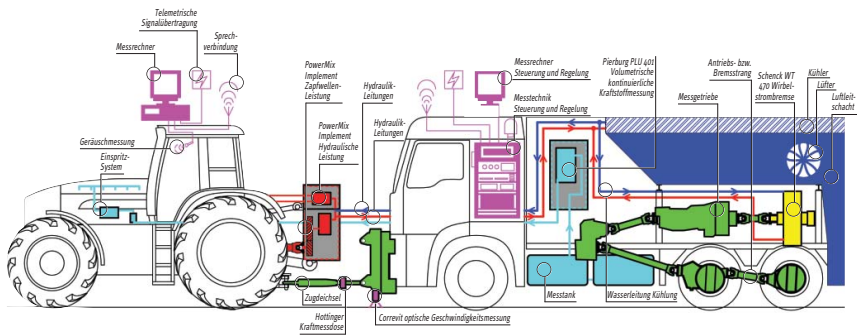


Bild 2.4: Testgespann aus Traktor und Messwagen im DLG PowerMix (Nach [Neu03])

ckung des Motors ergeben sich daraus Abweichungen bei den Geschwindigkeiten und Fahrwegen der unterschiedlichen Testkandidaten. Während die zurückgelegte Strecke von untergeordneter Bedeutung ist, wirkt sich die Geschwindigkeit auf den spezifischen Kraftstoffverbrauch sowie die Qualität des Arbeitsprozesses aus. Den Kennlinien in Abbildung 2.5 ist zu entnehmen, dass sich der spezifische Verbrauch des Beispieltractors bei Drückung vom Punkt maximaler Leistung bei  $2000 \frac{1}{\text{min}}$  auf den Punkt maximalen Drehmoments bei  $1500 \frac{1}{\text{min}}$  um 7 % verbessert. Prüflinge mit starker Motordrückung sind demnach im Vorteil.

Durch die Versuchsführung mit einem zeitfesten Kräfteverlauf bleibt die Geschwindigkeitsabhängigkeit des Arbeitsprozesses, wie in Kapitel 2.4.2 dargelegt, unberücksichtigt. Ebenso werden Trägheitskräfte der Anbaugeräte nicht abgebildet. Hier müsste bei der Kräftevorgabe ein Modell zur Berücksichtigung der Geschwindigkeit und Beschleunigungen hinterlegt werden. Dieses Vorgehen ist nur sinnvoll, wenn der Prüfstand über die Dynamik verfügt, diese Kräfte auch zu stellen.

Um in der Zukunft unabhängig von Wetter und Verkehr zu sein und eine höhere Reproduzierbarkeit zu erzielen, wurde am DLG Testzentrum 2015 ein Scheitelrollenprüfstand in Betrieb genommen, auf dem seit März 2017 der PowerMix Test durchgeführt wird. Wie zuvor können hier neben dem Fahrtrieb parallel die Zapfwelle und die Hydraulik des Prüflings belastet werden. Zusätzlich können an dem neuen Prüfstand elektrische Verbraucher simuliert werden [Vol14]. Ob sich mit der Verlagerung der Messungen



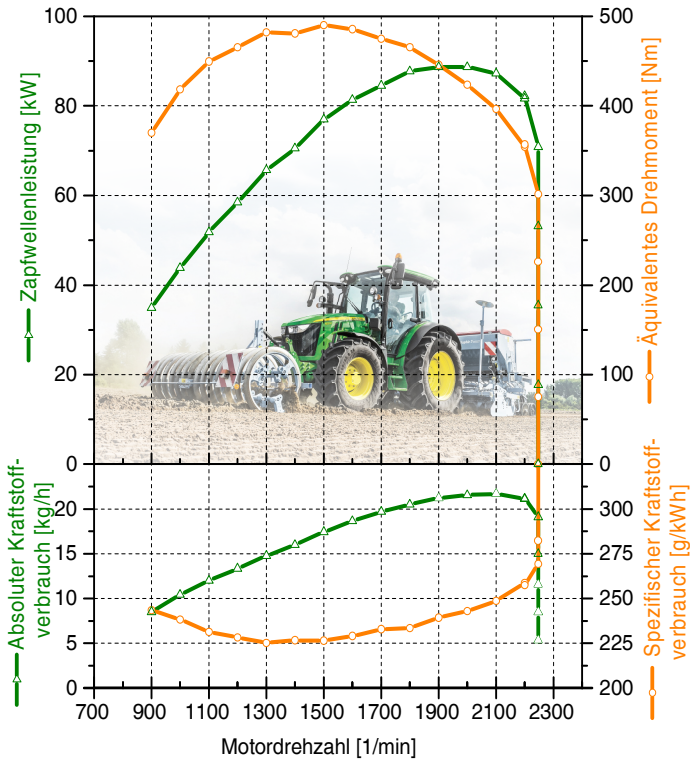


Bild 2.5: Motorkennlinien eines John Deere 5125R PowerShift Traktors (Nach [DLG17a])

auf den Prüfstand weitere Änderungen in der Versuchsführung ergeben, ist nicht öffentlich bekannt. Laut einer Pressemitteilung können auf dem Rollenprüfstand mit dem DLG PowerMix Wirkungsgradoptimierungen unter 1 % festgestellt werden [DLG17b].

## 2.5.2 Fahrzeugversuche mit Baumaschinen

Im Bereich der Baumaschinen existiert mit der DIN ISO 9245 eine Norm zur Leistung von Erdbaumaschinen, die jedoch lediglich Erläuterung zu Begriffen, Formelzeichen und Einheiten liefert [DIN95]. Ein Vorgehen zur Ermittlung der Leistung ist nicht beschrieben.

In Japan beschreibt das Construction Mechanization Association Standards Committee drei Testprozeduren zur Ermittlung des Kraftstoffver-

brauchs von Baggern, Planierraupen und Radladern. Mit diesen Versuchen sollen in Japan unter anderem die klimapolitischen Ziele zur  $CO_2$  Reduktion nachgewiesen werden. In den Testprozeduren sind die Rahmenbedingungen, die Anforderungen an die Messtechnik, die Versuchsdurchführung und die Auswertung beschrieben [JCM07a] [JCM07b] [JCM07c].

Weil die Umweltbedingungen den Kraftstoffverbrauch beeinflussen, wird aus Luftdruck und Temperatur in Abhängigkeit des Motorluftsystems ein Faktor berechnet, der innerhalb eines vorgegebenen Toleranzbereichs liegen muss. Für die Messtechnik wird der maximale Messfehler der Sensoren vorgegeben sowie geeignete Kraftstoffmesssysteme beschrieben. Vor den Versuchen soll das Fahrzeug warm gefahren und die Motordrehzahlen mit den Herstellerangaben abgeglichen und gegebenenfalls korrigiert werden. Tankfüllstand und Norm für den verwendeten Kraftstoff sind ebenfalls definiert. Die Versuchsdurchführung gliedert sich in Abhängigkeit des Fahrzeugtyps in drei bis vier Teile.

Beim Bagger sind dies ein  $90^\circ$  Ladezyklus, das Erstellen eines Planums, eine Fahrt und eine Leerlaufphase. Die Arbeitszyklen werden so schnell wie möglich mit Vollgas durchgeführt. Dabei wird der Löffel dicht über dem Boden geführt, ohne diesen zu berühren und bleibt leer. Theoretische Grabtiefe, Entladehöhe sowie Länge des Planums sind jeweils stufenweise in Abhängigkeit des Löffelvolumens klassiert. Der Kraftstoffverbrauch beim Fahren wird unter Vollgas im Kriechgang auf Beton oder einer anderen verdichteten, ebenen Fläche ermittelt. Nur in der fünf minütigen Leerlaufphase wird der Motor mit minimaler Drehzahl betrieben [JCM07a].

Für Planierraupen sind ausschließlich stationäre Versuche auf festem Untergrund definiert. Schwere Arbeiten sollen bei maximaler Geschwindigkeit im kleinsten Gang mit einer Zuglast von 40 bis 60 % der Fahrzeugmasse abgebildet werden. Für leichte Arbeiten wird der zweite Gang und eine Zuglast von 8 bis 12 % gewählt. Bei stufenlosen Getrieben werden für diesen Test drei Sollgeschwindigkeiten in Abhängigkeit der Fahrzeugleistung angegeben. Ob die Zuglast über eine Strecke von mindestens 30 m gegen ein reelles Haufwerk oder mithilfe eines Prüfstandes aufgebracht werden soll, geht aus der vorliegenden Übersetzung nicht hervor. Schließlich wird noch der Verbrauch im Stand bei einer Leerlaufphase gemessen [JCM07b].

Die Testprozedur für Radlader sieht ebenfalls drei Teilversuche vor. Ein kurzer Y-Ladezyklus mit einem  $60^\circ$  Winkel zwischen Haufwerk und Abladestelle soll die Hauptarbeit der Maschine repräsentieren. Die Schaufel ist durchgehend mit einem Gewicht entsprechend dem Schaufelvolumen und der Dichte von Erde beladen. Der Grabvorgang und das Auskippen werden mit Wartezeiten ersetzt. Die Zugkraft während des Einstechvorgangs wird gegen die Betriebsbremse aufgebracht. Der Teilversuch für Transportfahrten wird ebenfalls mit einem Gewicht in der Schaufel auf Asphalt, Beton oder einem verdichteten Boden durchgeführt. Die Fahrt geht für Fahrzeuge bis 110 kW über eine Distanz von je 50 m, bei größeren Fahrzeugen über je 100 m vorwärts und rückwärts ohne zu Wenden. Der letzte Teil ist wie zuvor eine Leerlaufphase [JCM07c].

Der erforderliche Versuchsumfang ist für jeden Teilversuch genau vorgeschrieben. Alle transienten Zyklen werden mehrmals direkt hintereinander ausgeführt und bilden jeweils eine Versuchsreihe. Bei den stationären Versuchen und Fahrten besteht eine Reihe aus nur einem Versuch. Aus fünf Reihen werden die Kürzeste und die Längste gestrichen. Die Verbleibenden drei bilden die Datenbasis für die Auswertung. Der Kraftstoffverbrauch im Leerlauf wird nur einmal über 600 s gemessen. Eine Gewichtung für die Verrechnung der Teilergebnisse ist nicht angegeben.

Durch den Verzicht auf reale Grabvorgänge, die Fahrten auf festem Untergrund und das Ignorieren von Ausreißen sind einige wesentliche Störgrößen ausgeschlossen. Gleichzeitig werden damit jedoch auch grundlegende Einflussfaktoren vernachlässigt, was die Praxisrelevanz enorm einschränkt. Mit der Vorgabe des maximalen Messfehlers ist zumindest die messtechnische Genauigkeit bekannt. Unklar bleibt der Einfluss der verbleibenden Störgrößen und damit das Auflösungsvermögen des gesamten Prüfprozesses.

Die Firma Liebherr wirbt auf den Datenblättern ihrer Radlader mit dem Kraftstoffverbrauch im „*Liebherr-Normtest*“. Bei diesem Test wird die Anzahl der Ladespiele ermittelt, die mit 5 Litern Kraftstoff durchgeführt werden können. Dabei soll in einem 35 Sekunden Arbeitszyklus Material von einem Haufwerk zu einer 20 m entfernten Position mit einer Ausschütthöhe von 2,5 m transportiert werden. Da weder der Untergrund noch das

Umschlagsmaterial näher spezifiziert sind, wäre die Reproduktion der angegebenen Werte eher Zufall. Obwohl der Versuch sehr praxisnah scheint, ist für den Kunden nicht abschätzbar, wie gut die angegebenen Verbrauchswerte den eigenen Einsatzfall und seine fahrerischen Fähigkeiten repräsentieren. Von Liebherr abgeführte Durchschnittsverbräuche liegen um mindestens 40 % unter den Ergebnissen des „*Liebherr-Normtests*“. Abbildung 2.6) nach [Lie17, S.27]

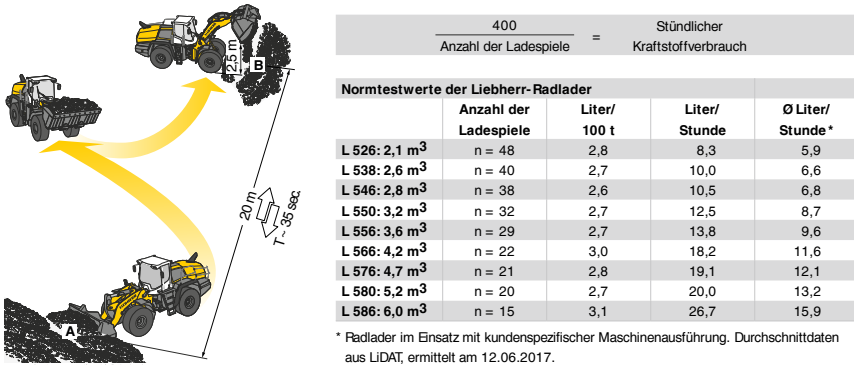


Bild 2.6: Arbeitszyklus und Ergebnisse des „*Liebherr-Normtests*“ (Nach [Lie17, S.27])

Der enorme Einfluss durch den Fahrer wird in Kapitel 2.5.4 ausführlich diskutiert. Bei den Untersuchungen von FRANK führten 73 verschiedene Fahrer jeweils drei Arbeitsaufgaben mit demselben Radlader durch [FSA12, S.6].

- Ein kurzer Ladezyklus mit Schotter in einer betonierten Halle auf einen Muldenkipper, ähnlich dem Liebherr-Normtest.
- Ein weiterer kurzer Ladezyklus über 10 – 15 m mit groben Felsbrocken in einem Steinbruch zur Beladung eines Knicklenkers.
- Ein Ladezyklus mit Schotter über 100 – 150 m zu einem höher gelegenen Bunker ebenfalls in einem Steinbruch.

Die beiden erstgenannten Aufgaben unterscheiden sich durch das Umschlagsmaterial und den Fahruntergrund, was im zweiten Fall die Füllung der Schaufel anspruchsvoller gestaltet. Die dritte Arbeitsaufgabe hat ei-

nen ausgeprägten Fahranteil inklusive einer Höhenänderung. Um bei seinen Messungen eine hohe Reproduzierbarkeit zu erreichen entschied sich FRANK vorab dazu, zumindest den ersten Ladezyklus in einer Halle statt in einem Steinbruch durchzuführen. Als reproduzierbares Umschlagmaterial wählte er vornehmlich Schotter. Für die zweite Arbeitsaufgabe wurde zur Vermeidung von Reifenschäden und wegen der gleichmäßigeren Größe kein gesprengter Fels sondern vorverarbeitete Felsbrocken verwendet.

Als wesentliche Störeinflüsse berichtet FRANK von der wetterbedingten Änderung der Feuchte seines Umschlagmaterials [FSA12, S.7f.]. Außerdem hatte er den Verschleiß des Schotters unterschätzt, der immer mehr feinere Anteile aufwies. Infolgedessen erleichterte sich das Füllen der Schaufel. Beide Einflüsse führen zudem zu einer Erhöhung der Dichte, wodurch die umgeschlagene Masse und damit die Produktivität unbeabsichtigt anstieg. Auswirkungen auf den Fahrweg und damit die Fahrwiderstände werden nicht erwähnt. Des Weiteren musste er während seiner Messungen zeitweise mit technischen Problemen am Fahrzeug und der Messtechnik zurecht kommen. Da über den Versuchszeitraum kein Fahrer wiederholt die Arbeitsaufgabe als Referenz durchgeführt hat, kann der Einfluss der genannten Störeinflüsse nicht quantifiziert werden.

Mit dem Fokus einen repräsentativen Einzelzyklus für die Simulation zu finden, hat auch DEITERS Versuche mit einem Radlader in einem kurzen Ladezyklus durchgeführt (Abbildung 2.7a). Dabei wurde ein Haufwerk aus losem Gestein mit Korngrößen von etwa 2 bis 200 mm Durchmesser zu einem neuen, circa 15 m entfernten Haufwerk transportiert. Diese Arbeitsaufgabe wurde von erfahrener Personal im Dauerbetrieb auf festgefahrener, trockenen Boden mit kleineren Unebenheiten durchgeführt. Bei der Auswertung der rund 150 Arbeitszyklen machte sich die Verlagerung des Haufwerks durch einen Drift bei den zurückgelegten Wegstrecken bemerkbar. Des Weiteren schwankten die Zykluszeiten im Bereich von  $\pm 4$  s, was bei den kurzen Zyklen  $\pm 12\%$  entspricht. Zur Generierung des repräsentativen Einzelzyklus skalierte DEITERS die Dauer der Einzelabschnitte auf den arithmetischen Mittelwert. Dabei berücksichtigte er die daraus resultierenden höheren Geschwindigkeiten sowie Beschleunigungskräfte und passte dementsprechend auch die Kraftverläufe seiner Messungen an. Das

am Beispiel eines Radladers beschriebene Vorgehen kann nach DEITERS auf andere mobile Arbeitsmaschinen übertragen werden [Dei09]. Wie repräsentativ der generierte Einzelzyklus für Radlader anderer Hersteller, abweichende Größen oder für weitere Fahrer ist, war nicht Teil der Untersuchung.

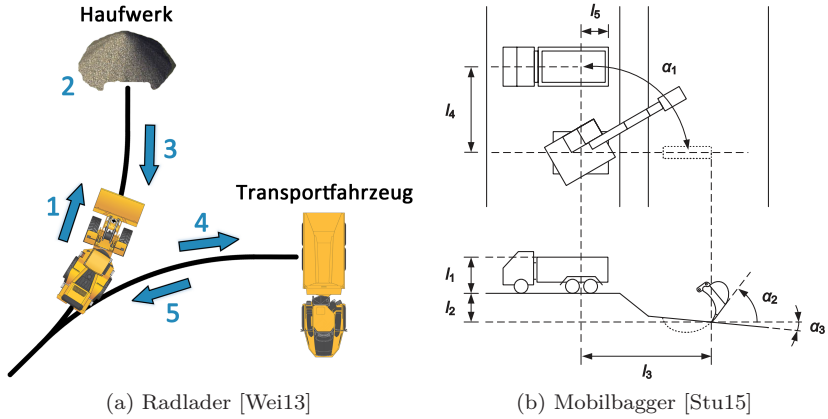


Bild 2.7: Arbeitszyklen von Baumaschinen

Für ihre Forschungen zu Hydrauliksystemen von Baggern haben sowohl MELCHINGER als auch HOLLÄNDER sowie STURM reale Messungen für verschiedene Grabeinsätze durchgeführt und auf ein Simulationsmodell übertragen. MELCHINGER hat dazu die Arbeitsbewegungen aus Einzelmessungen mit zwei verschiedenen Baggern auf Baustellen ausgewertet. Er beschränkte sich dabei auf die Koordinaten der Löffelspitze und des Schau-feldrehpunktes an verschiedenen Stationen des Arbeitsspiels wie beispielsweise vor und nach der Grabbewegung. Im Simulationsmodell werden diese Zustände mit optimalen Bewegungsabläufen verbunden. Eine derartige Beschreibung der Bewegung soll zu einer weitestgehenden Unabhängigkeit bezüglich der Kinematik des Baggerarms führen [Mel92]. Eine Übertragbarkeit auf andere Baggergrößen ist jedoch nur eingeschränkt sinnvoll. HOLLÄNDER stand für seine Versuche ein 30 t Raupenbagger mit zwei Fahrern zur Verfügung. Der Boden des Versuchsgeländes wird als homogen und grabfähig der Bodenklasse 3 zugeordnet. Des Weiteren wurden Dichte

und Wassergehalt bestimmt. Mit den genau beschriebenen Arbeitsaufgaben *Leistungseinsatz*, *Grabenaushub* und *Planum erstellen* werden in seinen Versuchen sowohl das Feinsteuerverhalten unter Teillast wie auch die Maximalleistung unter Volllast abgeprüft. Als repräsentativen Belastungszyklus für das Simulationsmodell wählt HOLLÄNDER den Einzelzyklus, der den arithmetischen Mittelwerten der Versuchsreihe am ehesten entspricht. Mit der Variation von Entladeart, Schwenkwinkel, Antriebsleistungsmodi, Hubhöhe und Grabtiefe wurde der Leistungseinsatz besonders detailliert analysiert. Die Auswertung dieser Parametervariationen zeigt die große Streuung innerhalb der Versuche. In einigen Fällen sind klare Tendenzen und Abhängigkeiten von dem veränderten Parameter erkennbar, wohingegen bei manchen Diagrammen eher zufällig verteilte Punktwolken analysiert wurden.

Obwohl beim *Leistungseinsatz* vom ersten Fahrer Schaufelfüllung und Zykluszeit um bis zu 10 % vom Mittelwert abweichen bleiben die Schwankungen im gewichtsspezifischen Kraftstoffverbrauch unter  $\pm 1\%$  [Hol98, S.82]. Bei dem zweiten Versuchsfahrer streuen dagegen fast alle Zielgrößen mit  $\pm 10\%$  [Hol98, S.102]. Zu den Abweichungen in den Arbeitsaufgaben *Grabenaushub* und *Planum erstellen* im Teillastbereich werden keine Angaben gemacht. HOLLÄNDER bezeichnet die ermittelten Lastspielzeiten als weitestgehend bodenunabhängig, da lediglich der Grababschnitt von den Bodenbedingungen beeinflusst wird. Dieser Einfluss ist allerdings erheblich, weil in diesem Abschnitt, der circa ein Viertel der Zykluszeit ausmacht, die höchsten Belastungen auftreten und auch die Schaufelfüllung hiervon bestimmt wird [Hol98, S.113]. Zudem ist zu erwarten, dass sich Dichte und Schaufelfüllung auch auf den Rest des Zyklus auswirken.

Davon unabhängig hält HOLLÄNDER seine Versuchsführung für den Vergleich verschiedener Herstellermodelle geeignet. Voraussetzungen sind eine homogene, steinfreie Bodenstruktur, die parallele Durchführung von identischen Versuchen am gleichen Ort und ein Lastspiel, das wenige feste Vorgaben beinhaltet sowie fließende Bewegungen zulässt. Weiter schlägt er vor alle Versuche zusätzlich mit einem Referenzfahrzeug durchzuführen um sich bei den Ergebnisauswertung auf diese Referenz zu beziehen. Eine ähnliche Vorgehensweise ist bei Mähdreschern bekannt, deren Leistung

wesentlich von den Guteigenschaften des Jahres und vom Standort des Versuchsfeldes abhängen [Hol98, S.148ff.].

Zur Bewertung der Energieeffizienz von Antriebssystemen mobiler Arbeitsmaschinen entwickelt und beschreibt STURM eine Methode am Beispiel eines Mobilbaggers [Stu15]. Hierzu leitete er aus realen Messungen mit einem Mobilbagger der 15 bis 20 t Klasse Bedienregeln für ein Fahrermodell ab. Als Arbeitsaufgabe wählt STURM das in Abbildung 2.7b skizzierte 90° LKW Ladespiel ähnlich dem *Leistungseinsatz* von HOLLÄNDER [Stu15, S.35f.]. Die Fahrzeugmessungen wurden von drei professionellen Fahrern in einem Steinbruch mit lockerem Boden gefahren. Mit dem Fokus auf der entwickelten Methode sind die Rahmenbedingungen der Messungen weniger detailliert beschrieben. Wie stark sich die Einzelzyklen innerhalb einer Versuchsreihe unterscheiden ist nicht dokumentiert. Dafür geht STURM auf den beachtlichen Einfluss des Fahrers ein (vgl. Abschnitt 2.5.4). Damit bekräftigt er seinen Ansatz die Effizienzuntersuchung in die Simulation zu übertragen. Das Fahrzeug und die auftretenden Grabkräfte werden ebenfalls mithilfe von Modellen in der Simulation nachgebildet. Somit können der Fahrereinfluss und alle weiteren Störgrößen eliminiert werden. Die Qualität der Ergebnisse wird dabei durch die Modellgüte beschränkt.

### 2.5.3 Fahrzeugversuche mit Flurförderzeugen

Im Bereich der Flurförderzeuge wurde durch den Verein deutscher Ingenieure die Richtlinie VDI 2198 herausgegeben. Hierin ist beschrieben, welche Angaben das Typenblatt eines Flurförderzeuges beinhalten soll und wie diese zu ermitteln sind. Zur Ermittlung des Energie-/Kraftstoffverbrauchs wird ein Zyklus definiert, bei dem das Fahrzeug über eine Stunde mit einer Nennlast zwischen zwei Punkten pendelt und dort jeweils einen Hub durchführt [VDI02]. Die Arbeitsgeschwindigkeit ist an die fest vorgegebene Zyklenzahl anzupassen, wodurch Produktivitätsunterschiede ausgeschlossen werden. Da die Last durchgehend auf dem Fahrzeug bleibt fehlen die Arbeitsabschnitte Leerfahrten und Absenken ohne Last. Bei den Umgebungsbedingungen wird ein trockener, ebener Boden mit einer Reibzahl zwischen Reifen und Fahrbahn definiert, wie er in den meisten Lagerhallen vorkommen dürfte. Die verhältnismäßig geringe Anzahl an



Störgrößen sowie die Mittelung über eine Stunde lässt eine relativ hohe Reproduzierbarkeit der Zyklusergebnisse erwarten.

Um die Leistungsfähigkeit ihrer Produkte klarer herausstellen zu können, hat die Linde Material Handling GmbH einen Gabelstapler Leistungstest aus typischen Praxisfällen in Industrie- und Handelsunternehmen abgeleitet. In diesem Versuch muss ein Stapler zwölf Gitterboxen mit jeweils ca. 800 kg schnellstmöglich von einem LKW in ein Lager entladen und anschließen wieder beladen (Abbildung 2.8a). Der TÜV Nord Mobilität GmbH & Co. KG bescheinigt in einem Zertifikat, dass mit diesem Leistungstest Umschlagsleistung und Energieverbrauch unter realistischen Einsatzbedingungen ermittelt werden [Lin09]. Weitere zertifizierte Leistungstests existieren für Niederhubwagen sowie für Schubmaststapler (Abbildung 2.8b, 2.8c). Wie groß hierbei der Fahrereinfluss ist, bleibt unklar. Weiterhin fehlt eine Anpassung der Palettengewichte an die Nennlast des Fahrzeuges.

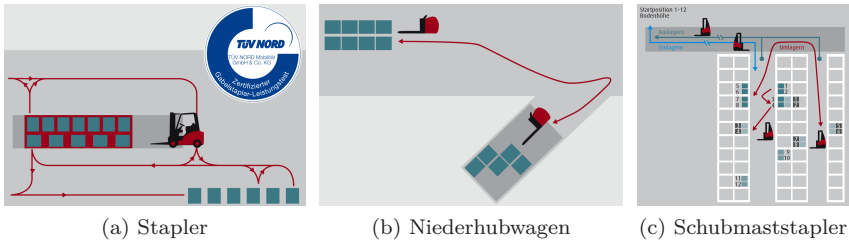


Bild 2.8: TÜV zertifizierte Leistungszyklen für Flurförderzeuge von Linde [Lin10]

In einer weiteren Pressemitteilung wird von drei abweichenden Arbeitsaufgaben für Schwerstapler der 16 t Klasse berichtet. Durch die Einzelversuche Streckentransport über eine längere Distanz, Umschlag über kurze Distanzen und Einsätze mit hoher Leistungsabfrage mit Anbaugeräten sollten die sehr unterschiedlichen Einsatzfelder dieser Geräteklasse abgebildet werden [Lin13]. Eine Zertifizierung durch den TÜV ist hier nicht ersichtlich. Außerdem ist die Beschreibung der Versuchsführung nicht detailliert genug um die Testergebnisse in unabhängigen Versuchen nachzuvollziehen.

Die Redaktion der VerkehrsRUNDSCHAU definiert für ihren Staplertest einen Parcours zur Be- und Endladung eines Sattelauflegers mit 32 Pa-

letten mit je 800 kg Prüfgewicht [Kra14]. DERENBACH führt dagegen in seinem „Staplercheck“ die Be- und Entladung eines LKWs mit lediglich sechs Europaletten unbekanntes Gewichtes mit maximaler Geschwindigkeit durch [Der09]. Die vielen leicht abgewandelten Varianten zeigen, dass die VDI Richtlinie aufgrund der Mängel bei der Repräsentativität nicht allgemein anerkannt ist.

### **2.5.4 Bedienerinflüsse bei Gesamtfahrzeuguntersuchungen**

Die Bedienung einer mobilen Arbeitsmaschine lässt sich je nach Vorkenntnissen und Komplexität der Maschine innerhalb weniger Stunden oder Tage erlernen. Die Effizienz im Umgang mit der Maschine kann über Erfahrung und Trainings jedoch auf das Dreifache gesteigert werden. Bei der Produktivität sind sogar Steigerungen bis auf das Achtfache möglich [FSA12, S.16]. Der Fahrer übernimmt eine zentrale Position im Regelkreis des Arbeitsprozesses mit wesentlich mehr Ein- und Ausgangsgrößen als beispielsweise bei einem PKW. Besonders herausfordernd ist dabei die zueinander abgestimmte parallele Bedienung mehrerer Funktionselemente der Maschine. Außerdem kommt zum Verbrauch die Produktivität als weitere Optimierungsgröße hinzu. Entsprechend groß sind die Unterschiede im Vergleich von derselben Maschine und gleicher Arbeitsaufgabe mit unterschiedlichen Fahrern. [FSA12]

Im Rahmen seiner Untersuchungen hat FRANK 73 Fahrer über jeweils eine Stunde mit demselben Radlader in drei verschiedenen Arbeitsaufgaben beobachtet. Die hauptberuflichen Radladerfahrer übertrafen dabei Arbeiter, die nur gelegentlich einen Radlader bedienen um das Zweieinhalbfache bei der Kraftstoffeffizienz und das Vierfache bei der Produktivität. Der Abstand zu frisch angelernten Personen mit nur wenigen Stunden Bedienungserfahrung war nochmals deutlich größer. Aber auch zwischen den Arbeitsspielen desselben Fahrers mussten Schwankungen um  $\pm 10\%$  bis  $\pm 20\%$  festgestellt werden [FSA12, S.17].

In der Dissertation von HOLLÄNDER werden dagegen lediglich zwei Fahrer bei der Bedienung eines Kettenbaggers miteinander verglichen. Bei sehr guten Bedienern limitiert das Hydrauliksystem die Umschlagsleistung. Entsprechend gering sind auch die Unterschiede von Zyklus zu Zyklus.

Während der bessere Fahrer zwischen den Einzelzyklen Schwankungen von  $\pm 1\%$  bis  $\pm 2\%$  bei Kraftstoffverbrauch und Umschlagsleistung zeigte [Hol98, S.82], wurden bei dem zweiten Baggerführer  $\pm 3\%$  beim Verbrauch und  $\pm 10\%$  bei der Umschlagsleistung verzeichnet [Hol98, S.102]. Zwischen den beiden als professionell betitelten Versuchsfahrern traten Leistungsunterschiede von bis zu  $30\%$  auf [Hol98, S.XI]. Des Weiteren ist der Versuchsauswertung zu entnehmen, wie stark bei den Fahrern die Bewegungsbahnen des Baggerarms voneinander abweichen. Außerdem musste HOLLÄNDER erfahren, dass ein enges Toleranzfeld bei den Arbeitsvorgaben wie Grabtiefe, Schwenkwinkel oder Hubhöhe die Fahrer verunsicherte und zu schlechteren Umschlagsleistungen führte [Hol98, S.148].

Auch STURMS Fahrversuche mit einem Mobilbagger und drei professionellen Fahrern ergaben Unterschiede im Kraftstoffbedarf von bis zu  $17\%$  bei einer um bis zu  $23\%$  geringeren Zykluszeit [Stu15, S.103]. Nach GIRMSCHIED kann alleine die Motivation des Fahrers bereits zu einer Leistungsreduktion von bis zu  $25\%$  führen [Gir10, S.5].

In den angeführten Beispielen wurde den Maschinenführern eine Arbeitsaufgabe mit einigen Rahmenbedingungen und Zielvorgaben erteilt. Bei der Festlegung der Systemgrenze wurden diese bewusst mit eingeschlossen (vgl. Kapitel 3). Dabei zeigt sich der enorme Einfluss des Fahrers auf Verbrauch und Produktivität einer mobilen Arbeitsmaschine. Je nach den persönlichen Fertigkeiten treten auch starke Unterschiede bei der Wiederholung der einzelnen Arbeitsspiele auf.

Im Bereich der Simulation wurde von STURM „*ein regelbasiertes Fahrermodell entwickelt, das die Bedienregeln des Baggervorgangs abbildet und für jedes zu bewertende Konzept einen individuellen, realistischen Arbeitszyklus generiert*“ [Stu15]. Diese Lösung ist im Bereich der Simulation sicher eleganter als die Vorgabe von Bewegungen oder Kräften über die Zeit. Eine Übertragung auf den realen Maschineneinsatz, was letztendlich einer autonom arbeitenden mobilen Arbeitsmaschine entspricht, ist zum aktuellen Zeitpunkt nur mit einem immensen Mehraufwand an Sensorik umsetzbar. Zudem wird die Geschwindigkeit und Arbeitsqualität eines realen Fahrers nur selten erreicht.

## 2.6 Forschungsbedarf

Produktivität und Energieeffizienz bestimmen maßgeblich die Gesamtkosten einer mobilen Arbeitsmaschine und sind deshalb die stärksten Argumente beim Verkauf an den Endkunden. Daneben sind für viele Kunden die Investitionskosten, die Lebensdauer sowie die Verfügbarkeit entscheidende Faktoren. [FSA12, S.1f.]

Betriebseigenschaften wie Produktivität oder Energieeffizienz lassen sich nur im realen Fahrzeugbetrieb zuverlässig ermitteln. Um eine Vergleichbarkeit derartiger Kennzahlen zu gewährleisten, empfiehlt sich eine Messung in einem Prüfprozess mit repräsentativen Arbeitszyklen. Mit den hieraus gewonnen charakteristischen Werten können verschiedene Modelle verglichen oder auch die Verbesserung beziehungsweise der Nutzen von Variantenentwicklungen nachgewiesen werden.

Eine mögliche Methode zur Bestimmung repräsentativer Arbeitszyklen findet sich bei SCHERER [Sch17a],[Sch17b]. In Kapitel 2.5 sind zudem einige Beispiele von Prüfprozessen für mobile Arbeitsmaschinen vorgestellt. Während bei den Leistungstests der Linde Material Handling GmbH, beim DLG PowerMix und bei SCHERER vor allem auf die Repräsentativität Wert gelegt wird, existieren bei den anderen genannten Prüfzyklen nur wenige Informationen zur Übertragbarkeit und Aussagekraft der Ergebnisse für die Endkundenanwendung.

Insbesondere Maschinenhersteller versprechen für ihre aktuelle Maschinengeneration Verbrauchsvorteile von bis zu 30%, ohne dabei auf den zugrunde liegenden Prüfprozess einzugehen [Bom17, S.14], [Ham10, S.2], [Mer16, S.8] [New13, S.5], [Vol13]. Im alltäglichen, realen Maschineneinsatz können diese Werte jedoch nur selten erreicht werden. Möglicherweise basieren diese Vergleichsangaben auf Einzelwerten ohne jegliche statistische Absicherung. Eventuell sind die angeführten Werte lediglich in einer Simulationen ermittelt oder wurden unter optimalen Bedingungen gemessen: Eine Maschine in hervorragendem Wartungszustand in Kombination mit einer perfekt abgestimmten Produktionskette und einem erfahrenen, hochkonzentrierten Bediener. Diese Bedingungen lassen sich jedoch nur über kurze Zeit aufrechterhalten. Arbeitsprozesse, Material, und Betriebs-

phasen sind nicht immer repräsentativ und können erheblich vom kundenindividuellen Maschineneinsatz abweichen. Die enormen Unterschiede beim Vergleich der Messergebnisse aus dem „*Liebherr-Normtests*“ mit den Daten des Flottenmanagementsystems LiDAT in Abbildung 2.6 stellt die Aussagekraft sowie Übertragbarkeit dieser Prüfergebnisse für den realen Betrieb eindeutig in Frage. Derartige Unterschiede sind auch für andere Prüfprozesse für mobile Arbeitsmaschinen vorstellbar. Während PKW-Fahrzyklen häufig zu geringe Lastanforderungen stellen, werden bei mobilen Arbeitsmaschinen eher der Teillastbetrieb und Leerlaufphasen unzureichend berücksichtigt.

Messgrößen und Messstellen der Untersuchungen sind in den meisten Fällen benannt. Dahingegen fehlen zumeist Angaben zur Genauigkeit der Messtechnik und eine Fehlerrechnung. Lediglich die drei Prüfprozesse der JCMA stellen hier klare Anforderungen [JCM07a] [JCM07b] [JCM07c]. Bei HOLLÄNDER sind die Messunsicherheiten zumindest teilweise angegeben [Hol98, S.216]. Für den DLG PowerMix lassen sich die Genauigkeiten des Messwagens zum Teil aus den Typenbezeichnungen der Messaufnehmer recherchieren [DF03, S.5]. Ohne Angaben zur Genauigkeit der Messtechnik sind Untersuchungsergebnisse unterschiedlicher Messaufnehmer jedoch nicht vergleichbar. Wird beispielsweise der Kraftstoffverbrauch über das Motorsteuergerät ermittelt, könnten Verbrauchsunterschiede in einem Fahrzeugvergleich ebenso gut auf die Messungenauigkeit statt auf den veränderten Einflussfaktor zurückzuführen sein. Gerade bei kleinen Unterschieden kann selbst hochwertige Messtechnik an ihre Grenzen kommen.

Informationen zur Streuung der Einzelversuche sind ebenfalls selten, die Darstellung uneinheitlich. Die Bedeutung der Streuung soll anhand des Beispiels in Abbildung 2.9 veranschaulicht werden. Die Wirkung einer Weiterentwicklung wird experimentell überprüft. Dazu werden jeweils fünf Einzelversuche vor und nach der Modifikation mit zwei unterschiedlichen Prüfprozessen durchgeführt. Die einzelnen Beobachtungswerte sind getrennt nach Prüfprozess in die Diagramme in Abbildung 2.9 eingetragen. Infolge der Streuung des Prüfprozesses überschneiden sich die Beobachtungswerte und lassen sich dadurch nicht klar voneinander trennen. Die

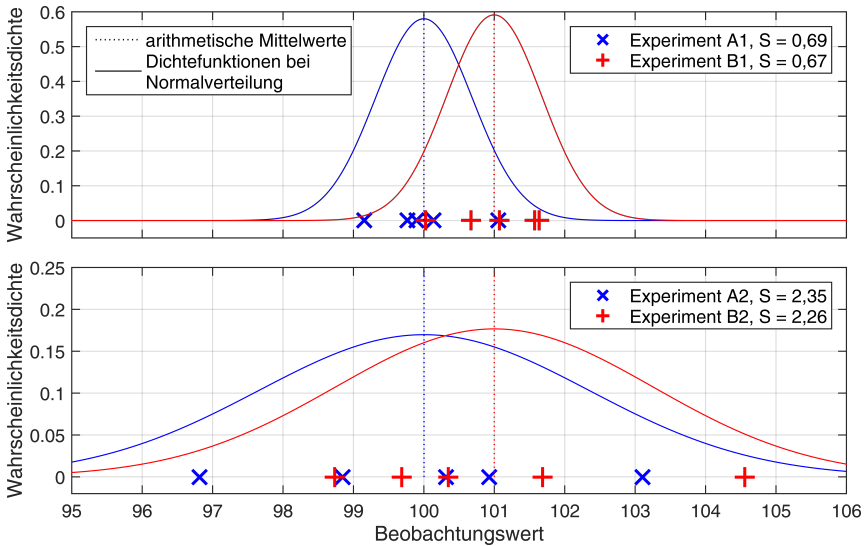


Bild 2.9: Bedeutung der Streuung bei einem Vergleich mit je fünf Einzelversuchen

Dichtefunktionen veranschaulichen die Streuung innerhalb eines Experiments unter Annahme einer Normalverteilung.

In beiden Prüfprozessen wird von A zu B eine Anhebung des arithmetischen Mittels um 1 % festgestellt. Ursache und Zuverlässigkeit dieser Ergebnisse klärt eine statistische Auswertung mithilfe der einfaktoriellem Varianzanalyse. Diese bestätigt für den ersten Prüfprozess bei einer Fehlerwahrscheinlichkeit von 4,9 %, dass die Weiterentwicklung tatsächlich eine signifikante Änderung bewirkt. Für den zweiten Prüfprozess liegt diese Fehlerwahrscheinlichkeit jedoch bei 51 %. Angesichts der großen Streuung bei der Untersuchung kann mit dem zweiten Prüfprozess kein signifikanter Unterschied festgestellt werden. Die Abweichung der Mittelwerte ergibt sich nicht zwangsläufig durch den Einfluss der Weiterentwicklung sondern ist wahrscheinlich der Streuung des Prüfprozesses geschuldet. Trotz gleicher Mittelwertdifferenz ergeben sich bedingt durch die Streuung der Beobachtungswerte ganz unterschiedliche Schlussfolgerungen.

Indessen zeigen das Ishikawa-Diagramm in Abbildung 2.1 sowie die Erläuterungen dieses Hauptkapitels, dass gerade bei mobilen Arbeitsmaschinen wesentlich mehr Einfluss- und Störgrößen als beispielsweise bei Fahrzeugen

des Personen und Güterverkehrs wirken. Die Beschreibung der Streuung bei den Untersuchungen ist demzufolge umso wichtiger.

DEITERS nutzt zur Darstellung der Streuung die empirische Verteilungsfunktion, Boxplots sowie den *relativen Variationskoeffizienten*. Letzterer ist im Gegensatz zum Variationskoeffizienten kaum verbreitet. Für die Zyklusdauer seiner 150 Einzelversuche kann bei DEITERS ein Variationskoeffizient von 6,7% hergeleitet werden. Die Dauer der Teilabschnitte erreicht Variationskoeffizienten von bis zu 17%. Dabei wurden die Messungen an nur einem Ort mit demselben Fahrer innerhalb weniger Stunden und damit unter nahezu konstanten Bedingungen durchgeführt. [Dei09, S.52f.]

Bei HOLLÄNDER wird die Streuung dagegen mithilfe der Spannweite der Zielgrößen angegeben. Je nach Fahrer und Zielgröße zeigen sich erhebliche Unterschiede bei den Erwartungswerten und der Streuung eines Experimentes. Obwohl HOLLÄNDER Fahrer, Maschineneinstellung und Arbeitsaufgabe in mehreren Stufen variiert hat, sind nur die Ergebnisse einer Arbeitsaufgabe mit konstanter Fahrzeugkonfiguration fahrerindividuell veröffentlicht. Die starke Streuung im realen Fahrversuch bewog DEITERS und HOLLÄNDER wie auch STURM und MELCHINGER ihre Untersuchungen in ein Simulationsmodell zu verlagern. [Hol98, S.65]

FRANK stellt Auszüge seiner Messergebnisse als Punktwolken in Punktdiagrammen dar [FSA12, S.10ff.]. Eine Ableitung statistischer Kenngrößen wäre nicht nur mühselig, sondern ist angesichts der fehlenden Achsenkalibrierung in diesem Fall unmöglich. Außerdem ist die Verteilungsdichte bei dieser Darstellung ab einer gewissen Punkteanzahl nicht mehr differenzierbar. Punktdiagramme eignen sich vielmehr zur Darstellung von Korrelationen und werden zu diesem Zweck auch bei DEITERS, HOLLÄNDER und STURM verwendet.

Über den DLG-PowerMix-Rollenprüfstand ist zu lesen, dass Wirkungsgradoptimierungen im Bereich kleiner 1% aufgelöst werden können [DLG17b]. Unklar bleibt bei dieser Angabe, ob damit die Wiederhol- oder die Vergleichspräzision gemeint ist und welcher Versuchsaufwand dafür erforderlich wäre. Mit dem Messwagen wurden die einzelnen Experimente bei der DLG bislang drei Mal wiederholt [FD05].

Zusammenfassend weisen alle aufgeführten Prüfprozesse unvollständige sowie in der Darstellung uneinheitliche Informationen bezüglich ihrer Qualität auf. Damit bleibt unklar,

- unter welchen Bedingungen die Untersuchungen durchgeführt wurden,
- ob sich die Ergebnisse auf den kundenindividuellen Fahrzeugeinsatz übertragen lassen,
- ob geeignete Messtechnik eingesetzt wurde,
- ob beobachtete Unterschiede tatsächlich signifikant sind,
- welche Unsicherheit dem Ergebnis innewohnt.

Demzufolge bedarf es einer einheitlichen Vorgehensweise die Qualität von Prüfprozessen für mobile Arbeitsmaschinen festzustellen und damit ihre Eignung für Untersuchungen nachzuweisen. Eine entsprechende Methode wird in Kapitel 4 vorgestellt und exemplarisch in Kapitel 5 angewendet. Prüfprozesse, deren Qualität in dieser Form dokumentiert wurde, eignen sich zur Anwendung in künftigen Untersuchungen oder als Basis für die Entwicklung weiterer Prüfprozesse. Des Weiteren unterstützt Kapitel 3 mit einigen grundlegenden Entscheidungsalternativen bei der Entwicklung eines zweckmäßigen Prüfprozesses.



### 3 Festlegen von Rahmenbedingungen und Versuchsablauf

Nach Definition der Prüfprozessanforderungen und des zu untersuchenden Arbeitsprozesses werden der Versuchsablauf und die Rahmenbedingungen der Experimente detaillierter ausgearbeitet. Dazu sollten aus einer Einfluss- und Störgrößenanalyse die Einflussfaktoren auf die Zielgröße bekannt sein. Einflussgrößen können gezielt variiert werden, um deren Wirkung genauer zu untersuchen, werden jedoch mit dem Ziel ihren Einfluss auszuschließen häufig konstant gehalten. Die Störgrößen, beispielsweise Umwelteinflüsse wie der Wind oder der intermittierende Betrieb eines Kühlgebläses, sind dagegen kaum zu steuern. Während sich technische Störgrößen häufig durch eine geschickte Versuchsführung oder Manipulation zu Einflussgrößen abwandeln lassen, können Umweltbedingungen lediglich gemessen und zulässige Wertebereiche definiert werden. Auf einen Teil der Umweltbedingungen kann durch die Wahl des Versuchsorts Einfluss genommen werden. So haben sich über die Jahre einige Regionen herstellerübergreifend für die Durchführung von Hitze-, Kälte- und Höherenerprobungen etabliert. Eine weitere Möglichkeit den Störgrößeneinfluss zu Gunsten der Prüfprozesspräzision zu reduzieren ist die Synthese einzelner Prozesselemente.

Bei der Konkretisierung des Prüfprozesses steht sowohl die realitätsnahe Abbildung des Arbeitsprozesses als auch die Reduzierung der Störgrößen im Fokus. Zur Unterstützung des Detaillierungsprozesses werden im Folgenden einige wesentliche Entscheidungsalternativen diskutiert:

- Versuchsablauf: stationär oder transient
- Systemgrenzen für die Untersuchung
  - genutzte Antriebsstränge

- Prozesskette
- Fahrer- und Bedienregeln
- Anpassung der Arbeitsumgebung
  - Untergrund
  - Umschlagmaterial
  - Abtrennung zur Umwelt
  - Pflanzen
- Randomisierung der Versuchsreihenfolge

### 3.1 Versuchsablauf: stationär oder transient

Mit der Auswahl des Arbeitsprozesses ist der Versuchsablauf längst nicht festgeschrieben. Wenn Dynamik und Schwingungen nicht im Vordergrund stehen, ist eine Abbildung des Arbeitsprozesses mit mehreren stationären Betriebspunkten aufgrund einiger beachtlicher Vorteile in Erwägung zu ziehen. Bei stationären Versuchen werden alle Fahrzeugeinstellungen und soweit möglich alle externen Einflüsse konstant gehalten. Sofern keine Resonanzen auftreten, klingen Systemschwingungen ab und es stellen sich gleichbleibende Systemgrößen ein. Trägheits- und Beschleunigungskräfte treten im Idealfall nicht auf und spielen deshalb keine oder zumindest eine untergeordnete Rolle. Ebenso wird der Bedienerinfluss auf ein Minimum reduziert. Diese Bedingungen vereinfachen die Versuchsauswertung erheblich, entsprechen allerdings nur bedingt der üblichen Nutzung der Maschine. Insbesondere bei realen Versuchen ist es außerdem nahezu unmöglich alle externen Einflüsse, wie Bodenbeschaffenheit und Wettereinflüsse konstant zu halten. Darüber hinaus sind die Aufgaben mobiler Arbeitsmaschinen häufig zyklischer Natur und nur teilweise kontinuierliche Prozesse.

Mit einer transienten Versuchsführung kann die alltägliche Arbeit einer mobilen Arbeitsmaschine in der Regel besser abgebildet werden. Durch dynamische Vorgänge wirken Beschleunigungs- und Trägheitskräfte, die zu Lastspitzen und Schwingungen führen. Selbst träge Systemgrößen wie

Temperaturen variieren über den Versuchsverlauf. Während des Versuchs ändern sich die Systemparameter, wodurch das System wesentlich umfangreicher geprüft wird. Bei der Versuchsauswertung wird die Zuordnung von Einflüssen erschwert. Werte und Ergebnisse unterliegen einer stärkeren Streuung.

## 3.2 Systemgrenzen für die Untersuchung

Weiter müssen die Systemgrenzen für die Untersuchung definiert werden. Bei einigen Untersuchungen, beispielsweise von Mobilbaggern, ist es denkbar sich auf einen der beiden Antriebsstränge zu beschränken. Dazu muss bekannt sein, ob Fahr- und Arbeitsantrieb im üblichen Arbeitsprozess gleichzeitig Leistung übertragen. Falls das zutrifft kann der Einfluss- und Störgrößenanalyse entnommen werden, inwiefern eine Wirkung auf die Zielgröße vorhanden ist. Kann ein Einfluss nicht ausgeschlossen werden, sollten beide Antriebsstränge bei der Untersuchung genutzt werden. Diverse mobile Arbeitsmaschinen erfüllen ihre Aufgaben innerhalb einer Prozesskette in enger Zusammenarbeit mit anderen Maschinen. Beispielsweise müssen die meisten Erntemaschinen das Erntegut permanent oder zyklisch auf ein Transportfahrzeug überladen. Ebenso muss ein Asphaltfertiger durchgehend mit Einbaumaterial beschickt werden. Tatsächlich macht es einen Unterschied, ob ein Bagger gezielt einen LKW belädt oder willkürlich auf eine Halde entleert. In den meisten Untersuchungen wird von einer optimalen Prozesskette ausgegangen und die Zusammenarbeit mit anderen Gliedern der Kette ausgeklammert. Dies kann etwa durch die Wahl des untersuchten Zeitfensters oder eine Anpassung der Prozesskette erfolgen und hat den Vorteil Einflüsse von anderen Maschinen oder Logistikprozessen auszuschließen. Andererseits ist dadurch die Aussagekraft des Untersuchungsergebnisses auf normale, suboptimale Rahmenbedingungen eventuell beschränkt. Manchmal ist speziell das Maschinenverhalten bei Wartezeiten oder die Zusammenarbeit mit anderen Maschinen der Prozesskette das maßgebliche Unterscheidungsmerkmal.

Noch ausschlaggebender ist die Entscheidung den Fahrer als Regler in die Untersuchung zu integrieren. Wird lediglich ein Anfangszustand und ein

Endzustand beschrieben, obliegt es diesem den Verlauf der Einflussgrößen individuell vorzugeben. In der Umsetzung einer solchen Aufgabe hat der Maschinenbediener viele Freiheiten, um die Arbeit auf die geschickteste Weise zu erledigen. Er reagiert auf das Maschinenverhalten, passt seine Bedienvorgaben an und wird so Teil der Beobachtung. Entsprechend groß stellt sich aber auch die Streuung zwischen einzelnen Versuchswiederholungen und verschiedenen Fahrern ein. Bei dieser Versuchsführung wird neben der Maschine das Zusammenspiel mit dem Fahrer betrachtet. Insbesondere die Bedienbarkeit und das Feinsteuerverhalten, welche beispielsweise durch Dosierbarkeit und Reaktionszeiten des Systems bestimmt werden, können so untersucht werden. Das System kann auf diese Weise unter realistischen Bedienvorgaben sein volles Leistungspotential entfalten.

Da diese starke Abhängigkeit vom Fahrer zur objektiven Fahrzeugbeurteilung häufig nicht erwünscht ist, wird bei einigen Fahrzeugversuchen ein fester Bewegungs- oder Geschwindigkeitsverlauf über die Zeit vorgegeben. Ein bekanntes Beispiel dafür ist der neue europäische Fahrzyklus (NEFZ) aus dem PKW Bereich. Die großen Verbrauchs- und Emissionsabweichungen zum realen Betrieb und die damit verbundene öffentliche Diskussion deuten auf einige Nachteile dieser Versuchsführung hin. Der vorgegebene Bewegungs- oder Geschwindigkeitsverlauf ist vielmehr eine Minimalanforderung und erlaubt es dem Fahrzeug nicht sein Leistungsvermögen zu nutzen. Dabei ist jedem klar, dass ein Sportwagen nicht wie ein Kleinwagen betrieben wird. Weil bei mobilen Arbeitsmaschinen die Produktivität im Vordergrund steht, ist eine derartige Vorgabe in den seltensten Fällen geeignet.

Diesem Punkt wird beim Traktorenvergleichstest DLG PowerMix besser Rechnung getragen. Mithilfe eines Prüfstandes, dem DLG Zugleistungsmesswagen ergibt sich die Möglichkeit einen Momenten-/Kräfteverlauf sowohl auf den Fahr- als auch auf den Arbeitsantriebsstrang des Prüflings aufzuprägen. Die Last wird dabei auf die Maximalleistung des Motors skaliert [Neu03]. In den Versuchen fährt der Traktor mit dem Tempomat, wobei in den meisten Versuchen der Kräfteverlauf so hoch ist, dass das Fahrzeug überwiegend unter die voreingestellte Geschwindigkeit gedrückt wird. Je nach Drückung des Fahrzeuges ergeben sich dabei über die Versuchszeit

unterschiedliche Fahrstrecken und damit eine Abweichung in der geleisteten Arbeit. Hier ist eine Vorgabe der Kräfte und Momente über den Weg statt über die Zeit zweckmäßiger. Mit einer strecken- oder ereignisbasierten Versuchsführung wird eine höhere Praxisrelevanz erwartet.

Selbst wenn bei der Festlegung der Systemgrenzen der Fahrer von der Untersuchung ausgeschlossen wurde, muss die Maschine bedient werden. Sobald der Fahrer eine regelnde Funktion übernimmt, ergibt sich eine Bedienerabhängigkeit. Wird der Maschinenführer als Regelglied zum Nachfahren eines festen Werteverlaufs eingesetzt, ist sein Einfluss beschränkt. Gleiches gilt jedoch in diesem Fall häufig auch für die Praxisrelevanz, weil die Maschine nicht entsprechend ihres Leistungsvermögen betrieben wird. Falls der Versuch fahrerunabhängig bleiben soll, darf dieser nur steuernd auf die Maschine wirken. Dies kann mithilfe klarer, einfacher Verhaltensregeln sowie mit Anschlägen und festen Steuerpunkten realisiert werden. Kann mittels eines Fahrroboters oder elektrischer Schnittstellen eine automatisierte Übertragung der Bedienervorgaben realisiert werden, ist die Simulation des Fahrerverhaltens möglich. Da sich das Fahrermodell reproduzierbar verhalten wird, ist in einem solchen Fall eine Regelung unproblematisch. Je nach Maschine und Arbeitsprozess ist für die automatisierte Bedienung jedoch ein immenser Mehraufwand an Sensorik erforderlich.

### **3.3 Anpassung der Arbeitsumgebung**

Durch die Verlagerung des Prozesses in eine andere Arbeitsumgebung können einige Rahmenbedingungen eventuell besser gesteuert werden. Dafür muss jedoch häufig auch der Arbeitsprozess angepasst werden.

Sofern die Leistungsübertragung zwischen Fahrbahn und Boden nicht im Untersuchungsfokus steht, ist die Verlagerung von Fahrbewegungen auf unbefestigten, nachgiebigen Böden auf einen befestigten Untergrund wie Beton oder Asphalt ein wesentlicher Schritt zur Steigerung der Präzision. Durch diese Anpassung der Rahmenbedingungen ergibt sich jedoch ein wesentlich geringerer Fahrwiderstand. Außerdem kann eine landwirtschaftliche Zugmaschine beispielsweise keine Bodenbearbeitung mehr durchführen. Beide Unterschiede zum ursprünglichen Prozess können in einer Simu-

lation berechnet und mithilfe einer Belastungseinheit dem Untersuchungsfahrzeug aufgeprägt werden. Dank Modellierung und Simulation des Fahrwiderstandsanteils des Bodens und der Kräfte aus dem Arbeitsprozess werden diese Einflüsse reproduzierbar. Modellgüte und Charakteristik der Belastungseinheit bestimmen die Qualität des Ersatzsystems und damit die Praxisrelevanz.

Bei Arbeitsprozessen mit Materialumschlag wird empfohlen, das ursprüngliche Umschlagsmaterial durch standardisiertes, sortiertes Material wie Schotter, Split, Kies oder Sand zu ersetzen. Diese Stoffe haben eine definierte Zusammensetzung und Korngrößenverteilung, so dass unabhängig vom Hersteller mit vergleichbaren Eigenschaften gerechnet werden kann. FRANK beobachtete allerdings bei wiederholter Nutzung desselben Materials eine Änderung der Korngrößenverteilung [FSA12, S.7f.] und damit der Dichte und des Eindringwiderstandes. Folglich sollte das Material in regelmäßigen Abständen erneuert werden.

Daneben ändert sich auch bei diesen Stoffen wetterbedingt der Wassergehalt und damit relevante Eigenschaften. Die Verlagerung der Untersuchungen in einen überdachten Bereich kann eine Veränderung der Materialfeuchte begrenzen. In einem geschlossenen Raum lassen sich zusätzlich Windeinflüsse ausschließen und die Umgebungstemperatur einstellen. Sehr großer Platzbedarf kann gegebenenfalls mit einem Rollen- oder Nabenprüfstand auf ein umsetzbares Maß reduziert werden. Bei kombinierter Belastung des Fahr- und Arbeitsantriebes ist in diesem Fall jedoch ein Ersatzsystem für den Arbeitsprozess erforderlich.

Wie beim Materialumschlag ist bei der Verarbeitung von Pflanzen die Verwendung einheitlicher Ersatzstoffe anzustreben. Aufgrund der Menge, Form und Größe ist dies in den meisten Fällen höchst unwirtschaftlich. Stattdessen sollten bei großen Pflanzen, wie bei Forstarbeiten, die charakteristischen Eigenschaften eines Baumes und deren Toleranzbereiche festgelegt und erfasst werden. Gleichzeitig sind durch eine erhöhte Versuchsanzahl die Einflüsse durch die Schwankungen der Eigenschaften zu glätten.

Bei Feldfrüchten ist die Erfassung der individuellen Pflanzencharakteristika nicht durchführbar. Für diese Fälle existiert der Ansatz einen Vergleich

zu einer Referenzmaschine anstelle einer absoluten Messung durchzuführen. Diese Vorgehensweise ist von Mähdreschern bekannt [Hol98, S.148ff.]. Prüfling und Referenzmaschine führen den Arbeitsprozess im Wechsel auf demselben Feld aus. Durch den ständigen Wechsel und die Größe des Feldes werden die Inhomogenitäten des Bewuchses weitestgehend ausgeglichen. Wichtig ist bei diesem Konzept die Referenzmaschine, insbesondere die Verschleißkomponenten, in technisch einwandfreiem Zustand zu halten. Des Weiteren ist zu klären, inwieweit die Leistungen typgleicher Maschinen voneinander abweichen. Dieser Ansatz kann für einen Überblick des aktuellen Marktes oder für den Vergleich zweier Maschinengenerationen eingesetzt werden. Mangels eines absoluten Ergebnisses und der Alterung der Referenzmaschine ist dies tendenziell eine lokale und kurzfristige Lösung.

Kommt es auf Basis dieser Überlegungen zu Anpassungen des Versuchs zugunsten der Präzision, so ist zu überprüfen, ob durch Quereinflüsse auch anderer Kriterien betroffen sind. Außerdem ist zu untersuchen, inwieweit die Aussagekraft für den normalen Fahrzeugeinsatz beeinträchtigt wird. Eventuell besteht die Möglichkeit auftretende Abweichungen beispielsweise mithilfe eines mathematischen Modells oder in Kombination mit einem Prüfstand weitestgehend zu kompensieren. Eine Lösungsmöglichkeit zeigt das Testzentrum der Deutschen Landwirtschafts-Gesellschaft mit ihrem PowerMix Zugkraftmesswagen [Neu03]. Abbildung 2.4 zeigt diesen Prüfstand in Form eines LKWs, der dem Versuchsfahrzeug sowohl Zugkräfte als auch Zapfwellenleistung und hydraulische Leistung abfordern kann. Gerade bei Prüfständen sollte allerdings bedacht werden, ob die erforderliche Dynamik erreicht werden kann und ob das Schwingungsverhalten bei den Versuchen eine Rolle spielt.

### 3.4 Randomisierung der Versuchsreihenfolge

Beim Aufstellen des Versuchsplans sollten unvorhergesehene und zeitliche Einflüsse durch eine Randomisierung der Versuchsreihenfolge berücksichtigt werden. Auf diese Weise werden Störeinflüsse bewusst in die Streuung der Ergebnisse aufgenommen und nicht irrtümlich als Effekte interpretiert.

tiert. Des Weiteren wird in Kapitel 4 gezeigt, dass über die Anzahl an Wiederholungen die Untersuchungsgenauigkeit deutlich gesteigert werden kann. Zusätzlich ist es sinnvoll über die gesamte Versuchszeit einen einfachen Referenzversuch regelmäßig zu wiederholen. Dieser kann auch als Vorkonditionierung vor allen anderen Versuchen genutzt werden und damit ebenfalls die Streuung reduzieren. Anhand dieser Referenzmessungen lassen sich im Nachhinein Veränderungen der Versuchsbedingungen besser nachvollziehen. Zudem lässt sich mit diesen Referenzversuchen die Präzision sowohl innerhalb einer Versuchsreihe, als auch über längere Zeiträume genauer abschätzen. Mithilfe einer Prüfprozessanalyse der Referenzversuche kann die Prüfprozesseignung der Messkampagne bestätigt werden. Zusammenfassend sollte während der Versuchsplanung durchgehend die Streuung der Versuchsergebnisse, die Präzision, berücksichtigt werden. Je kleiner Anzahl und Ausprägung der Störgrößen gehalten werden kann, desto zuverlässiger und feiner können Unterschiede erfasst werden. Die Reduktion der Störgrößen steht häufig im Konflikt mit dem Wunsch nach einer hohen Praxisnähe. Die Herausforderung besteht somit darin unter möglichst praxisrelevanten Rahmenbedingungen ein präzises Untersuchungsergebnis zu generieren.



## 4 Beurteilung der Prüfprozesseignung

Nachdem unter Berücksichtigung der besonderen Eigenschaften einer mobilen Arbeitsmaschine ein Prüfprozess definiert wurde, sollte dessen Eignung zur Untersuchung der Zielgröße sichergestellt werden. Die Qualitätsanforderungen an den Prüfprozess sollten bereits zu Beginn dessen Entwicklung festgelegt worden sein und lassen sich den Bereichen **Aussagekraft**, **Genauigkeit** und **Präzision** zuordnen.

Unter dem Kriterium Aussagekraft sind zahlreiche Aspekte zusammengefasst, um Übertragbarkeit und Gültigkeit der Prüfergebnisse für den kundenindividuellen Fahrzeugeinsatz zu beurteilen. Aufgrund des häufig vielseitigen Aufgabenspektrums einer Fahrzeugart sowie der großen Varianz an Nutzungs- und Betriebsbedingungen ist eine allgemeingültige Quantisierung der Aussagekraft kaum möglich. Deshalb sollten den Adressaten einer Veröffentlichung Informationen zur Entstehung und Durchführung des Prüfprozesses zur Verfügung gestellt werden. Beim Vergleich des im Prüfprozess repräsentierten Einsatzprofil mit den eigenen, spezifischen Bedingungen kann sich ein Fachmann selbst ein Urteil zur Aussagekraft der Ergebnisse für seine Anwendung bilden. Adressaten mit weniger Praxisbezug sind dagegen über die Schilderung eines methodischen sowie plausiblen Vorgehens bei der Entwicklung des Prüfprozesses zu überzeugen.

Genauigkeit und Präzision der Prüfergebnisse lassen sich demgegenüber weit besser gegen die Anforderungen prüfen. Die mögliche Abweichung der gemessenen Werte von den tatsächlichen Werten wird mit der Genauigkeit angegeben. Diese ist vor allem bei der Angabe absoluter Ergebnisse und dem Vergleich der Messwerte mehrerer Sensoren, beispielsweise in verschiedenen Fahrzeugen relevant. Ursachen für Diskrepanzen sind Kalibrierungsfehler, systematische Abweichungen wie Nichtlinearitäten und

zufällige Abweichungen. Die Ermittlung der Genauigkeit sowie die Kalibrierung der Sensoren erfolgt an Stützstellen mithilfe von Normalen oder durch eine parallele Messung mit einem geeichten Messsystem. Da die Kalibrierung im Fahrzeug nicht immer möglich ist, muss diese teilweise vor dem Einbau und mit großem zeitlichen Abstand erfolgen.

Mindestens genauso wichtig ist eine geringe Streuung bei der Messung, die Präzision. Die Varianz der Messwerte wird durch den Einfluss der Störgrößen hervorgerufen. Folglich ist eine hohe Präzision mit einem geringen Störgrößeneinfluss gleichzusetzen. Diese Eigenschaft des Prüfprozesses lässt sich weiter in Wiederholpräzision und Vergleichspräzision untergliedern. Wird bei mehreren Versuchsreihen keiner der Einflussfaktoren gezielt geändert, zeigt die Streuung der Messwerte die Wiederholpräzision des Prüfprozesses. Zur Überprüfung der Reproduzierbarkeit werden Einflussgrößen wie Maschine, Bediener, Messsystem oder Ort variiert, um deren Einfluss auf das Ergebnis und damit auf die Varianz des Prüfprozesses zu ermitteln. Daraus ergibt sich die Vergleichspräzision in Abhängigkeit der jeweiligen Einflussgrößen.

Ohne Informationen zu diesen beiden Kriterien können bei der Auswertung der Versuchsergebnisse Abweichungen der Ergebnisse nicht beurteilt werden. Diese könnten von der gezielten Variation einer Einflussgröße herühren, könnten aber ebenso gut durch Messrauschen oder eine unbemerkte Veränderung der Rahmenbedingungen verursacht worden sein. Deshalb sollte der Experimentator so früh wie möglich die Prüfprozesseignung für sein Untersuchungsziel sicherstellen.

Abbildung 4.1 zeigt den Ablauf zur Beurteilung der Eignung eines Prüfprozesses in einem Flussdiagramm. Die Rauten stellen die wesentlichen Entscheidungskriterien im Beurteilungsprozess dar. Können diese Kriterien nicht erfüllt werden, zeigen die roten Rechtecke die direkten Konsequenzen, die letztendlich zu einer negativen Bewertung des gesamten Prüfprozesses führen. Mögliche Abhilfemaßnahmen bei der Entwicklung sind orange hinterlegt. Sind alle Qualitätsmerkmale erfüllt, ist die grundsätzliche Eignung des analysierten Prüfprozesses nachgewiesen. Insbesondere die Präzision sollte jedoch in jeder Untersuchung bestätigt werden.

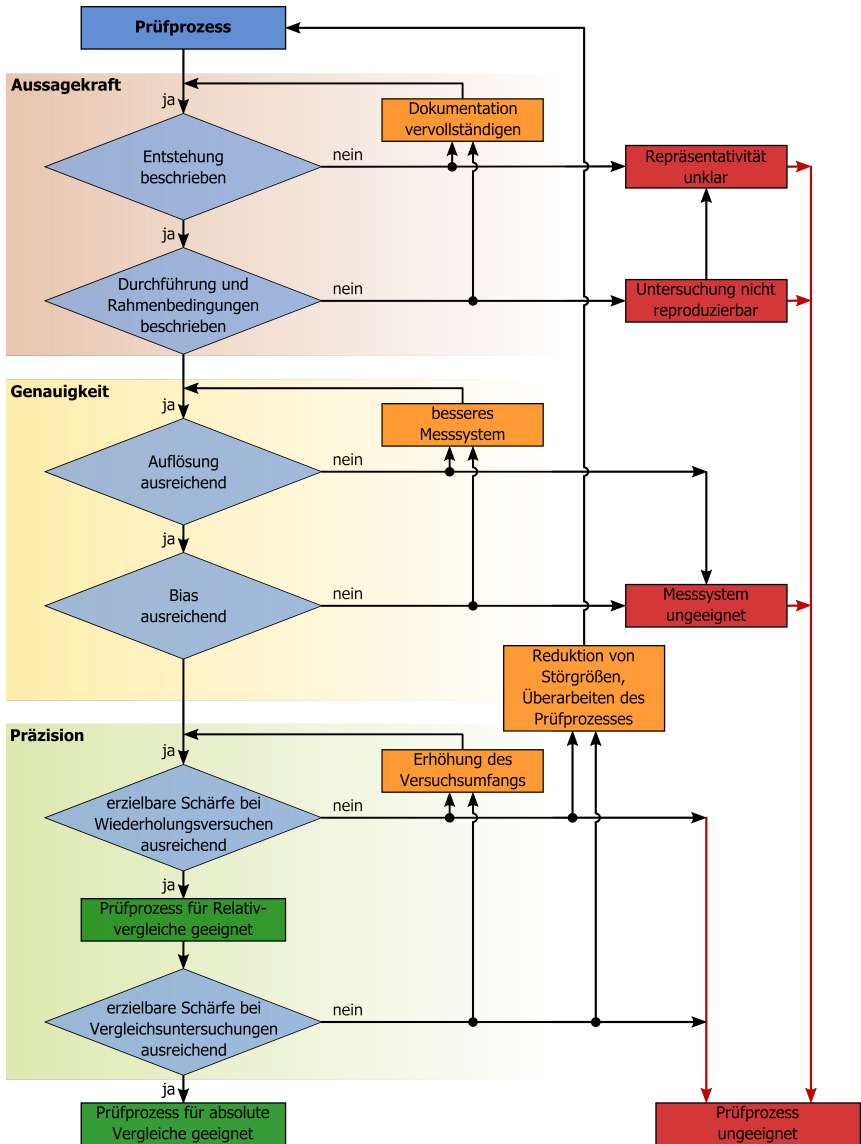


Bild 4.1: Ablauf zur Beurteilung der Prüfprozesseignung

## 4.1 Aussagekraft des Prüfprozesses

Die zahlreichen Nutzungsmöglichkeiten, Einsatz- und Betriebsbedingungen einer mobilen Arbeitsmaschine lassen sich in einem Prüfzyklus nur bruchstückhaft abbilden. Zusätzlich müssen zur Begrenzung der Streuungen im Prüfprozess häufig Vereinfachungen getroffen und Schnittstellen substituiert werden. Folglich stellt sich für den Interessenten die Frage, inwieweit sich Ergebnisse des Prüfprozesses auf seine reale, individuelle Fahrzeugnutzung übertragen lassen.

Diese Beurteilung kann nur vom Interessenten selbst geleistet werden. Urheber und Anwender des Prüfprozesses müssen dafür jedoch alle relevanten Informationen zur Verfügung stellen. Gleichzeitig sollte der Urheber bei der Entwicklung des Prüfprozesses eine hohe Übereinstimmung mit den gängigsten Einsatzprofilen des Untersuchungsfahrzeuges sicherstellen. Ein Verfahren zum Bestimmen und Clustern dieser Einsatzprofile ist bei SCHERER zu finden [Sch17a], [Sch17b]. Unabhängig von der Vorgehensweise ermöglicht die Beschreibung derselben eine maßgeblich fundierteres Urteil über die Aussagekraft des Prüfprozesses.

Tabelle 4.1: Informationen zur Entstehung des Prüfprozesses

- Vorgehensweise zur Entwicklung des Prüfprozesses
- Einsatzprofile
  - Umfang der Datenerhebung beträchtete Fahrzeugmodelle, Anzahl Betriebe, Anzahl Fahrer, Dauer, Saison, ...
  - Einsatzbetrieb Charakterisierung der Einsatzbetriebe
  - Arbeitsaufgaben Zeitanteile der verschiedenen Arbeitsaufgaben in den Einsatzprofilen des Fahrzeuges
  - Betriebsphasen Zeitanteile einzelner Prozessschritte, beispielsweise Standzeiten, Fahranteile, Arbeitsanteile, kombinierte Anteile, charakteristische Teilschritte, ...

Falls die Angaben zu den Einsatzprofilen nicht messtechnisch erfasst wurden, sollen diese abgeschätzt und dementsprechend als Schätzung gekennzeichnet werden. Eine Beurteilung der Übertragbarkeit der Prüfprozesseergebnisse auf den individuellen Anwendungsfall des Interessenten ermöglichen Informationen zu folgenden Aspekten:

Tabelle 4.2: Informationen zur Durchführung einer Untersuchung

- Versuchsträger
  - Fahrzeug                      Fahrzeugklasse, Herstellerbezeichnung
  - Fahrzeuggröße                charakteristische Größen des Fahrzeuges wie Motorleistung, Gewicht, Nutzlast, Schaufelgröße, ...
- Fahrer                              Erfahrung, Fertigkeit, Routine mit dem Versuchsfahrzeug
- Vorkonditionierung
  - Fahrzeug                      Fahrzeugzustand, vorbereitende Arbeitsprozesse, Fahrzeugeinstellungen, ...
  - Umgebung                      Fahrbahn, Untergrund, Umschlagsmaterial, Pflanzenbestand, ...
- Prüfprozess
  - Zielgröße                      Messgrößen, Messpunkte, Messverfahren
  - Arbeitsaufgabe                Beschreibung der Arbeitsaufgabe
  - Vereinfachungen              Vereinfachungen von Arbeitsprozessschritten, Substitution von Kontaktstellen, ...
  - Lastniveau                      Belastungsgrad von Fahrzeug und Bediener, limitierender Faktor
  - Betriebsphasen                Zeitanteile einzelner Prozessschritte
  - Durchführung                Anzahl der Versuchswiederholungen, Vorgaben für den Fahrer

- Qualität Anforderungen an die Qualität der Arbeitsergebnisse
- Umweltbedingungen (ggf. als Wertebereich)
  - Wetter Temperatur, Luftdruck, Wind, Niederschlag, ...
  - Kontaktstellen Fahrbahn, Untergrund, Umschlagsmaterial, Pflanzenbestand, ...

Tabelle 4.2 dient in dieser Form der Dokumentation einer konkreten Untersuchung. Auf diese Angaben beziehen sich die experimentell ermittelten Werte zur Präzision. Durch die Definition von zulässigen Wertebereichen eignet sich diese Aufstellung jedoch ebenso für die Beschreibung eines Prüfprozesses.

Eine weitere Maßnahme zur Verbesserung der Aussagekraft zeigt der Traktorentest DLG PowerMix. Zur Abdeckung eines möglichst breiten Nutzungsspektrums werden hier sieben verschiedene dynamische Arbeitsprozesse unter zumeist zwei Lastniveaus geprüft. Dank Veröffentlichung der Einzelergebnisse können sich Interessenten ein individuell gewichtetes Gesamtergebnis berechnen. Die Vorgehensweise zur Entwicklung der einzelnen Prüfzyklen ist allerdings auch hier nur grob skizziert [DF03]. Ebenso sind Vorkonditionierung, Umweltbedingungen und Versuchsumfang nicht mit den Testergebnissen dokumentiert.

## 4.2 Unsicherheit von Messergebnissen

Jedes Messergebnis ist mit einer gewissen Messunsicherheit belegt. Diese Unsicherheit setzt sich aus systematischen und zufälligen Messabweichungen zusammen. Ein Teil der systematischen Abweichungen lässt sich beispielsweise durch eine Justierung korrigieren. Weil der wahre Wert des Ergebnisses nicht bekannt ist und nicht alle Störeinflüsse ausgeschlossen oder messtechnisch erfasst werden können, muss die verbleibende Messabweichung aus Versuchen oder Annahmen abgeschätzt werden. Die Größe dieser Messabweichungen beschränkt die Eignung des Prüfprozesses für eine Messaufgabe. [Die14, S.24ff.]

Aus dem Bereich der Qualitätssicherung in der Fertigung existieren einige Normen, Leitfäden und Richtlinien, die eine Vorgehensweise für den Eignungsnachweis eines Prüfprozesses aufzeigen. Die „Prüfmittelfähigkeit und Messunsicherheit im aktuellen Normenumfeld“ wird von DIETRICH in seinem Buch „Prüfprozesseignung“ ausführlich diskutiert [Die14]. Unter Prüfprozess ist in diesem Umfeld die Vermessung oder die attributive Überprüfung der Geometrie von Teilen aus der Produktion zu verstehen. Die Einhaltung der geforderten geometrischen Toleranzen kann nur dann sichergestellt werden, wenn auch die Streuung des Prüfprozesses bekannt ist. Diese wird experimentell mittels mehrfacher Vermessung von Produktionsteilen durch mehrere Prüfern und unter verschiedenen Umweltbedingungen ermittelt. Zudem wird die Genauigkeit im Toleranzbereich der Teile mit Normalen untersucht, die sich auf nationale oder internationale Normale zurückführen lassen. In Summe sollte die Standardabweichung des Prüfprozesses maximal  $\frac{3}{80}$  der geforderten Bauteiltoleranz erreichen [Die14, S.56ff.]. Je nach Richtlinie und Rahmenbedingungen wird dieser Schwellenwert bis auf ein Sechzigstel herabgesetzt. Die wichtigsten Dokumente zu diesem Thema sind die *Measurement System Analysis* im amerikanischen Raum [A.I10] und der *VDA Band 5* [VDA11] beziehungsweise die ISO 22514-7 im europäischen Markt [ISO12]. Letztere leiten sich aus dem *Guide to the expression of uncertainty in measurement* (GUM) ab [DIN99]. Da diese Regelwerke durchweg auf Prüfprozesse für die Kontrolle von geometrischen Eigenschaften von Produktionsteilen gegen eine Fertigungstoleranz ausgerichtet sind, ist eine direkte Übertragbarkeit auf Gesamtfahrzeugversuche nicht möglich. Bei Fahrzeuguntersuchungen wird vorwiegend das Ziel verfolgt mit dem Prüfprozess eine signifikante Abweichung festzustellen, während bei der Kontrolle des Produktionsergebnisses die Einhaltung einer Toleranz überprüft wird. Die Zielgröße bei Fahrzeugversuchen beschreibt häufig Kräfte oder die Bewegungen des Fahrzeugs, die Dynamik. Teilweise sind die Versuchsergebnisse keine Einzelwerte, sondern ein charakteristischer Werteverlauf. Da es sich zumeist um abgeleitete Einheiten handelt, existieren außerdem in der Regel keine Normale wie bei geometrischen Größen. Zudem ist die Mehrzahl der Prüfprozesse, bei denen das Versuchssystem mit der Umgebung interagiert, nicht zerstörungs-

frei. Gerade diese Interaktion ist jedoch eine wesentliche Charakteristik im Arbeitsprozess einer mobilen Arbeitsmaschine. Selbst eine einfache Fahrt über einen nachgiebigen Untergrund führt zu irreversiblen Veränderungen des Bodens und damit einer Abwandlung der Bedingungen. Ebenso beeinflusst der Prüfprozess auch das Untersuchungsobjekt. Mit jeder Versuchsdurchführung ist ein Wärmeeintrag in das System verknüpft, der sich auf die Eigenschaften der Maschine auswirkt. Das Erreichen eines Zustandes thermischer Beharrung kann mehrere Stunden in Anspruch nehmen, was für die meisten Untersuchungen einen unverhältnismäßig großen Aufwand bedeutet.

Aus diesen Gründen ist für die Beurteilung der Prüfprozesseignung von Fahrzeuguntersuchungen mit mobilen Arbeitsmaschinen ein abweichendes Vorgehen erforderlich.

### 4.2.1 Genauigkeit des Prüfprozesses

Unter dem Kriterium Genauigkeit werden die technischen Daten des Messsystems zur Erfassung der Zielgröße hinsichtlich ihrer Eignung für den Prüfprozess betrachtet. Bei der Kontrolle von Fertigungsteilen darf laut DIETRICH die Auflösung der Messtechnik „höchstens 5 % der Toleranz betragen.“ Bei Erfüllung dieser Forderung „ist in der Regel der Anteil der Messabweichung aufgrund der Auflösung vernachlässigbar klein“ [Die14, S.47].

Für die Reproduzierbarkeit von Untersuchungsergebnissen oder bei der Verwendung mehrerer Messsysteme ist die absolute Genauigkeit besonders wichtig. Diese setzt sich aus systematischer Messabweichung vom Normal  $|B_i|$  und Standardabweichung der Messung  $S_M$  zusammen. In den Sensordatenblättern findet sich allerdings häufig nur eine Genauigkeitsklasse oder die maximale Abweichung. Diese sind als Absolutwert für das Bias  $|B_i|$  einzusetzen. Streuungsangaben sind dagegen selten. Bei Einzelbeispielen einer Kraftmessdose und eines Drucksensor ist die Standardabweichung jeweils etwa eine Größenordnung kleiner als das Bias. Wenn für die Messgröße Normale oder ein Vergleichsmesssystem mit bekannter Abweichung zur Verfügung stehen, kann das Bias auch experimentell in Wiederholungsversuchen bestimmt werden. Die Standardabweichung des Messsystems wird



als unbekannter Bruchteil der Streuung des Prüfprozesses berücksichtigt. Letztere wird grundsätzlich experimentell ermittelt.

Nach den Erläuterungen von DIETRICH sollten Bias und Standardabweichung des Messsystems entsprechend Gleichung 4.1 begrenzt sein. Alternativ wird noch die Forderung „ $|Bi| \leq 10\%$  der Toleranz“ als Faustregel genannt. [Die14, S.63f.]

$$T \geq 40 \cdot S_M + 10 \cdot |Bi| \quad (4.1)$$

Die Anforderung „ $|Bi| \leq 10\%$  der Toleranz“ lässt sich nur für  $S_M \approx 0$  mit Gleichung 4.1 in Einklang bringen. Bei Gesamtfahrzeuguntersuchungen wird die Messtechnik im Regelfall fest eingebaut und automatisiert ausgelesen. Gleichzeitig wirken beispielsweise größere Temperaturschwankungen und damit Störeinflüsse auf die Sensorik. Weil Angaben zur Sensorstreuung häufig nicht verfügbar sind, soll hier den beiden Summanden jeweils die Hälfte der Toleranz zugestanden werden. Daraus ergibt sich:

$$\frac{1}{2} \cdot T \geq 10 \cdot |Bi| \quad (4.2)$$

Die verwendete Bezugsgröße „Toleranz“ spielt bei Untersuchungen mit Gesamtfahrzeugen allerdings keine Rolle. Stattdessen interessiert hier die kleinste feststellbare Abweichung  $\Delta Y_{min}$ . Dies ist eine der grundlegenden Anforderungen an den Prüfprozess. Um die Empfehlungen aus der geometrischen Überprüfung von Fertigungsprodukten auf Gesamtfahrzeuguntersuchungen zu übertragen, ist ein Bezug zwischen der Toleranz  $T$  und  $\Delta Y_{min}$  herzustellen.

In Kapitel 4.4.3 wird dargelegt wie mithilfe der Varianzanalyse die kleinste feststellbare Abweichung, die erzielbare Schärfe, abgeschätzt werden kann. Basis dieser Abschätzung sind die Versuchsanzahl sowie die Streuung des Prüfprozesses. Für das Messsystem ergibt sich daraus die Anforderung, diese Prüfprozessstreuung hinreichend genau zu erfassen. Eine vergleichbare Forderung existiert bei der Überprüfung von Fertigungsprodukten für die Streuung des Fertigungsprozesses. Eine Fertigung kann nur kostengünstig gesteuert werden, wenn die Produktmerkmale und deren prozessbedingte Streuung überprüft werden können [Die14, S.8]. Für einen zuverlässigen

Fertigungsprozess sollte die Prozessstreuung  $\sigma_P$  höchstens  $\frac{1}{6}$  der Toleranz erreichen. Bei einer normalverteilten Streuung sind so mindestens 99,7% der Produkte innerhalb der Spezifikation. Demnach kann für die Toleranz  $T$  auch die Prozessstrebereite  $6\sigma$  eingesetzt werden [Die14, S.58].  $\sigma$  steht für die Streuung der Grundgesamtheit und wird mithilfe der Standardabweichung von Stichproben  $S$  angenähert (vgl. Kapitel 4.3.2).

$$T = 6\sigma \approx 6S \quad (4.3)$$

Die Beziehung zwischen  $S$  und  $\Delta Y_{min}$  kann nach Umstellung von Gleichung 4.81 berechnet werden.

$$S = \sqrt{\frac{n}{2 \cdot F_{1-\alpha}(1, 2(n-1))}} \cdot \Delta Y_{min} \quad (4.4)$$

Mit der Empfehlung Stichproben mit mindestens  $n = 25$  Einzelversuche pro Experiment zu generieren und dem höchsten verwendeten Signifikanzniveau  $\alpha = 0,1\%$  ergibt sich der Faktor 1,0087. Gleichzeitig wird innerhalb der Auswertung mit Versuchsreihen mit jeweils lediglich  $n = 5$  Einzelversuchen und einem Signifikanzniveau von  $\alpha = 5\%$  gearbeitet. Gilt auf dieser Ebene die gleiche Anforderung an die Schärfe, liefert Gleichung 4.4:

$$S = 0,6857 \cdot \Delta Y_{min} \quad (4.5)$$

Aus den Gleichungen 4.3 und 4.5 folgt:

$$T \approx 6 \cdot 0,6857 \cdot \Delta Y_{min} \approx 4\Delta Y_{min} \quad (4.6)$$

Mit dieser Gleichung lassen sich nun die Empfehlungen bezüglich Auflösung und Genauigkeit aus der geometrischen Teileprüfung auf Gesamtfahrzeuguntersuchungen übertragen. Demnach sollten sowohl die Auflösung als auch die maximale Abweichung der Messtechnik höchstens 20% beziehungsweise ein Fünftel der kleinsten festzustellenden Abweichung  $\Delta Y_{min}$  betragen. Diese Empfehlung ist etwas konservativer als DIETRICH'S Vorschlag einzustufen und gilt für Varianzanalysen mit einem Stichprobenumfang von mindestens  $n = 25$  Einzelversuche pro Experiment.

Berechnet sich die Zielgröße als Funktionswert aus mehreren einzelnen Messgrößen  $f(x, y, z, \dots)$ , summieren sich auch deren Unsicherheiten. Der größtmögliche Fehler wird mit der linearen Fehlerfortpflanzung berechnet. Gleichung 4.7 gilt unter der Annahme, dass die Abweichung weit kleiner als die Messwerte sind. Die Ermittlung des Größtfehler wird empfohlen, falls abhängige Messgrößen verwendet werden oder sich die Messunsicherheiten stark unterscheiden. Außerdem sollte das Gesamtbias mit der linearen Fehlerfortpflanzung errechnet werden. Für  $\Delta x, \Delta y, \Delta z$  wird dabei die maximale Abweichung eingesetzt. Bei unterschiedlichen Einheiten ist mit relativen Fehlergrenzen zu rechnen. [Tes17, S.42f.]

$$\Delta f = \left| \frac{\partial f}{\partial x} \right| \cdot \Delta x + \left| \frac{\partial f}{\partial y} \right| \cdot \Delta y + \left| \frac{\partial f}{\partial z} \right| \cdot \Delta z + \dots \quad (4.7)$$

Die Standardabweichung  $\sigma_f$  einer berechneten Größe ist mithilfe des Gauß'schen Fehlerfortpflanzungsgesetzes zu ermitteln (Gleichung 4.8). Die quadratische Addition berücksichtigt die geringe Wahrscheinlichkeit, dass maximale Abweichungen bei allen Messgrößen gleichzeitig auftreten. Vorausgesetzt sind unabhängige Messgrößen mit ähnlichen Messunsicherheiten. Sind diese Bedingungen nicht erfüllt, ist die lineare Fehlerfortpflanzung nach Gleichung 4.7 zu verwenden. [Tes17, S.40f.]

$$\sigma_f = \sqrt{\left( \frac{\partial f}{\partial x} \cdot \sigma_x \right)^2 + \left( \frac{\partial f}{\partial y} \cdot \sigma_y \right)^2 + \left( \frac{\partial f}{\partial z} \cdot \sigma_z \right)^2 + \dots} \quad (4.8)$$

Bei einem direkten Vergleich mit derselben Messtechnik in unverändertem Messaufbau ist die Wiederholpräzision entscheidend. In derartigen Anwendungsfällen können Abstriche bei den systematischen Messabweichungen und damit der absoluten Genauigkeit gemacht werden. Dafür muss die Wiederholpräzision über die Versuchsanzahl sowie die verwendete Messtechnik sicher gestellt sein.

## 4.2.2 Präzision des Prüfprozesses

Bei der Entwicklung eines Prüfprozesses mit reproduzierbaren, absoluten Ergebnissen sowie bei größeren Versuchsumfängen ist es sinnvoll die Prüf-

prozessstreuung auf Basis realer Experimente vorab zu bestimmen. Falls in diesen Voruntersuchungen die erforderliche Trennschärfe nicht erreicht werden kann, ist der Prüfprozess zu überarbeiten. Dafür kann schlimmstenfalls auf eine vollständige Wiederholung der Untersuchungen verzichtet werden.

Aufgrund der großen Anzahl und Streuung der Störgrößen bei einem Gesamtfahrzeugversuch kann innerhalb der Voruntersuchungen häufig nur ein Teil des Störspektrums erfasst werden. Zur Beurteilung der Wiederholpräzision im Versuchszeitraum sollte deshalb die auftretende Streuung während der Messungen mithilfe einer regelmäßig wiederholten Referenzversuchsreihe erfasst werden. Liegt die Standardabweichung dieser Referenzversuchsreihen über den Werten der Voruntersuchungen, so ist die erzielbare Messschärfe mit der aktuellen Varianz erneut zu berechnen um die Prüfprozesseignung zu bestätigen.

Je nach Zweck des Prüfprozesses lässt sich die Beurteilung der Prüfprozessstreuung etwas verkürzen. Bei reinen Wiederholungsversuchen kann die Voranalyse übersprungen und die Streuung erst nach den Messungen, zu Beginn der Auswertungen, bestätigt werden. Dabei besteht allerdings das Risiko, dass sich der Prüfprozess als ungeeignet erweist und damit die gesamte Messung nicht die geforderte Zuverlässigkeit besitzt oder sogar gänzlich unbrauchbar wird.

Das Vorgehen zur Beurteilung der Präzision auf Basis realer Versuche ist in Kapitel 4.4 detailliert beschrieben. Dort sind Anmerkungen zur Gewinnung der Datengrundlage sowie zur Auswertung von Referenzexperimenten zu finden. Die Prüfprozessanalyse basiert auf der statistischen Auswertung experimentell ermittelter Messwerte. Deshalb wird in den folgenden Unterkapiteln Bedeutung und Berechnung statistischer Kennzahlen erläutert sowie die Anwendung von Varianzanalysen im Kontext der Prüfprozessanalyse beschrieben.

Eine analytische Ermittlung der Messunsicherheit, wie im *Guide to the expression of uncertainty in measurement* (GUM) als Methode B beschrieben, setzt eine mathematische Modellierung der Messung voraus. Teilweise existieren entsprechende Modelle aus der Simulation des Fahrzeuges. Da die Streuungen der zahlreichen Einflussgrößen jedoch oft gar nicht oder

nur ungenau bekannt sind, kann eine analytische Abschätzung der Messunsicherheit weit neben den tatsächlichen Werten liegen. Außerdem sind die Modelle für Messsysteme oder Prüfprozesse wie ein Gesamtfahrzeugversuch häufig sehr komplex und nicht explizit darstellbar, sodass reale Versuche für eine zuverlässige Beurteilung letztendlich unumgänglich sind [Die14, S.194].

### 4.3 Statistische Methoden und Kennzahlen

Die Datenaufzeichnung eines Gesamtfahrzeugversuchs erfolgt zumeist in einem fortlaufenden, zeitdiskreten Messschrieb. Für die weitere Auswertung wird für jeden Versuch aus diesen Messdaten ein Beobachtungswert  $Y_{ijk}$ , beispielsweise ein Mittelwert oder ein Integral, abgeleitet. Die Indizes erlauben eine genaue Zuordnung der Beobachtungswerte als  $k$ -ter Einzelversuch in der Versuchsreihe  $j$  innerhalb von Experiment  $i$ . Die jeweiligen Großbuchstaben  $I$ ,  $J$ , und  $K$  stehen für den größten Wert des entsprechenden Index. Im besten Fall sind die Gruppen einer Ebene gleich groß, so dass die Indizierung der Gruppengröße entfallen kann.

$$I \quad \text{Anzahl der Experimente} \quad (4.9)$$

$$J_i \quad \text{Reihenanzahl in Experiment } i \quad (4.10)$$

$$K_{ij} \quad \text{Versuchsanzahl in Experiment } i, \text{ Reihe } j \quad (4.11)$$

$$K_i \quad \text{Versuchsanzahl jeder Reihe in Experiment } i \text{ bei homogener Versuchsanzahl} \quad (4.12)$$

$$K_i \cdot J_i = \sum_{j=1}^{J_i} K_{ij} \quad \text{Versuchsanzahl in Experiment } i \text{ bei homogener Versuchsanzahl} \quad (4.13)$$

$$K \cdot J = \sum_{j=1}^{J_i} K_{ij} \quad \text{Versuchsanzahl in jedem Experiment bei homogener Versuchs- und Reihenanzahl} \quad (4.14)$$

Mehrere unmittelbar hintereinander durchgeführte **Versuche** mit gleicher Konfiguration bilden eine **Versuchsreihe**. Jede Reihe wird durch ihren arithmetischen Mittelwert  $\bar{Y}_{ij}$  der Beobachtungswerte und eine Standardabweichung  $S_{ij}$  charakterisiert. Die Ursachen für die Ergebnisstreueung in-

nerhalb einer Reihe sind vom Prüfprozess abhängig. Wesentliche Quellen können der Bediener sowie die Veränderungen an den Umgebungsbedingungen durch den Versuch sein. Fahrzeug und Wetterbedingung werden sich in dieser kurzen Beobachtungszeit kaum verändern. Mess- und Auswerteprozess bleiben gleich.

Unter dem Begriff **Experiment** werden in dieser Arbeit Wiederholungen von Versuchsreihen zu einem anderen Zeitpunkt oder nach einer Unterbrechung durch andere Tests zusammengefasst. Einzige Voraussetzung ist, dass binnen eines Experiments kein Einflussfaktor gezielt variiert wurde. Dennoch lassen sich bei diversen Prüfprozessen mithilfe einer Varianzanalyse signifikante Abweichungen zwischen den Reihen feststellen. Neben Bediener und Rückwirkungen auf die Umgebungsbedingungen können nun auch Fahrzeug und Wetterbedingungen ausschlaggebende Ursachen dieser Streuung sein. Dies kann gegebenenfalls durch die Wahl des Auswerteverfahrens berücksichtigt werden. Abbildung 4.2 illustriert die in dieser Arbeit verwendete Nomenklatur zur Strukturierung einer Untersuchung.

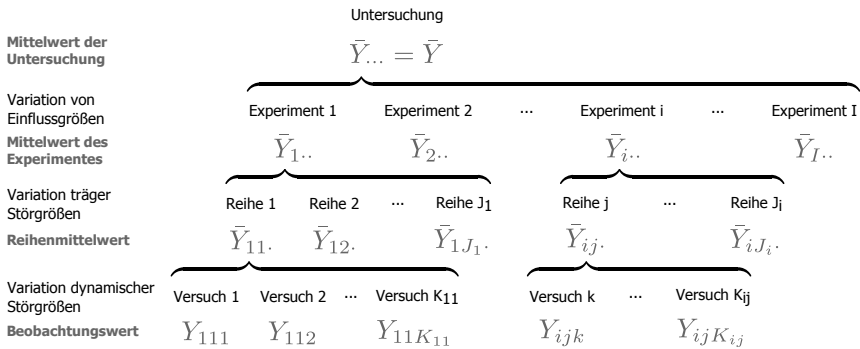


Bild 4.2: Gruppierungsebenen zur Strukturierung einer Untersuchung

Aufgrund der zahlreichen Störeinflüsse bei Fahrzeugversuchen mit mobilen Arbeitsmaschinen unterliegen die Messergebnisse einer nicht zu vernachlässigenden Streuung. Mithilfe der Statistik können Lage, Form und Ausprägung dieser Verteilung beschrieben werden. Aus diesem Grund wird zunächst ein kompakter Überblick über die verwendeten statistischen Kennzahlen und Methoden gegeben.

### 4.3.1 Lagemaße

Zur Bestimmung einer Approximation für eine stationäre oder eine mehrfach wiederholte Messung mit den Werten  $Y_1 \dots Y_K$  wird der Mittelwert  $\bar{Y}$  gebildet. Dieses Lagemaß spiegelt die zentrale Tendenz der Beobachtungen wider. Für den Mittelwert gibt es angepasst an den Datensatz verschiedene Definitionen mit den zugehörigen Bezeichnungen, wie geometrisches Mittel, harmonisches Mittel und Modus. Im Zusammenhang mit empirisch ermittelten Daten, die in einer metrischen Skalierung vorliegen, sind der Median und das **arithmetische Mittel** die am häufigsten benutzten Mittelwerte [Kro14, S. 19ff.]. Der letztere ist auch als Durchschnitt bekannt und berechnet sich als Quotient aus der Summe aller Werte und ihrer Anzahl.

$$\bar{Y} = \frac{1}{K} \cdot \sum_{k=1}^K Y_k \quad (4.15)$$

Um bei mehrfach indizierten Größen zu verdeutlichen, auf welche Werte sich die Mittelung bezieht, werden die entsprechende Indizes durch einen Punkt ersetzt. Demzufolge wird ein Reihenmittelwert wie in Gleichung 4.16, der Mittelwert eines Experiments gemäß Gleichung 4.17 dargestellt. In den Beispielen wird von homogenen Reihen- und Experimentumfängen ausgegangen.

$$\bar{Y}_{ij\cdot} = \frac{1}{K} \cdot \sum_{k=1}^K Y_{ijk} \quad (4.16)$$

$$\bar{Y}_{i\cdot\cdot} = \frac{1}{J \cdot K} \cdot \sum_{j=1}^J \sum_{k=1}^K Y_{ijk} \quad (4.17)$$

$$\bar{Y} = \bar{Y}_{\cdot\cdot\cdot} = \frac{1}{I \cdot J \cdot K} \cdot \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K Y_{ijk} \quad (4.18)$$

Mit der Annahme einer für die Grundgesamtheit repräsentativen Stichprobe gilt:

$$\bar{Y} = \mu \quad (4.19)$$

Der **Median** ist sehr robust gegen Extremwerte und Ausreißer und lässt sich auch auf ordinale Daten anwenden [Kro14, S. 22]. Zur Bestimmung werden die Daten zunächst nach der Größe in einer Rangwertreihe sortiert und der mittlere Wert bzw. der Durchschnitt der beiden mittleren Werte entnommen.

$$\tilde{Y} = \begin{cases} Y_{(\frac{k+1}{2})} & \text{für } k \text{ ungerade} \\ \frac{1}{2} \cdot \left( Y_{(\frac{k}{2})} + Y_{(\frac{k}{2}+1)} \right) & \text{für } k \text{ gerade} \end{cases} \quad (4.20)$$

Eine Mischform von arithmetischem Mittel und Median kann mithilfe eines **gewichteten Durchschnitts** gebildet werden. Dazu werden die Beobachtungswerte aufsteigend sortiert und der ersten Hälfte der Werte ihre jeweilige Position in der Reihe als Gewichtungsfaktor  $gf$  zugewiesen. Die zweite Hälfte der Beobachtungswerte erhält ihre Gewichtungsfaktoren durch ihre Position in einer absteigenden Sortierung. Wie in Tabelle 4.3 zu sehen, ergibt sich dadurch mit der Nähe zum Median eine zunehmende Gewichtung. Durch Potenzierung der Gewichtungsfaktoren kann der Einfluss des Median verstärkt oder bei einem Exponenten  $r$  kleiner 1 auch abgeschwächt werden.

$$\check{Y} = \frac{1}{K} \cdot \frac{\sum_{k=1}^K Y_k \cdot gf_k^r}{\sum_{k=1}^K gf_k^r} \quad (4.21)$$

Tabelle 4.3: Gewichtungsfaktoren gemäß Position in der Rangwertreihe

$Y_k$	$Y_1$	$Y_2$	$\dots$	$Y_{\frac{K}{2}-1}$	$Y_{\frac{K}{2}}$	$Y_{\frac{K}{2}+1}$	$\dots$	$Y_{K-1}$	$Y_K$
$gf_k^r$	$1^r$	$2^r$	$\dots$	$(\lceil \frac{K}{2} \rceil - 1)^r$	$\lceil \frac{K}{2} \rceil^r$	$\lfloor \frac{K}{2} \rfloor^r$	$\dots$	$2^r$	$1^r$

Mit dem Exponent Null liefert der gewichtete Durchschnitt das arithmetische Mittel, mit großen Exponenten zeigt sich schnell eine starke Annäherung an den Median. Zur Berechnung des gewichteten Mittels werden alle Beobachtungswerte mit ihrem zugehörigen Gewichtungsfaktor multipliziert und durch die Summe aller Gewichtungsfaktoren dividiert. Vorteile



dieses Mittelwertes sind die geringe Anfälligkeit gegen Ausreißer, wie beim Median, und die gute Abbildung des Messwertdurchschnitts bei kleiner Beobachtungszahl. Im letzten Fall wird auch der Einfluss der Diskretisierung gegenüber dem Median reduziert. Dennoch hat dieser Mittelwert in der Statistik keine Bedeutung oder ist zumindest in der Basisliteratur zur Statistik nicht zu finden.

Wenn bei Messdaten Ausreißer und Extremwerte zu erwarten sind, ist bei ausreichenden Messwerten der Median als aussagekräftigerer Mittelwert zu bevorzugen. Einen gut regulierbaren Kompromiss zwischen Median und arithmetischem Mittel bietet der gewichtete Durchschnitt. Bei einer symmetrischen Verteilung der Messwerte liefern arithmetisches Mittel und Median nahezu gleiche Werte. Ein großer Abstand zwischen diesen beiden Größen deutet auf Extremwerte und/oder eine asymmetrische Werteverteilung hin.

Mit dem Median wurde bereits der erste Rangwert eingeführt. Weitere Rangwerte sind das Minimum und das Maximum als erster und letzter Wert der Rangwertreihe. Charakteristische Werte innerhalb der Reihe lassen sich mithilfe des  $p$ -Quantils  $\tilde{Y}_p$  beschreiben. Der Median entspricht dem 0,5-Quantil, da die Hälfte der Werte kleiner als dieses Quantil ist. Insbesondere bei Boxplots wichtige Werte sind das **untere Quartil**  $\tilde{Y}_{0,25}$ , welches das Ende des unteren Viertels kennzeichnet, sowie das **obere Quartil**  $\tilde{Y}_{0,75}$ , das die Rangwertreihe nach 75 % der Werte teilt.

Um die Anwendung der hier und im folgenden dargestellten Berechnungsverfahren zu erleichtern, werden in den jeweiligen Abschnitten die entsprechenden Funktionsaufrufe für die Software MATLAB genannt. Dieses Programm der Firma MathWorks ist auf numerische Berechnungen beispielsweise bei der Auswertung von umfangreichen Messdaten ausgelegt und ist sowohl in der Industrie als auch an Hochschulen weit verbreitet.

In MATLAB können die vorgestellten Lagemaße mit den Funktionen in Gleichung 4.22 bis 4.24 berechnet werden.

$$\text{arithmetisches Mittel: } \bar{Y} \quad Y\_aM = \text{mean}(Y) \quad (4.22)$$

$$\text{Median: } \tilde{Y} \quad Y\_m = \text{median}(Y) \quad (4.23)$$

$$\text{p-Quantil: } \tilde{Y}_p \quad Y\_p = \text{quantile}(Y,p) \quad (4.24)$$

### 4.3.2 Streuungsmaße

Um nun zu beurteilen, wie gut ein Mittelwert die Stichprobe repräsentiert, werden die Abweichungen der einzelnen Beobachtungen zu diesem Mittelwert analysiert. Auch hier gibt es eine ganze Sammlung von Kennwerten, von denen die wichtigsten hier kurz vorgestellt werden sollen. Die gebräuchlichsten Streuungsparameter sind die empirische Varianz  $S^2$  und die **empirische Standardabweichung**  $S$  [Bur12, S. 92-98]. Die Varianz ist als Quotient aus der Summe aller quadratischen Abweichungen zum arithmetischen Mittelwert und der Anzahl der Freiheitsgrade definiert. Für Stichproben entspricht die Zahl der Beobachtungswerte minus eins den Freiheitsgraden. Die empirische Standardabweichung  $S$  hat die gleichen physikalischen Einheiten wie die Beobachtungsgröße und wird aus der Quadratwurzel der Varianz berechnet:

$$S = \sqrt{\frac{1}{K-1} \cdot \sum_{k=1}^K (Y_k - \bar{Y})^2} \quad (4.25)$$

Die Varianz ist inhaltlich schwieriger zu interpretieren, wird jedoch in vielen statistischen Methoden verwendet. Gelten die Kennwerte nicht nur für eine Stichprobe sondern für die Grundgesamtheit der Daten wird  $\sigma$  und  $\sigma^2$  für Standardabweichung und Varianz verwendet [Kro14, S. 34].

Die Streuung eines erwartungstreuen Mittelwerts wird mit

$$S_{\bar{Y}} = \frac{S}{\sqrt{K}} \quad (4.26)$$

abgeschätzt [Kro14, S.94]. Somit ergibt sich der Standardfehler der Differenz zweier Mittelwerte gemäß dem Fehlerfortpflanzungsgesetz entsprechend Gleichung 4.27.

$$\begin{aligned} S_{\bar{Y}_1 - \bar{Y}_2} &= \sqrt{\left(\frac{\partial f}{\partial \bar{Y}_1} \cdot S_{\bar{Y}_1}\right)^2 + \left(\frac{\partial f}{\partial \bar{Y}_2} \cdot S_{\bar{Y}_2}\right)^2} = \sqrt{\frac{S^2}{K_1} + \frac{S^2}{K_2}} \\ &= \sqrt{\frac{2}{K}} \cdot S \quad \text{für gleiche Stichprobenumfänge} \end{aligned} \quad (4.27)$$

Zum Vergleich der Streuung unterschiedlicher Größen wird der **Variationskoeffizient**  $V$  als Quotient aus empirischer Standardabweichung und arithmetischem Mittel berechnet [FKPT11, S. 69-76].

$$V = \frac{\sqrt{S^2}}{Y} \quad (4.28)$$

Eine alternative Herangehensweise zur Beurteilung der Streuung ist die Analyse der Differenz zweier Lagemaße. Die **Spannweite**  $R$  ist der Abstand zwischen Maximum und Minimum des Datensatzes. Bei Messwerten mit Ausreißern ist dieses Streuungsmaß kaum aussagekräftig, deutet jedoch auf solche Extremwerte hin. Für robustere Werte wird wiederum auf Quantile zurückgegriffen. Allgemein wird vom **Quantilabstand** gesprochen. Die Spanne zwischen unterem und oberem Quartil wird **Interquartilabstand**  $IQR$  oder zentraler Quartilabstand genannt [Bou13, S. 91ff.].

Wie für die Lagemaße existieren auch für einige Streuungsmaße vorgefertigte Funktionen in MATLAB:

$$\text{empirische Standardabweichung: } S \quad S = \text{std}(Y) \quad (4.29)$$

$$\text{Varianz: } S^2 \quad S\_sq = \text{var}(Y) \quad (4.30)$$

$$\text{Interquartilabstand: } IQR \quad IQR = \text{iqr}(Y) \quad (4.31)$$

### 4.3.3 Visualisierung statistischer Kennzahlen

Um auch ohne die zahlreichen statistischen Kennwerte schnell einen Überblick über die Qualität eines Datensatzes zu bekommen, wird eine Visualisierung in Form eines **Boxplots** genutzt. Abbildung 4.3 zeigt die Häufig-

keitsverteilung einer Messung in einem Histogramm. In dieses Balkendiagramm ist eine angepasste Normalverteilung als Kurve eingezeichnet. Darunter ist horizontal ein Boxplot dargestellt, dessen charakteristische Punkte beschriftet sind.

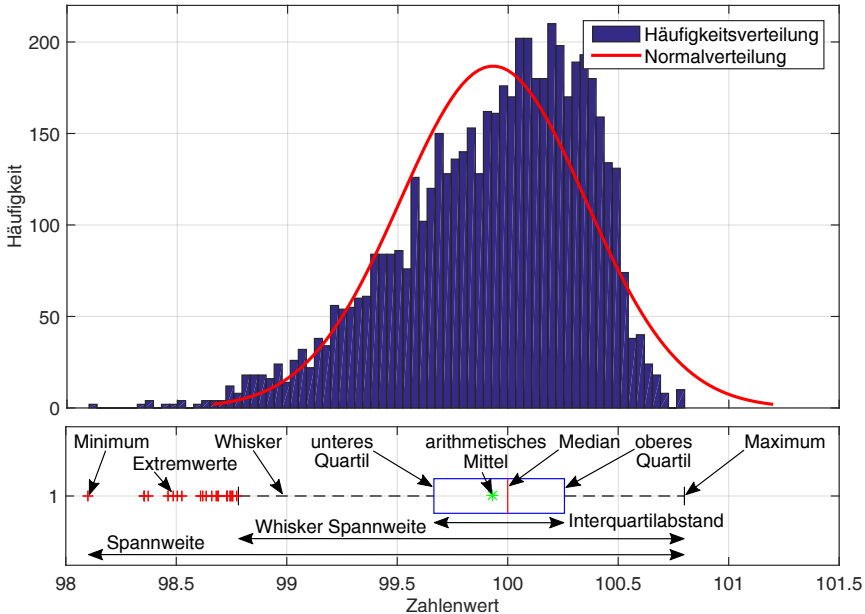


Bild 4.3: Statistische Kennwerte in Histogrammen und Boxplots

Median, unteres und oberes Quartil teilen die Menge der Beobachtungen in vier Abschnitte. Die Whisker begrenzen den Datensatz nach unten und oben. Ihre jeweilige Länge ist auf den 1,5-fachen Interquartilabstand begrenzt, was bei einer Normalverteilung 99,3% der Werte abdeckt. Kleinere und größere Werte werden als Ausreißer interpretiert und separat dargestellt. Sind alle Werte innerhalb dieser Grenzen, reichen die Whisker bis zum Minimum bzw. Maximum. Somit wird ein Eindruck von Spannweite und Interquartilabstand vermittelt. In dieser Arbeit wird zusätzlich das arithmetische Mittel als Stern illustriert. Über die Relativposition der Kennwerte lassen sich Rückschlüsse auf Form und Symmetrie der Verteilung ziehen.

Im gezeigten Beispiel in Abbildung 4.3 ist die asymmetrische Messwertverteilung an verhältnismäßig kurzen Abständen zwischen Median und oberem Quantil sowie zum rechten Whisker zu erkennen. Offensichtlich liegen die Werte in diesem Bereich sehr dicht beieinander. Nach oben gibt es keine Ausreißer, so dass der rechte Whisker hier das Maximum anzeigt. Dies deckt sich mit der Häufigkeitsverteilung des Histogramms. Bei den Abweichungen nach unten greift die Begrenzung der Whiskerlänge auf den 1,5-fachen Interquartilabstand. Auch wenn in diesem Beispiel nicht alle Werte links des Whiskers als Ausreißer interpretiert werden müssen, wird deutlich, dass es sich hierbei um wenige Extremwerte handelt. Dennoch haben diese einen erheblichen Einfluss auf das arithmetische Mittel, das klar vom Median abweicht.

### 4.3.4 Einfaktorielle Varianzanalyse

Aufgrund der zahlreichen Störeinflüsse bei Gesamtfahrzeugversuchen ist mit einer erheblichen Streuung der Beobachtungswerte eines Experiments zu rechnen. Dementsprechend ist die Auswertung nicht trivial. Das Auswerteverfahren soll zum einen Beobachtungswerte, die mit derselben Konfiguration der Einflussfaktoren ermittelt wurden, demselben Experiment zuordnen. Zum anderen sollen jedoch selbst kleine Änderungen der Einflussfaktoren in Vergleichsexperimenten zuverlässig unterschieden werden. Selbstredend sollen diese Unterschiede auch quantisiert werden.

Diese Anforderungen können mithilfe der Varianzanalyse (ANOVA) erfüllt werden. Die einfaktorielle Varianzanalyse kann als Verallgemeinerung des doppelt t-Test für eine beliebige Zahl unabhängiger Stichproben gesehen werden [FKPT11, S.518]. Bei diesem statistischen Test werden mindestens zwei unabhängige Stichproben, in unserem Fall Experimente, daraufhin überprüft, ob diese aus derselben Grundgesamtheit stammen können oder signifikante Unterschiede vorliegen. Die in Kapitel 4.3 beschriebene Reihenanzahl wird dabei nicht beachtet. Somit entspricht in diesem Kapitel der Wert  $K$  der Versuchsanzahl im Experiment.

Die Darstellung eines Untersuchungsergebnisses wird in Kapitel 4.3.5 erläutert. Die Abweichungen zwischen den Experimenten wird aus den Mit-

telwerten mit der zugehörigen Unsicherheit berechnet. Letztere kann aus der Varianz des Prüfprozesses abgeleitet werden.

In den folgenden Abschnitten werden zunächst das Vorgehen und die Zusammenhänge für die Analyse beliebiger Beobachtungswerte erläutert. Dies entspricht einem gängigen Anwendungsfall der Varianzanalyse mit den Schritten:

- Aufstellen von Nullhypothese und Alternativhypothese
- Überprüfung der Voraussetzungen der Testmethode
- Festlegen des Signifikanzniveaus
- Berechnen der Prüfgröße
- Vergleich der Prüfgröße mit der erwarteten statistischen Verteilung
- Testentscheidung

Zusätzlich zu diesem üblichen Vorgehen wird zur besseren Einordnung des Ergebnisses die Fehlerwahrscheinlichkeit für das Verwerfen der Nullhypothese für die berechnete Prüfgröße ermittelt.

Zur Veranschaulichung der Methode wird hier nochmals das Beispiel aus Abbildung 2.9 verwendet. Die Beobachtungswerte dieser Experimente sind in Tabelle 4.4 dargestellt.

Tabelle 4.4: Beobachtungswerte für das Anwendungsbeispiel der Bilder 2.9 und 4.4

	<b>Experiment</b>			
	<b>A1</b>	<b>B1</b>	<b>A2</b>	<b>B2</b>
<b>1. Beobachtungswert</b>	99,16	101,08	100,93	104,55
<b>2. Beobachtungswert</b>	101,05	101,64	98,84	101,69
<b>3. Beobachtungswert</b>	99,90	100,02	100,31	99,68
<b>4. Beobachtungswert</b>	100,13	101,58	96,82	100,35
<b>5. Beobachtungswert</b>	99,76	100,67	103,11	98,73
<b>arithmetisches Mittel</b>	100,00	101,00	100,00	101,00
<b>Standardabweichung</b>	0,69	0,67	2,35	2,26

Zunächst wird die **Nullhypothese**  $H_0$  aufgestellt, dass jede der betrachteten Gruppen denselben Erwartungswert  $\mu_i$  besitzt. Die Ergebnisse der verglichenen Experimente unterscheiden sich nur infolge der Prüfprozessstreuung. Die **Alternativhypothese**  $H_1$  ist erfüllt, wenn die Erwartungswerte von mindestens zwei Gruppen signifikant voneinander abweichen. Der Unterschied der Ergebnisse lässt sich nicht ausschließlich auf die Streuung zurückführen, sondern muss durch einen Einflussfaktor verursacht worden sein.

$$\begin{aligned} H_0 : \quad & \mu_1 = \mu_2 = \dots = \mu_i \\ H_1 : \quad & \mu_s \neq \mu_t \end{aligned} \tag{4.32}$$

Die Testentscheidung beruht auf einem Vergleich zwischen einer berechneten Prüfgröße, welche die charakteristische Verteilung der Beobachtungswerte einer Untersuchung widerspiegelt, und einer erwarteten statistischen Verteilung. Natürlich kann es bei dieser Testentscheidung auch zu einer Fehlentscheidung kommen. Wird die Nullhypothese verworfen, obwohl diese richtig ist, wird von einem Fehler 1. Ordnung gesprochen. Die Wahrscheinlichkeit für eine derartige Fehlentscheidung ist  $\alpha$  [BV08, 126f.]. Üblicherweise wird die maximal geduldete Fehlerwahrscheinlichkeit zur Verwerfung der Nullhypothese vorab als Signifikanzniveau festgelegt. Typische Werte für  $\alpha$  sind 5 %, 1 % oder 0,1 %. Der Vergleichswert der Teststatistik ergibt sich aus diesem Signifikanzniveau sowie dem Untersuchungsumfang. In den aufgeführten MATLAB-Funktionen wird die Fehlerwahrscheinlichkeit direkt im Ausgabewert  $p$  berechnet.

Der umgekehrte Fall, die Nullhypothese zu behalten, obwohl diese eigentlich abgelehnt werden müsste, ist ein Fehler 2. Ordnung mit der Wahrscheinlichkeit  $\beta$ . Bei Signifikanztests wird  $\beta$  jedoch normalerweise nicht berechnet. Der Fehler 2. Ordnung wird bei gleichgroßen Gruppenumfängen minimal [Hed16, S.583]. Folglich bedeutet das Behalten der Nullhypothese lediglich, dass mit den vorhandenen Daten keine signifikanten Unterschiede nachgewiesen werden können. Weitere Daten könnten diese Annahme widerlegen.

Nachdem die Fragestellung nach signifikanten Unterschieden als statistisches Testproblem formuliert wurde, ist zu klären, ob die vorliegenden Daten die **Voraussetzungen** zur Anwendung der einfaktoriellem Varianzanalyse erfüllen. Dazu sollten die Messdaten von Ausreißern befreit, unabhängig und normalverteilt sein sowie Varianzhomogenität aufweisen. Bei HUBER ist zusätzlich noch eine Gruppengröße von mindestens 20 Beobachtungswerten als Voraussetzung genannt [Hub14, S.63]. Außerdem ist hier mit Bezug auf BRAY und MAXWELL [Bra85, S.34] sowie auf PERREAULT und DARDEN [PD75, S.334] die Varianzanalyse bei ähnlich großen, umfangreichen Testgruppen als besonders Fehlerrobust selbst bei Verletzung dieser Bedingungen bezeichnet.

Die Bedingung gleicher Varianzen der Experimente sollte bei einer Gesamtfahrzeuguntersuchung selbstverständlich sein. Schließlich wird in der Regel bei einem Vergleich bewusst ein Einflussfaktor des Fahrzeugs oder das gesamte Fahrzeug, nicht aber die Störeinflüsse verändert. Dennoch kann Varianzhomogenität auftreten. Diese kann durch Änderungen in der Versuchsführung wie auch durch sehr kleine oder stark variierende Experimentumfänge hervorgerufen werden. Abhilfe kann folglich eine Angleichung oder eine Erhöhung der Versuchszahl pro Experiment sowie eine pedantische Einhaltung der Versuchsführung leisten. In der Literatur werden Mindestgruppenumfänge von 20 [Hub14, S.64] bis 30 [BV08, S.130], [FKPT11, S.390,438] empfohlen. Hintergrund ist die zunehmende Annäherung der T-Verteilung an die Normalverteilung für große Stichproben. Auf dieser Basis wird hier für Gesamtfahrzeugversuche die Empfehlung von mindestens 25 Einzelversuchen pro Experiment ausgesprochen. Bei einheitlichen Experimentumfängen von jeweils über 25 Beobachtungswerten, akkurater Versuchsdurchführung und Randomisierung der Versuchsrangfolge ist Varianzhomogenität ein Indiz dafür, dass die Veränderung des Einflussfaktors auch Auswirkungen auf die Störeinflüsse hat. Aus statistischer Sicht ist Varianzhomogenität bei gleicher Gruppengröße aller Vergleichsgruppen nicht so schwerwiegend und damit die Verwendung der Varianzanalyse nicht ausgeschlossen [Hed16, S.583].

Die Überprüfung eines Experiments auf **Varianzhomogenität** kann mit dem Levene-Test durchgeführt werden [Hub14, S.63]. Dieser Test ent-



spricht der einfaktoriellen Varianzanalyse mit dem Unterschied, dass die einzelnen Messungen vorab auf den Gruppenmittelwert transformiert werden. Wenn  $Y_{ik}$  dem  $k$ -ten Beobachtungswert aus dem Experiment  $i$  entspricht, so ist die Transformation auf den Gruppenmittelwert entsprechend Gleichung 4.33 durchzuführen.

$$X_{ik} = |Y_{ik} - \bar{Y}_i| \quad (4.33)$$

Die Testhypothesen des Levene-Tests zur Varianz sind analog zu 4.32 in Gleichung 4.34 aufgestellt.

$$\begin{aligned} H_0 : \sigma_1^2 &= \sigma_2^2 = \dots = \sigma_j^2 \\ H_1 : \sigma_s^2 &\neq \sigma_t^2 \end{aligned} \quad (4.34)$$

Da Prüfgröße und Ablehnungsbedingung der Varianzanalyse entsprechen, wird hier auf die später folgenden Gleichungen 4.42 und 4.43 verwiesen. In MATLAB ist dieser Test in der Funktion `vartestn` implementiert. Diese errechnet für je einen Vektor aller Beobachtungswerte  $Y$  und der Gruppenzuordnungen  $Gr$  die Fehlerwahrscheinlichkeit  $p$  zur Ablehnung der Nullhypothese. Die Transformation wird innerhalb der Routine durchgeführt.

$$p = \text{vartestn}(Y, Gr, 'TestType', 'LeveneAbsolute') \quad (4.35)$$

Die Anwendung auf die Untersuchungsdaten des Beispiels aus Tabelle 4.4 ergibt eine Fehlerwahrscheinlichkeit von 7,4% für die Ablehnung der Nullhypothese. Bei einem Signifikanzniveau von 5% wird die Nullhypothese demzufolge beibehalten. Es ist davon auszugehen, dass die Experimente der Untersuchung homogene Varianzen besitzen. Die Visualisierung der Werteverteilungen in Abbildung 4.4 vermittelt einen anderen Eindruck. Diese Diskrepanz lässt sich mit der geringe Teststärke des Levene-Test bei kleinen Gruppengrößen erklären.

Die Annahme einer **Normalverteilung** innerhalb eines jeden Experiments lässt sich ebenfalls mit einem statistischen Test stützen. Ein hierzu geeigneter Anpassungstest ist der Kolmogorov-Smirnov-Test. Hier lautet

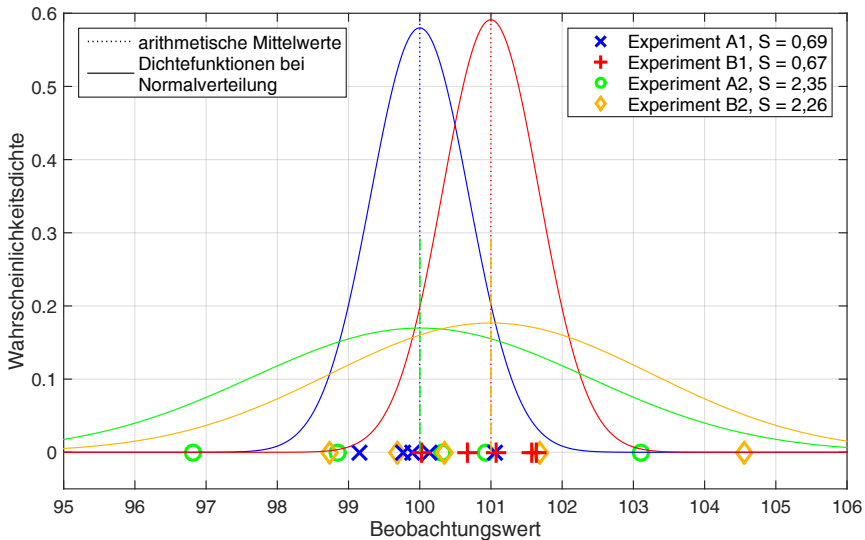


Bild 4.4: Beobachtungswerte des Anwendungsbeispiels Tabelle 4.4 und deren Verteilung

die Nullhypothese, dass die Wahrscheinlichkeitsverteilung der Normalverteilung entspricht.

$$\begin{aligned}
 H_0 : F_j(Y) &= F_{norm}(Y) \\
 H_1 : F_j(Y) &\neq F_{norm}(Y)
 \end{aligned}
 \tag{4.36}$$

Dieser Test beruht auf der Idee, die empirischen Verteilungsfunktionen der Messergebnisse und einer Normalverteilung mit demselben Erwartungswert  $\mu_i$  und der selber Varianz  $\sigma_i^2$  miteinander zu vergleichen. Die Anzahl der Beobachtungswert sowie das Signifikanzniveau, typischerweise  $\alpha = 5\%$ , bestimmen die akzeptierte Abweichung, bei deren Überschreitung die Nullhypothese abzulehnen ist. Für detailliertere Informationen zum Kolmogorov-Smirnov-Test wird an dieser Stelle auf HEDDERICH verwiesen [Hed16, S.461-465]. In MATLAB steht dieser Test unter dem Funktionsnamen *kstest* zur Verfügung. Zur Nutzung der Funktion muss vorab eine Normalverteilungsfunktion mit Erwartungswert und Varianz der Versuchs-

daten erstellt werden oder alternativ die Beobachtungswerte mithilfe der linearen Transformation in Gleichung 4.37 standardisiert werden.

$$Z_{ik} = \frac{Y_{ik} - \bar{Y}_i}{S_i}. \quad (4.37)$$

Die Funktion `kstest` gibt für den Vektor der Beobachtungswerte  $Z_i$  die Testentscheidung  $h$  sowie die berechnete Fehlerwahrscheinlichkeit  $p$  zur Ablehnung der Nullhypothese aus (Gleichung 4.38). Mit  $a$  wird das Signifikanzniveau für die Testentscheidung  $h$  festgelegt.

$$[h,p]=kstest(Z\_i,'Alpha',a) \quad (4.38)$$

Eine der zahlreichen Alternativen zum Kolmogorov-Smirnov-Test ist der Shapiro-Wilk-Test. Dieser Anpassungstest besitzt selbst bei kleinen Stichproben eine hohe Teststärke. Die Teststatistik beruht auf dem Vergleich der Varianz einer normalverteilten Stichprobe mit der tatsächlichen Varianz der untersuchten Stichprobe. [Hed16, S.466]

Dieser Test steht im *Mathworks File Exchange* auch als MATLAB Funktion zur Verfügung. Eingangsgröße ist für jedes Experiment ein Vektor  $Y_i$  mit den Beobachtungswerten dieser Versuchsgruppe sowie das Signifikanzniveau  $a$ .

$$[h,p]=swtest(Y\_i,a) \quad (4.39)$$

Für die Experimente des Anwendungsbeispiel aus Tabelle 4.4 liefert der Kolmogorov-Smirnov-Test Fehlerwahrscheinlichkeiten von über 91 %, der etwas stärkere Shapiro-Wilk-Test ermittelt Werte über 54 %. Beide Tests befürworten das Beibehalten der Nullhypothese, die Experimente seien normalverteilt.

Anders sieht es für das Datenbeispiel von Abbildung 4.3 aus. Hier ergibt sich für ein Signifikanzniveau  $a = 0,05$  beim Kolmogorov-Smirnov-Test die Testentscheidung  $h = 1$ , da die Fehlerwahrscheinlichkeit für die Ablehnung der Normalverteilungsannahme  $p = 7,89 E-23$  die akzeptierte Wahrscheinlichkeit  $a$  deutlich unterschreitet. Der Shapiro-Wilk-Test kommt mit  $p = 0$  zur selben Testentscheidung. Für  $h = 1$  ist die Normalverteilungsannahme

verletzt und eine rein zufällige Verteilung der Beobachtungswerte in Frage gestellt.

Können Varianzhomogenität oder Normalverteilung auch bei etwa gleichgroßen, umfangreichen Vergleichsgruppen nicht bestätigt werden, sollte für die weitere Auswertung auf einen verteilungsunabhängigen Test wie den Kruskal-Wallis-Test in Kapitel 4.3.9 zurückgegriffen werden [FKPT11, S.528].

Die **Unabhängigkeit der Beobachtungswerte** ist gegeben, wenn die Werte aus der Messung zur Zeit  $t_2$  nicht von einer vorhergehenden Messung zum Zeitpunkt  $t_1$  beeinflusst werden. Bei technischen Systemen ist es jedoch nicht unwahrscheinlich, dass sich der Systemzustand zwischen den Zeitpunkten  $t_1$  und  $t_2$  verändert hat. Beispiele dafür sind die Systemtemperaturen oder der Verschleißzustand der Maschine. Ebenso könnten sich Konzentration, Motivation und Übung des Bedieners ändern. Da der Systemzustand von den vorausgegangenen Betriebspunkten abhängt, ist es Bestandteil einer geschickten Versuchsführung diese Schwankungen minimal zu halten oder durch einen Wechsel der Versuchsfolge diese Störgrößeneinflüsse zu randomisieren.

GREENHOUSE und GEISSNER haben ein Korrekturverfahren erarbeitet, um die Varianzanalyse auch auf abhängige Beobachtungswerte einzusetzen. Dieses Verfahren wird in den Kapiteln 4.3.7 und 4.4.4 näher erläutert und eingesetzt.

Nachdem überprüft wurde, ob die Messungen zur Verwendung der Varianzanalyse geeignet sind, geht es an die Berechnung der **Teststatistik**. Das Testkonzept besteht darin, die mittlere Abweichungsquadratsumme zwischen den Gruppen  $MQS_{zw}$  der mittleren Abweichungsquadratsumme innerhalb der Gruppen  $MQS_{in}$  gegenüber zu stellen. Letztere wird aus der Quadratsumme  $QS_{in}$  und der Anzahl an Freiheitsgraden  $df_{in}$  berechnet. Auf die gleiche Weise wird auch die Streuung zwischen den Versuchsreihen mit der Quadratsumme  $QS_{zw}$  ermittelt. Die entsprechenden Formeln zur Berechnung dieser Quadratsummen sowie mögliche Vereinfachungen sind in den Gleichungen 4.40 und 4.41 angegeben. Dabei steht  $i$  für die Gruppennummer der insgesamt  $I$  Experimente mit jeweils  $K_i$  Werten.  $\bar{Y}_i$  ist

der arithmetische Mittelwert und  $S_i^2$  die Stichprobenvarianz der Gruppe  $i$ , während  $\bar{Y}$  dem arithmetischen Mittel aller Beobachtungswerte entspricht.

$$QS_{zw} = \sum_{i=1}^I \sum_{k=1}^{K_i} (\bar{Y}_i - \bar{Y})^2 = \sum_{i=1}^I K_i \cdot (\bar{Y}_i - \bar{Y})^2 \quad (4.40)$$

$$QS_{in} = \sum_{i=1}^I \sum_{k=1}^{K_i} (Y_{ik} - \bar{Y}_i)^2 = \sum_{i=1}^I (K_i - 1) \cdot S_i^2 \quad (4.41)$$

Analog zur in Gleichung 4.25 dargestellten Standardabweichung müssen auch diese Quadratsummen zur Normierung durch ihre Freiheitsgrade dividiert werden. Bei der Varianz zwischen den Versuchen ist das die Anzahl der Versuchsreihen  $I$  um eins reduziert, wohingegen die Varianz innerhalb der Gruppen  $I \cdot (K - 1)$  Freiheitsgrade besitzt. Der Quotient der normierten Varianzen ist die Prüfgröße der Varianzanalyse. Für diese Prüfgröße wird eine F-Verteilung mit  $I - 1$  und  $I \cdot (K - 1)$  Freiheitsgraden erwartet.

$$F = \frac{MQS_{zw}}{MQS_{in}} = \frac{\frac{1}{I-1} \cdot QS_{zw}}{\frac{1}{I \cdot (K-1)} \cdot QS_{in}} = \frac{\frac{1}{I-1} \cdot \sum_{i=1}^I K_i \cdot (\bar{Y}_i - \bar{Y})^2}{\frac{1}{I \cdot (K-1)} \cdot \sum_{i=1}^I (K_i - 1) \cdot S_i^2} \quad (4.42)$$

Falls die berechnete Prüfgröße das  $(1 - \alpha)$ -Quantil der  $F(I - 1, I \cdot (K - 1))$ -Verteilung überschreitet wird die Nullhypothese verworfen, da die Variabilität zwischen den Versuchsreihen deutlich größer als innerhalb der Reihen ist [FKPT11, S.519-528]. Diese Feststellung stände im Widerspruch zu der Annahme, dass alle Beobachtungswerte der gleichen Grundgesamtheit entstammen. Folglich lautet die Ablehnungsbedingung:

$$F > F_{1-\alpha}(I - 1, I \cdot (K - 1)) \quad (4.43)$$

In MATLAB wird die einfaktorielle Varianzanalyse mit der Funktion *anova1* aufgerufen. Wie zuvor beim Levene-Test ist auch hier  $Y$  ein Vektor mit allen Beobachtungswerten und  $Gr$  ein Vektor mit der Gruppenzuordnung. Ausgabewert ist die Fehlerwahrscheinlichkeit  $p$  für das Verwerfen der Null-

hypothese. Eine hohe Fehlerwahrscheinlichkeit bedeutet, dass keine prägnanten Unterschiede zwischen den Experimenten festzustellen sind.

$$p = \text{anova1}(Y, Gr) \quad (4.44)$$

Der Beispieldatensatz aus Tabelle 4.4 erfüllt alle Anforderungen zur Anwendung der Varianzanalyse. Lediglich die Gruppengröße liegt deutlich unter den Empfehlungen. Die Berechnung der Prüfgröße und der Teststatistik ist in den Gleichungen 4.45 bis 4.49 angedeutet.

$$MQS_{zw} = \frac{1}{4-1} \cdot \left( 5 \cdot (100 - 100,5)^2 + 5 \cdot (101 - 100,5)^2 + \dots \right) \quad (4.45)$$

$$MQS_{in} = \frac{1}{4 \cdot (5-1)} \cdot \left( (5-1) \cdot 0,69^2 + (5-1) \cdot 0,67^2 + \dots \right) \quad (4.46)$$

$$F = \frac{MQS_{zw}}{MQS_{in}} = \frac{\frac{1}{3} \cdot 4,98}{\frac{1}{16} \cdot 46,20} = 0,57 \quad (4.47)$$

$$F_{1-0,05}(4-1, 4 \cdot (5-1)) = F_{0,95}(3, 16) = 3,24 \quad (4.48)$$

$$F \not> F_{0,95}(3, 16) \quad \Rightarrow \text{keine signifikanten Unterschiede} \quad (4.49)$$

Die berechnete Fehlerwahrscheinlichkeit liegt bei 64 %. Die Ergebnisse dieser Untersuchung sind infolge der ausgeprägten Streuung der Experimente A2 und B2 statistisch nicht zu unterscheiden. Werden die Experimente A2 und B2 beispielsweise wegen Fehler bei der Versuchsdurchführung verworfen, kann für ein Signifikanzniveau  $\alpha = 0,05$  eine signifikante Abweichung zwischen den Experimenten A1 und B1 festgestellt werden. MATLAB ermittelt für diesen Fall die Fehlerwahrscheinlichkeit 4,9 %.

$$F = \frac{MQS_{zw}}{MQS_{in}} = \frac{\frac{1}{2-1} \cdot 2,49}{\frac{1}{2 \cdot (5-1)} \cdot 0,46} = 5,36 \quad (4.50)$$

$$F_{1-0,05}(2-1, 2 \cdot (5-1)) = F_{0,95}(1, 8) = 5,32 \quad (4.51)$$

$$F > F_{0,95}(1, 8) \quad \Rightarrow \text{signifikante Unterschiede} \quad (4.52)$$

### 4.3.5 Darstellung von Untersuchungsergebnissen

Im voran stehenden Beispiel wurde ein signifikanter Unterschied zwischen den Experimenten A1 und B1 festgestellt. Der Betrag dieser Abweichung ist allerdings nicht genau bekannt. Bei einer Wiederholung der beiden Experimente würden sicher andere Mittelwerte und damit eine abweichende Differenz ermittelt. Entsprechend Gleichung 4.26 ist für die Mittelwerte dieser Experimente mit der Standardabweichung  $S_{\bar{Y}_i} = \frac{0,69}{\sqrt{5}} = 0,31$  zu rechnen. Die Streuung einer Differenz wurde ebenfalls bereits in Kapitel 4.3.2, Gleichung 4.27 hergeleitet. Infolge der Fehlerfortpflanzung ergibt sich für dieses Beispiel bei der Differenz die Streuung  $S_{\bar{Y}_A - \bar{Y}_B} = \sqrt{\frac{2}{5}} \cdot 0,69 = 0,436$ . Diese Standardabweichung ist die Messunsicherheit  $u$  der Beispieluntersuchung. Weil die Messunsicherheit nur 68 % der zu erwartenden Ergebnisse überdeckt wird daraus die erweiterte Messunsicherheit  $U$  abgeleitet.

$$U = f \cdot u \quad (4.53)$$

In der Praxis ist ein Erweiterungsfaktor beziehungsweise Überdeckungsfaktor  $f = 2$  üblich. Unter Annahme einer Normalverteilung überdeckt die erweiterte Messunsicherheit damit 95 % der Ergebnisse. Für das Beispiel erhält man auf diese Weise  $U = \pm 0,87$ . Diese Berechnung erweist sich als grobe Schätzung der erzielbaren Schärfe entsprechend Gleichung 4.81 mit  $F_{1-\alpha} = 4$ . Hätten die Mittelwerte der Experimente A1 und B1 bei der selben Streuung einen Abstand  $|\bar{Y}_B - \bar{Y}_A| = 0,87$ , könnte bei einem Signifikanzniveau von  $\alpha = 5\%$  kein signifikanter Unterschied festgestellt werden. Nach Einsetzung der ermittelten Varianzen in Gleichung 4.79 zeigt sich, dass mit der einfaktoriellen Varianzanalyse erst ab einer Differenz von  $\pm 0,99$  eine signifikante Abweichung festgestellt werden könnte. Dieser Wert erweist sich somit als bessere Schätzung für die erweiterte Messunsicherheit  $U$ .

Wären die Experimente mit unterschiedlichen Messsystemen erfasst worden, müssten noch die Bias beider Systeme addiert werden. Das vollständige Untersuchungsergebnis ist nach Gleichung 4.54 anzugeben. Dabei ist die erweiterte Messunsicherheit mit maximal zwei gültigen Stellen darzu-

stellen. Das Untersuchungsergebnis ist auf die gleiche Auflösung zu runden. [Die14, S.174], [Tes17, S.18, 50f., 64ff.] nach [DIN99]

$$\text{Untersuchungsergebnis} = (\bar{Y} \pm U) [\text{Einheit}] \quad (4.54)$$

Demnach ist das vollständige Untersuchungsergebnis des Anwendungsbeispiels:

$$\bar{Y}_B - \bar{Y}_A = (1,00 \pm 0,99) \text{ Nm}$$

Zur Veranschaulichung wurde hier willkürlich eine Einheit für das synthetische Anwendungsbeispiel gewählt. Diese Darstellung verdeutlicht, dass der tatsächliche Unterschied wesentlich geringer als bei der analysierten Untersuchung sein könnte.

### 4.3.6 Mehrfaktorielle Varianzanalyse

Die Aufteilung der Streuungsanteile in systematische und unsystematische Einflüsse lässt sich auf weitere Einflussfaktoren erweitern. Durch die Zuordnung der Varianzen zu den verschiedenen Einflüssen wird der unsystematische Anteil kleiner und die Gesamtstreuung genauer erklärt. Die systematischen Einflüsse gliedern sich in Haupteffekte, die direkt durch einzelne Einflussfaktoren verursacht werden sowie Interaktionseffekte, die erst in Kombination mehrerer Faktoren wirken. In den Gleichungen 4.55 bis 4.57 ist beispielhaft die Streuungszersetzung für ein, zwei und drei Einflussfaktoren dargestellt.

$$QS_T = QS_A + QS_R \quad (4.55)$$

$$QS_T = \underbrace{QS_A + QS_B + QS_{A \times B}}_{\text{systematischer Anteil}} + \underbrace{QS_R}_{\text{unsystematischer Anteil}} \quad (4.56)$$

$$\begin{aligned} \underbrace{QS_T}_{\text{Gesamtstreuung}} &= \underbrace{QS_A + QS_B + QS_C}_{\text{Haupteffekte}} \\ &+ \underbrace{QS_{A \times B} + QS_{B \times C} + QS_{A \times C} + QS_{A \times B \times C}}_{\text{Interaktionen}} \\ &+ QS_R \end{aligned} \quad (4.57)$$



Die jeweiligen Abweichungsquadratsummen sowie die zugehörigen Freiheitsgrade für eine Untersuchung mit 2 Faktoren sind in Tabelle 4.6 aufgeführt. Die hierfür verwendete Nomenklatur für die Beobachtungswerte ist in Tabelle 4.5 skizziert. Faktor A wird in  $L$  Stufen, Faktor B in  $M$  Stufen variiert. Zur Anwendung in der dargestellten Form müssen für jede Kombination der Faktoren A und B  $N$  Beobachtungswerte vorliegen. Ebenso müssen in jeder Gruppe der Faktorstufe A  $M$  mal  $N$  beziehungsweise bei der Faktorstufe B  $L$  mal  $N$  Beobachtungen zur Verfügung stehen. Unter diesen Voraussetzungen wird von einem balancierten Design gesprochen. Weitere Bedingungen sind, wie bei der einfaktoriellen Varianzanalyse, Normalverteilung, Varianzhomogenität und Unabhängigkeit der Messdaten. Der Durchschnittswert aller Beobachtungswerte ist  $\bar{Y}$ . Die Mittelwerte der  $L$  Stufen des Faktors A sind  $\bar{Y}_{l.}$ , für den Faktor B dem entsprechend  $\bar{Y}_{.m.}$ .  $\bar{Y}_{lm.}$  ist das arithmetische Mittel einer Zelle mit der Faktorstufe  $l$  des Faktors A und der Ausprägung  $m$  des Faktors B.

Tabelle 4.5: Nomenklatur der Beobachtungswerte einer zweifaktoriellen Varianzanalyse

Faktor A	Faktor B			Mittelwerte Faktorstufen A
	Stufe $B_1$	...	Stufe $B_M$	
Stufe $A_1$	$\left. \begin{array}{c} Y_{111} \\ \vdots \\ Y_{11N} \end{array} \right\} \bar{Y}_{11.}$	...	$\left. \begin{array}{c} Y_{1M1} \\ \vdots \\ Y_{1MN} \end{array} \right\} \bar{Y}_{1M.}$	$\bar{Y}_{1.}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
Stufe $A_L$	$\left. \begin{array}{c} Y_{L11} \\ \vdots \\ Y_{L1N} \end{array} \right\} \bar{Y}_{L1.}$	...	$\left. \begin{array}{c} Y_{LM1} \\ \vdots \\ Y_{LMN} \end{array} \right\} \bar{Y}_{LM.}$	$\bar{Y}_{L.}$
Mittelwerte Faktorstufen B	$\bar{Y}_{.1.}$	...	$\bar{Y}_{.M.}$	$\bar{Y}$

Mit den Abweichungsquadratsummen  $QS$  und den zugehörigen Freiheitsgraden  $df$  aus Tabelle 4.6 werden die Prüfgrößen  $F$  der Haupteffekte und Interaktionen gebildet (Gleichungen 4.58 bis 4.60). Überschreitet die Prüfgröße das  $(1 - \alpha)$ -Quantil der F-Verteilung mit den entsprechenden Freiheitsgraden, so liegt ein signifikanter Einfluss durch den betrachteten Faktor vor.

Tabelle 4.6: Abweichungsquadratsummen und Freiheitsgrade einer Untersuchung mit 2 Faktoren

Streuungsanteil durch	Abweichungsquadratsumme	Freiheitsgrade
Faktorstufen A	$QS_A = M \cdot N \cdot \sum_{l=1}^L (\bar{Y}_{l..} - \bar{Y})^2$	$L - 1$
Faktorstufen B	$QS_B = L \cdot N \cdot \sum_{m=1}^M (\bar{Y}_{.m.} - \bar{Y})^2$	$M - 1$
Interaktion A mit B	$QS_{A \times B} = N \cdot \sum_{l=1}^L \sum_{m=1}^M (\bar{Y}_{lm.} - \bar{Y}_{l..} - \bar{Y}_{.m.} + \bar{Y})^2$	$(L-1) \cdot (M-1)$
innerhalb der Zellen (residual)	$QS_R = \sum_{l=1}^L \sum_{m=1}^M \sum_{n=1}^N (Y_{lmn} - \bar{Y}_{lm.})^2$	$L \cdot M \cdot (N - 1)$
Gesamtstreuung	$QS_T = \sum_{l=1}^L \sum_{m=1}^M \sum_{n=1}^N (Y_{lmn} - \bar{Y})^2$	$L \cdot M \cdot N - 1$

$$F_A = \frac{MQS_A}{MQS_R} = \frac{\frac{QS_A}{df_A}}{\frac{QS_R}{df_R}} \quad (4.58)$$

$$F_B = \frac{MQS_B}{MQS_R} = \frac{\frac{QS_B}{df_B}}{\frac{QS_R}{df_R}} \quad (4.59)$$

$$F_{A \times B} = \frac{MQS_{A \times B}}{MQS_R} = \frac{\frac{QS_{A \times B}}{df_{A \times B}}}{\frac{QS_R}{df_R}} \quad (4.60)$$

Dieses Vorgehen kann auf weitere Einflussfaktoren übertragen werden. Wegen des zunehmenden Berechnungsaufwands empfiehlt sich spätestens an dieser Stelle eine rechnerunterstützte Auswertung. In MATLAB steht hierfür die Funktion *anovan* zur Verfügung. *Y* ist wieder ein Vektor mit allen Beobachtungswerten und *Gr\_X* sind Vektoren um die Stufen des jeweiligen Faktors zuzuordnen. Durch die Ergänzung *'model',2* werden auch Interaktionseffekte mit zwei Einflussfaktoren ausgewertet.

$$p = \text{anovan}(Y, \{Gr\_A, Gr\_B, Gr\_C, \dots\}, 'model', 2) \quad (4.61)$$

### 4.3.7 Varianzanalyse für abhängige Stichproben

Wie bereits in Kapitel 4.3.4 angedeutet, sind die Beobachtungswerte von technischen Systemen nicht immer unabhängig. Fahrzeugzustand, insbesondere Temperaturen, Bedienerverhalten und Umgebungsbedingungen verändern sich häufig während des Messzeitraums in Abhängigkeit der vorausgegangenen Versuche. Diese Änderungen erhöhen als Störgrößen die Streuung und reduzieren damit die Messschärfe des Prüfprozesses.

Auch für abhängige Beobachtungswerte hält die Statistik ein adäquates Auswerteverfahren bereit. Dieses findet unter der Bezeichnung *Varianzanalyse mit Messwiederholungen* häufig in Arzneimittelversuchen sowie in den Sozialwissenschaften Verwendung. Hier soll die Wirkung der Einflussgröße von den Eigenschaften der Versuchspersonen getrennt werden.

Ein Beispiel ist eine Untersuchung zur Reaktionsfähigkeit von Probanden unter Einfluss von unterschiedlichen Mengen Kaffee. Dazu wiederholen die Versuchsteilnehmer einen Reaktionstest nach dem Konsum festge-

legter Kaffeemengen. Hierbei werden einige Probanden bereits ohne Kaffeeeinfluss eine deutlich höhere Reaktionsfähigkeit als andere Teilnehmer zeigen. Die Beobachtungswerte sind damit nicht nur von der Kaffeemenge, sondern auch vom jeweiligen Versuchsteilnehmer abhängig.

Mithilfe des Auswerteverfahrens *Varianzanalyse mit Messwiederholungen* können die abweichende Leistungsniveaus der Probanden herausgerechnet werden, um die Wirkung des Einflussfaktors Kaffee zu isolieren. Dazu werden die Versuchspersonen als Ausprägungen eines zusätzlichen Einflussfaktors  $V_{pn}$  in die Varianzanalyse aufgenommen. Der Einfluss der Versuchspersonen kann so von den unsystematischen zu den systematischen Streuungsanteilen verschoben werden. Weil in jeder Zelle aber nur ein Beobachtungswert zur Verfügung steht, kann der unsystematische Streuungsanteil nicht wie bei der zweifaktoriellen Varianzanalyse aus der Streuung innerhalb der Zellen berechnet werden. Stattdessen wird zur Berechnung der Prüfgröße  $F$  die Residualvarianz aus der Interaktion zwischen Einflussfaktor und Versuchsperson geschätzt. Diese Schätzung beinhaltet untrennbar die unsystematischen Streuungsanteile wie Messfehler. Damit ergibt sich die Prüfgröße entsprechend Gleichung 4.62 [RFHN06, S.103-121].

$$F_A = \frac{MQS_A}{MQS_{A \times V_{pn}}} = \frac{\frac{QS_A}{df_A}}{\frac{QS_{A \times V_{pn}}}{df_{A \times V_{pn}}}} = \frac{Q \cdot \frac{\sum_{l=1}^L (\bar{Y}_l - \bar{Y})^2}{L-1}}{\frac{\sum_{l=1}^L \sum_{q=1}^Q (Y_{lq} - \bar{Y}_l - \bar{Y}_{\cdot q} + \bar{Y})^2}{(L-1) \cdot (Q-1)}} \quad (4.62)$$

Wenn bei einer Untersuchung jede Kombination aus Versuchsperson und Stufe des Einflussfaktors mehrfach getestet wird, wäre eine Auswertung mit einer zweifaktoriellen Varianzanalyse denkbar. Weil jedoch vorab der Einfluss einer Versuchswiederholung nicht bekannt ist, sollte auch die Wiederholung als eigener Faktor in die Auswertung mit einfließen. In einer *zweifaktoriellen Varianzanalyse mit Messwiederholung auf beiden Faktoren* können die Einflüsse der untersuchten Einflussgröße (Kaffee), der Wiederholung sowie eine mögliche Interaktion dieser beiden Einflüsse erfasst werden. Die Prüfgröße  $F$  wird dabei jeweils aus dem untersuchten Einflussfak-

tor und der Interaktion zwischen Einflussfaktor und Versuchsperson gebildet (Gleichungen 4.63 bis 4.65).  $F_B$  lässt sich analog  $F_A$  in Gleichung 4.62 aufstellen. Die Berechnung von  $F_{A \times B}$  ist in Gleichung 4.66 dargestellt.

$$F_A = \frac{MQS_A}{MQS_{A \times V_{pn}}} \quad (4.63)$$

$$F_B = \frac{MQS_B}{MQS_{B \times V_{pn}}} \quad (4.64)$$

$$F_{A \times B} = \frac{MQS_{A \times B}}{MQS_{A \times B \times V_{pn}}} \quad (4.65)$$

$$F_{A \times B} = \frac{N \cdot \sum_{l=1}^L \sum_{m=1}^M (\bar{Y}_{lm.} - \bar{Y}_{l..} - \bar{Y}_{.m.} + \bar{Y})^2}{(L-1) \cdot (M-1)} \cdot \frac{\sum_{l=1}^L \sum_{m=1}^M \sum_{q=1}^Q (Y_{lmq} - \bar{Y}_{lm.} - \bar{Y}_{l.q} - \bar{Y}_{.mq} + \bar{Y}_{l..} + \bar{Y}_{.m.} + \bar{Y}_{..q} - \bar{Y})^2}{(L-1) \cdot (M-1) \cdot (Q-1)} \quad (4.66)$$

Voraussetzung einer Varianzanalyse für abhängige Stichproben ist die Varianzhomogenität „der Differenzen zwischen jeweils zwei Faktorstufen“, die Sphärizität [RFHN06, s.108]. Bei nur zwei Stufen des untersuchten Faktors ist diese Bedingung immer erfüllt. Eine Verletzung der Sphärizität kann mit dem Korrekturfaktor  $\varepsilon$  bei den Freiheitsgraden von Zähler und Nenner kompensiert werden. Das Korrekturverfahren nach GREENHOUSE und GEISSNER ist unabhängig vom Grad der Verletzung und gleichzeitig die konservativste Korrektur. Der Korrekturfaktor berechnet sich gemäß Gleichung 4.67 in Abhängigkeit der Anzahl der messwiederholten Faktorstufen [RFHN06, s.111]. Besteht eine Abhängigkeit zwischen den Beobachtungswerten der Faktorstufen von  $B$  ist folglich  $\varepsilon_B$  zu verwenden.

$$\varepsilon_A = \frac{1}{L-1} \quad \varepsilon_B = \frac{1}{M-1} \quad \varepsilon_{A \times B} = \frac{1}{(L-1) \cdot (M-1)} \quad (4.67)$$

Von BOX sowie von HUYNH und FELDT wurden ebenfalls Verfahren zur Bestimmung von  $\varepsilon$  entwickelt. Diese Korrekturen beziehen die Stärke der Verletzung der Sphärizität mit ein und sind entsprechend liberaler [RFHN06,

S.110f.]. Zugunsten der Zuverlässigkeit und Komplexität des Beurteilungsverfahrens erfolgt in dieser Arbeit die Abschätzung grundsätzlich ohne Prüfung der Sphärizität mit der konservativsten Korrektur nach GREENHOUSE und GEISSNER.

Auch bei der Aufzeichnung abhängiger Stichproben ist auf eine Randomisierung der Versuchsreihenfolge zu achten. Wenn die Ausprägung  $A_2$  immer nach  $A_1$  durchgeführt wird, könnte beispielsweise ein Trainingseffekt statt des Einflussfaktors Ursache für einen signifikanten Unterschied sein. Hier wird von Sequenzeffekten gesprochen. Zur Untersuchung derartiger Effekte kann die Versuchsreihenfolge als zusätzlicher Einflussfaktor in der Varianzanalyse mit aufgenommen werden.

### 4.3.8 Kovarianzanalyse

Die Kovarianzanalyse (ANCOVA) kombinierte eine Regression mit einer Varianzanalyse. Dieses Verfahren kann dazu genutzt werden um die Wirkung eines messbaren Störeinflusses herauszurechnen und damit die unsystematische Streuung zu reduzieren.

Ein Beispiel: In einer Studie wird die Wirksamkeit verschiedener Diäten untersucht. Je ein Drittel der zufällig gewählte Probanden führen dabei die gleiche Ernährungsumstellung durch. Die untersuchte Einflussgröße ist folglich der Faktor „Diät“ in drei Ausprägungen. Die Beobachtungsgröße ist der Gewichtsverlust nach einem Monat. Neben der Ernährung gibt es viele weitere Größen, wie Alter, Geschlecht, tägliche Bewegung oder den Body-Mass-Index, die alle einen Einfluss auf den Gewichtsverlust haben. Da diese Größen nicht gezielt gesteuert werden können, gelten sie als Störeinflüsse. Falls diese Störeinflüsse bei der Untersuchung erfasst wurden, kann ihr Einfluss auf die Beobachtungsgröße herausgerechnet werden. Dazu wird die Wirkung jedes Störeinflusses in einer Regression abgeschätzt und die Beobachtungswerte entsprechend korrigiert. Anschließend erfolgt die Varianzanalyse mit diesen modifizierten Beobachtungswerten. Mit sinnvoll gewählten Kovariaten kann die unsystematische Streuung so reduziert und damit die Teststärke erhöht werden.

Die mathematische Umsetzung dieser Idee weicht von der anschaulichen Erläuterung ab. Tatsächlich wird die Varianz der Kovariate direkt aus den

einzelnen Quadratsummen herausgerechnet. Dazu wird die Quadratsummenzerlegung für den Einflussfaktor (4.68), die Kovariate (4.69) sowie die Produktsumme (4.70) beider Einflüsse durchgeführt [Vos09].

$$\overbrace{\sum_{l=1}^L \sum_{m=1}^M (Y_{lm} - \bar{Y})^2}^{QS_T(Y)} = \overbrace{\sum_{l=1}^L M \cdot (Y_l - \bar{Y})^2}^{QS_{zw}(Y)} + \overbrace{\sum_{l=1}^L \sum_{m=1}^M (Y_{lm} - \bar{Y}_l)^2}^{QS_{in}(Y)} \quad (4.68)$$

$$\overbrace{\sum_{l=1}^L \sum_{m=1}^M (X_{lm} - \bar{X})^2}^{QS_T(X)} = \overbrace{\sum_{l=1}^L M \cdot (X_l - \bar{X})^2}^{QS_{zw}(X)} + \overbrace{\sum_{l=1}^L \sum_{m=1}^M (X_{lm} - \bar{X}_l)^2}^{QS_{in}(X)} \quad (4.69)$$

$$\begin{aligned} \overbrace{\sum_{l=1}^L \sum_{m=1}^M (X_{lm} - \bar{X})(Y_{lm} - \bar{Y})}^{PS_T} &= \overbrace{\sum_{l=1}^L M \cdot (X_l - \bar{X})(Y_l - \bar{Y})}^{PS_{zw}} \\ &+ \underbrace{\sum_{l=1}^L \sum_{m=1}^M (X_{lm} - \bar{X}_l)(Y_{lm} - \bar{Y}_l)}_{PS_{in}} \end{aligned} \quad (4.70)$$

Aus diesen Quadratsummen werden die modifizierten Quadratsummen für die Varianzanalyse berechnet. Mit der Anpassung der Quadratsummen geht auch eine Modifikation des Freiheitsgrades  $df'_{in} = df_{in} - 1 = M \cdot (L - 1) - 1$  einher. Die Varianzanalyse erfolgt mit den modifizierten Quadratsummen und Freiheitsgraden [Vos09].

$$QS'_T(Y) = QS_T(Y) - \frac{PS_T^2}{QS_T(X)} \quad (4.71)$$

$$QS'_{in}(Y) = QS_{in}(Y) - \frac{PS_{in}^2}{QS_{in}(X)} \quad (4.72)$$

$$QS'_{zw}(Y) = QS_{zw}(Y) + \frac{PS_{in}^2}{QS_{in}(X)} - \frac{PS_T^2}{QS_T(X)} \quad (4.73)$$

Neben den Voraussetzungen der Varianzanalyse müssen die Messdaten einer Kovarianzanalyse zusätzliche Anforderung erfüllen. Die Kovariable

muss mit der Beobachtungsgröße korrelieren. Dabei muss der Regressionskoeffizient aller Messdaten etwa den Koeffizienten der einzelnen Gruppen entsprechen. Außerdem soll keine signifikante Interaktion zwischen Kovariable und Faktor existieren. Andernfalls kommt es zu einer Verzerrung der Ergebnisse und somit zu Fehlinterpretationen [Hub14, S.69ff.]. Laut BORTZ ist die Kovarianzanalyse bei homogenem Stichprobenumfang und normalverteilten Residuen jedoch ausgesprochen robust gegen die Verletzung einzelner Bedingungen [BS10, S.311f.].

### 4.3.9 Kruskal-Walis-Test

Der Kruskal-Walis-Test ist als Verallgemeinerung des Wilcoxon-Rangsummen-Tests ein verteilungsunabhängiger Test. Alle Beobachtungswerte werden in einer Rangliste sortiert und die Rangsummen der einzelnen Vergleichsgruppen  $RS_j$  berechnet. Die Prüfgröße des Kruskal-Walis-Test in Gleichung 4.74 besitzt bei größeren Testumfängen näherungsweise eine  $\chi^2$ -Verteilung mit  $J - 1$  Freiheitsgraden. Für kleinere Prüfumfänge existieren entsprechende Wertetabellen. [Hed16, S.604-616]

$$H = \frac{12}{\sum_{j=1}^J K_j \cdot \left( \sum_{j=1}^J K_j + 1 \right)} \cdot \sum_{j=1}^J \frac{RS_j^2}{K_j} - 3 \cdot \left( \sum_{j=1}^J K_j + 1 \right) \quad (4.74)$$

Bei homogener Gruppengröße  $K$  aller  $J$  Vergleichsgruppen vereinfacht sich die Teststatistik zu:

$$H = \frac{12}{K^2 \cdot (KJ + 1)} \cdot \sum_{j=1}^J RS_j^2 - 3 \cdot (KJ + 1) \quad (4.75)$$

Ist dieser berechnete H-Wert größer als ein Schwellwert, weicht mindestens eine der untersuchten Gruppen signifikant von den anderen Gruppen ab. In Tabelle 4.7 sind die kritischen Schranken für ein Signifikanzniveau von  $\alpha = 5\%$  bei gleich großen Gruppenumfängen gelistet. Die Werte wurden aus [Hed16, S.605] übernommen. Dort sind auch die genauen Werte für drei Gruppen inhomogener Größe mit bis zu 5 Beobachtungswerten zu finden.



Laut HEDDERICH ist bei allen grau gedruckten Werten die Verwendung der  $\chi^2$ -Verteilung zulässig. Da die  $\chi^2$ -Verteilung stets größere Werte als die Tabellenwerte annimmt, wird der Test dadurch konservativer. Die Nullhypothese wird häufiger beibehalten, die Wahrscheinlichkeit Fehler 1. Ordnung sinkt,  $\beta$  für Fehler 2. Ordnung steigt.

Tabelle 4.7: Kritische Schranken für  $\alpha = 5\%$  beim Kruskal-Wallis-Test

Werte pro Gruppe	Gruppenanzahl			
	3	4	5	6
<b>3</b>	5,292	6,850	8,330	9,758
<b>4</b>	5,546	7,162	8,685	10,147
<b>5</b>	5,668	7,323	8,874	10,359
<b>6</b>	5,738	7,421	8,991	10,492
<b>7</b>	5,783	7,486	9,071	10,583
<b>8</b>	5,815	7,533	9,128	10,649
<b>9</b>	5,838	7,568	9,172	10,700
<b>10</b>	5,856	7,596	9,206	10,740
<b>20</b>	5,929	7,711	9,353	10,911
<b>50</b>	5,968	7,775	9,435	11,008
$\infty$	5,990	7,813	9,485	11,067
$\chi^2$	5,992	7,815	9,488	11,071

Der Test setzt keine Varianzhomogenität oder Normalverteilung voraus und kann deshalb alternativ zur Varianzanalyse eingesetzt werden. Einzige Voraussetzung ist, dass alle Daten zufallsmäßig und unabhängig der selben Grundgesamtheit entstammen [Hed16, S.447]. Aufgrund der geringen Anforderungen an die Beobachtungswerte, ist die Teststärke des Kruskal-Wallis-Test geringer als bei der Varianzanalyse. Sie erreicht aber immerhin eine asymptotische Effizienz von 95% im Vergleich zur Varianzanalyse [Hed16, S.604-616]. Laut HEDDERICH können insbesondere bei einer kleinen Datengrundlage mit  $K \lesssim 20$  verteilungsunabhängige Tests, wie der von Kruskal-Wallis wirksamer als die sonst optimalen parametri-

schen Tests sein [Hed16, S.442]. Die Umformung und Rückwärtsrechnung des verteilungsfreien Tests zur Abschätzung der erreichbaren Schärfe wie bei der einfaktoriellen Varianzanalyse ist allerdings nicht möglich.

In MATLAB wird der Kruskal-Wallis-Test unter Verwendung einer  $\chi^2$ -Verteilung aufgerufen. Die Eingangsvektoren sind Vektoren analog zur Varianzanalyse.

$$p = \text{kruskalwallis}(Y, Gr) \quad (4.76)$$

## 4.4 Quantifizierung der Präzision

In Kapitel 4.2 wurde an die Unsicherheit von Messergebnissen erinnert. Auflösung und Genauigkeit des Messsystems, die zu systematischen Abweichungen führen, werden anhand der Kriterien in Kapitel 4.2.1 beurteilt. Im Folgenden werden die zufälligen Messabweichungen, die eine Streuung der Messwerte im Prüfprozess verursachen, evaluiert. Hierzu sollte zunächst die Wiederholpräzision und je nach Anforderungen an den Prüfprozess die Vergleichspräzision bestimmt werden.

Die mathematische Quantisierung der Präzision variiert je nach Quelle. Häufig wird die Standardabweichung genannt, teilweise wird aber auch Varianz, mittlere Abweichungsquadratsumme, Spannweite oder reziproke Standardabweichung verwendet. In dieser Arbeit wird als Maß für die Präzision die Standardabweichung sowie der vierfachen Interquartilabstand angegeben. Letzterer beinhaltet bei normalverteilten Beobachtungswerten 99,3% der Werte, ist im Gegensatz zur Spannweite robust gegen Ausreißer und vermittelt somit einen guten Eindruck zur Bandbreite der Messdaten. Die Ermittlung und Verwendung des Variationskoeffizienten ist nur unter zwei Voraussetzungen sinnvoll. Zum einen sollten die Daten Verhältnisskaliert sein, also einen natürlichen, absoluten Nullpunkt besitzen [Bou13, S.16f, S.101ff.]. Zum anderen sollte die Streuung der Daten proportional zu ihrem Mittelwert ansteigen. Während die erste Voraussetzung bei physikalischen Größen im Allgemeinen erfüllt ist, trifft die zweite Annahme jedoch bei den meisten Gesamtfahrzeugversuchen nicht zu. Dementsprechend ist eine Abschätzung der Präzision anderer Wertebereiche auf Basis des Va-

riationskoeffizienten oder ein Vergleich unterschiedlicher Größen kritisch zu betrachten.

Für den Experimentator ist vor allem von Interesse, ob ein erwarteter Effekt auf die Zielgröße zuverlässig nachgewiesen werden kann. Hierbei ist jedoch die Anzahl der Einzelversuche und Versuchsreihen pro Experiment von großer Bedeutung. Mithilfe von Varianzanalysen wird abgeschätzt, welche Messschärfe unter Annahme eines bestimmten Versuchsaufwands erreicht werden kann. Dazu wird auf Basis von Referenzversuchen ein Toleranzband errechnet innerhalb dessen keine signifikanten Unterschiede erkannt werden können. Dies ist das maßgebliche Ergebnis zur Beurteilung der Prüfprozesseignung.

#### 4.4.1 Datenbasis für die Prüfprozessbeurteilung

Zur Beurteilung der Präzision wird der Prüfprozess entsprechend der Prozessbeschreibung in mehreren Versuchsreihen mit jeweils einigen Versuchswiederholungen durchgeführt und messtechnisch erfasst. Unmittelbar hintereinander ausgeführte Einzelversuche bilden bei konstant gehaltenen Einflussgrößen eine Versuchsreihe. Zwischen den Versuchsreihen werden andere Prüfprozesse ausgeführt oder pausiert, wie es im regulären Testbetrieb zu erwarten ist. Alle Versuche beziehungsweise Versuchsreihen mit identischer Konfiguration der Einflussgrößen werden in einem Experiment zusammengefasst.

Bei Prüfprozessen mit stationären Betriebspunkten für mehrere Maschineneinstellungen, sollte die Eignung über den gesamten Variationsbereich überprüft werden. Dazu sind mindestens die Streuungen bei Minimum, Maximum und arithmetischen Mittel des Variationsbereichs zu erfassen. Gegebenenfalls ist ein Prüfverfahren sowie Auswertung mit Direktvergleichen, entsprechend Kapitel 4.4.4 sinnvoll.

Varianzanalysen, die zur Versuchsauswertung und zur Abschätzung der erzielbaren Schärfe verwendet werden, basieren unter anderem auf der Annahme normalverteilter Beobachtungswerte. Bei großem Stichprobenumfang kann die Streuung zufälliger Beobachtungswerte mit einer Normalverteilung beschrieben werden. Ab etwa 25 Werten sind die Abweichungen dieser Approximation akzeptabel [Hub14, S.64], [FKPT11, S.390,438].

Deshalb sollten pro Experiment mindestens 25 Einzelversuche durchgeführt werden. Falls diese Versuche in unmittelbarer Folge absolviert werden, besteht eine hohe Wahrscheinlichkeit, dass zu vergleichende Experimente unter abweichenden Störeinflüssen aufgenommen werden. Darum empfiehlt sich die Aufsplittung eines Experiments in mehrere Versuchsreihen. Je nach Anforderungen an den Prüfprozess sollte die Durchführung dieser Versuchsreihen über mehreren Tage verteilt sein. Mit dem Kruskal-Wallis-Test wird analysiert, ob sich die Störeinflüsse zwischen den einzelnen Versuchsreihen signifikant verändert haben. Auch dieser Test gewinnt mit zunehmendem Gruppenumfang an Stärke. Als guter Kompromiss zwischen erfasstem Störgrößenspektrum und Teststärke hat sich eine Aufsplittung in fünf oder mehr Versuchsreihen mit jeweils mindestens fünf Einzelversuchen erwiesen.

Zur Bewertung der Reproduzierbarkeit müssen ebenfalls mehrere Experimente berücksichtigt werden, bei denen beispielsweise Maschine, Bediener, Messsystem, Örtlichkeit, Umweltbedingungen sowie Kombinationen davon variieren. Zur Bestimmung der Prüfprozessstreuung müssen die zulässigen Bedingungen aus der Prozessbeschreibung bekannt sein und über den gesamten Wertebereich variiert werden. Da sich einige Einflussgrößen, wie der Bediener oder die Eigenschaften eines Erdbodens kaum quantitativ beschreiben lassen, ist eine voll umfängliche Überprüfung kaum möglich. Außerdem ergibt sich durch die Kombination der Einflussgrößen in ihren verschiedenen Ausprägungen ein praktisch nicht zu bewältigender Experimentekatalog. Folglich kann die Reproduzierbarkeit höchstens durch die Auswahl einiger Kombinationen abgeschätzt werden.

Mit der Thematik einer möglichst geschickten Auswahl beschäftigt sich das Fachgebiet *statistische Versuchsplanung* (design of experiments, DoE) [Sie10]. Dabei sollte ein ausgewogener, orthogonaler Versuchsplan angestrebt werden. Das bedeutet, dass jede Faktorstufe eines Einflussfaktors gleich oft getestet wird und keine Kombination aus jeweils zwei Spalten des Versuchsplans miteinander korreliert. Für je zwei Faktorstufen erfüllt eine Feldkonstruktion nach PLACKETT-BURMANN diese Bedingungen und kann dabei sehr effizient die Hauptwechselwirkungen aufzeigen. Zur Aufschlüsselung von Haupteffekten und Zweifachwechselwirkungen müssen weitere

Versuche mit einer sogenannten Faltung des Versuchsplans durchgeführt werden. Auf diese Weise lässt sich nach einem groben Screening die Auflösungsstufe mit geringem Mehraufwand erhöhen. Eine mögliche Alternative wäre ein Versuchsplan nach dem YATES-Standard. Eine wichtige Voraussetzung für die Analyse ist die Möglichkeit die verschiedenen Einflussfaktoren beliebig zu kombinieren. Aufwendige Stufenwechsel können durch Umsortierung der Experimentreihenfolge reduziert, aber nicht umgangen werden. [Sie10, S.25-33]

Um mit einem überschaubaren Versuchsaufwand die Streuung eines Prüfprozesses festzustellen, empfiehlt sich demzufolge ein Versuchsplan mit lediglich zwei Stufen pro Einflussfaktor. Zweckmäßige Faktorstufen sind der untere sowie der obere Grenzwert des zulässigen Bereichs. Aus der Menge der untersuchten Einflussfaktoren leitet sich die erforderliche Experimentanzahl ab. Bei Versuchsplänen nach Plackett-Burmann ist dies ein Vielfaches von vier. Damit die Wechselwirkungsanteile die Haupteffekte nicht zu stark verfälschen, sollte der Versuchsplan mindestens drei Experimente mehr als Faktoren haben.

Beispielhaft ist in Tabelle 4.8 ein teilfaktorieller Versuchsplan für fünf Faktoren auf zwei Stufen in acht Experimenten dargestellt. Darin wird jeder Faktor jeweils viermal mit den beiden Stufen + und – abgeprüft. Zum Erreichen einer höheren Ergebnisqualität kann dieser Versuchsplan zusätzlich gefaltet, also mit invertierten Faktorausprägungen durchgeführt werden. Alternativ könnte auch direkt ein größerer Versuchsplan mit beispielsweise 12 Experimenten aufgestellt werden. Eine Erhöhung der Experimentanzahl zu Lasten der Experimentumfänge ist nur bis ca. 25 Einzelversuche pro Experiment akzeptabel, da andernfalls Erwartungswert und Streuung innerhalb des Experiments zunehmend unzuverlässig werden.

In der praktischen Anwendung ist ein derartiger, idealer Versuchsplan nur mit sehr großem Aufwand umsetzbar. Je nach Zielgröße ist beispielsweise der Wechsel des Messsystems oder des Fahrzeuges sehr kostenintensiv. Folglich können nicht alle Faktorkombinationen und Interaktionen beleuchtet werden. Des Weiteren bleibt unklar, ob bei schwer skalierbaren Einflussfaktoren wie Bediener oder Untergrund tatsächlich obere und untere zulässige Ausprägung geprüft werden konnte. Hier ist es häufig angebracht

Tabelle 4.8: Teilfaktorierter Versuchsplan für fünf Faktoren auf zwei Stufen in acht Experimenten nach PLACKETT-BURMANN

Experiment Nummer	Faktor				
	A	B	C	D	E
1	+	-	+	-	-
2	+	+	-	+	-
3	+	+	+	-	+
4	-	+	+	+	-
5	-	-	+	+	+
6	+	-	-	+	+
7	-	+	-	-	+
8	-	-	-	-	-

mehr als zwei Faktorstufen zu testen. Insbesondere bei den Untersuchungen zur Reproduzierbarkeit sollte klar sein, dass als Ergebnis lediglich eine Abschätzung geliefert wird.

Für die statistische Auswertung der Experimente muss für jeden Einzelversuch ein unabhängiger **Beobachtungswert** abgeleitet werden. Bei dynamischen Zyklen kann diese Zielgröße als Integral aus dem zeitkontinuierlichen Messschrieb gebildet werden. Beispiele sind die zurück gelegte Strecke, der verbrauchte Kraftstoff oder die geleistete Arbeit pro Versuch. Für stationäre Untersuchungen wird pro Versuch aufs Neue ein Betriebspunkt angefahren und gehalten, bis sich die Zielgröße nicht mehr wesentlich verändert. Messrauschen sowie Inhomogenitäten der Umweltbedingungen innerhalb eines Versuchs verursachen eine Streuung der Messwerte im fortlaufenden Messschrieb. Über den näherungsweise stationären Abschnitt eines jeden Versuchs wird ein Messfenster gelegt, innerhalb dessen für jedes erfasste Merkmal ein Mittelwert als statistische Einheit gebildet wird. Bei den Auswertungen der vorliegenden Gesamtfahrzeugversuchen hat sich hierfür aus Sicht des Autors der gewichtete Durchschnittswert nach Gleichung 4.21 mit dem Exponent zwei bewährt. Dieses Versuchsergebnis wird im Folgenden als Beobachtungswert verwendet.

## 4.4.2 Varianz der Beobachtungswerte eines Referenzexperiments

Auf Ebene der Experimente wird untersucht ob sich die Streuung innerhalb einer Versuchsreihe signifikant von der Streuung zwischen den Versuchsreihen unterscheiden. Ein schneller Überblick kann mithilfe von Boxplots der Reihen gewonnen werden. Weicht eine Box in Position oder Ausdehnung wesentlich von den anderen Boxplots ab, so darf angenommen werden, dass während dieser Reihe ein zusätzlicher Einfluss wirkte. Kann dieser Einfluss einem Prozessfehler zugeordnet werden, darf der entsprechende Versuch beziehungsweise diese Reihe aus der weiteren Auswertung entfernt werden. Andernfalls ist für diese Reihe von einer weiteren Störgröße auszugehen.

Eine objektivere Aussage kann mittels einer statistischen Auswertung erlangt werden. Da die Abschätzung der Varianz erst bei größeren Stichproben mit mehr als zwanzig Beobachtungswerten pro Vergleichsgruppe zuverlässig ist, wird hierzu der Kruskal-Walis-Test anstatt der Varianzanalyse genutzt. Die Beobachtungswerte eines Experiments werden entsprechend Kapitel 4.3.9 in eine Rangliste sortiert, um mit den Rangsummen der Versuchsreihen  $RS_j$  den Kruskal-Walis-Test durchzuführen.

Signifikante Abweichungen innerhalb des Experiments zeigen, dass die Störgrößen zwischen den Versuchsreihen ausgeprägter als innerhalb einer Reihe wirken. Ungewiss bleibt in diesem Fall, ob mit dem Referenzexperiment die gesamte Streubreite dieser Störgrößen erfasst wurden. Deshalb wird bei signifikanten Abweichungen innerhalb des Experiments bei der Beurteilung der Prüfprozesseignung ein größerer Unsicherheitsfaktor verwendet. Selbst wenn innerhalb der Referenzmessungen keine signifikanten Abweichungen zwischen den Reihen festgestellt wurden, so gibt das keine Sicherheit, dass die Ausprägung der Störgrößen auch in weiteren Reihen annähernd gleich bleibt. Dennoch erfolgt aufgrund der vermeintlich höheren Zuverlässigkeit der Experimentstreuung die Prüfprozessbeurteilung mit einem kleineren Faktor.

Die Abschätzung der erzielbaren Schärfe erfolgt auf Basis der geplanten Versuchsanzahl und der **Varianz der Referenzexperimente**. Wird lediglich die Wiederholbarkeit beurteilt, so wird direkt die Varianz des Refe-

renzexperimente verwendet. Wurden mehrere Betriebspunkte untersucht, sind die Varianzen innerhalb der Experimente mit dem Levene-Test auf Homogenität zu untersuchen. Bei Inhomogenitäten sollte die Standardabweichung zwischen den ermittelten Betriebspunkten interpoliert oder direkt die größte aufgetretene Streuung verwendet werden. Weicht die Varianz nicht signifikant zwischen den Betriebspunkten ab, ist die durchschnittliche Standardabweichung zu ermitteln. Da hier die Berücksichtigung aller Einzelwerte bei unterschiedlichen Experimentumfängen zu einer unberechtigten Gewichtung führen würden, werden jeweils nur die Experimentergebnisse miteinander verrechnet.

Für die Analyse der Reproduzierbarkeit gilt ein ähnliches Vorgehen. Die Streuung des Prüfprozesses ergibt sich hier aus der Varianz der Erwartungswerte sowie der Streuung innerhalb der Experimente. Da alle Prüfprozesse mit den gleichen Betriebsparametern durchgeführt wurden, sollten alle Experimente näherungsweise gleiche Varianzen aufweisen. Bei Varianzhomogenität und gleichen Experimentumfängen entspricht die Gesamtstreuung der Streuung des Prüfprozesses. Nicht selten zeigt sich jedoch eine Abhängigkeit der Varianz von den veränderten Prüfbedingungen. Auch wenn die ANOVA bei gleichen Gruppengrößen recht robust gegen eine Verletzung der Varianzhomogenität ist [Hed16, S.583], sind formal die Voraussetzungen für die Varianzanalyse nicht erfüllt. Deshalb sollte bei Varianzinhomogenität in Betracht gezogen werden die Rahmenbedingungen des Prüfprozesses enger einzuschränken.

Mit dem Ziel die maximal auftretende Streuung abzuschätzen, wird bei Varianzinhomogenität zwischen den Experimenten die maximal gemessene Experimentvarianz eingerechnet. Zur Berücksichtigung in der Gesamtvarianz muss die Quadratsumme der Abweichungen innerhalb der Experimente von der Gesamtstreuung abgezogen und dafür die Quadratsumme des Experiments mit der größten Varianz entsprechend der Experimentanzahl aufaddiert werden (Gleichung 4.77). Aus der so korrigierten Abweichungsquadratsumme  $QS_{TK}$  wird die erwartete Varianz  $S_{Ref}^2$  abgeschätzt. Die Anzahl der Freiheitsgrade bleibt dabei unverändert.



$$\begin{aligned}
 QS_{TK} = & \overbrace{\sum_{l=1}^L \sum_{m=1}^M \sum_{n=1}^N (Y_{lmn} - \bar{Y})^2}^{QS_T} - \overbrace{\sum_{l=1}^L \sum_{m=1}^M \sum_{n=1}^N (Y_{lmn} - \bar{Y}_{lm.})^2}^{QS_R} \\
 & + L \cdot M \cdot \max \left( \sum_{n=1}^N (Y_{lmn} - \bar{Y}_{lm.})^2 \right)
 \end{aligned} \quad (4.77)$$

Des Weiteren ist es sinnvoll die Streuungsanteile aufgrund der verschiedenen, gezielt variierten Einflussfaktoren zu ermitteln. Durch eine mehrfaktorielle Varianzanalyse können die signifikanten Streuungseinflüsse identifiziert werden. Diese Informationen können beispielsweise zur Optimierung des Prüfprozesses oder zur differenzierteren Auswertung von Versuchsergebnissen genutzt werden.

### 4.4.3 Erzielbare Schärfe

Die in Kapitel 4.3.4 beschriebene Varianzanalyse ist ein statistisches Verfahren zur Auswertung von zwei oder mehr Vergleichsgruppen hinsichtlich signifikanter Abweichungen zwischen den Gruppen. Für die Beurteilung der Prüfprozesseignung soll abgeschätzt werden, bis zu welcher Differenz mögliche Vergleichsgruppen voneinander unterschieden werden können. Dazu stehen jedoch lediglich die statistischen Eigenschaften einer oder mehrerer Referenzgruppen zur Verfügung. Aus diesen Informationen soll ermittelt werden, wie stark bei einer Veränderung der Versuchsbedingungen oder am Versuchsobjekt der Effekt sein müsste, um ihn sicher von den Referenzversuchen zu unterscheiden. Diese Information kann mit einigen Annahmen durch die Rückwärtsrechnung der Varianzanalyse gewonnen werden.

Dazu wird die Prüfgröße aus Gleichung 4.42 für zwei Gruppen aufgestellt und vereinfacht. Die erste Gruppe  $Y_1$  enthält die Beobachtungswerte eines Referenzexperiments, während sich die zweite Gruppe  $Y_2$  aus den hypothetischen Vergleichsmessungen zusammensetzt.

$$F = \frac{\frac{n_1 \cdot n_2}{n_1 + n_2} \cdot (\bar{Y}_1 - \bar{Y}_2)^2}{\frac{1}{n_1 + n_2 - 2} \cdot \left( (n_1 - 1) \cdot S_1^2 + (n_2 - 1) \cdot S_2^2 \right)} \quad (4.78)$$

Mit der Auflösung dieser Gleichung nach der Differenz der beiden Mittelwerte, steht bereits eine passende Funktion zur Berechnung der möglichen Untersuchungsgenauigkeit zur Verfügung.

$$\pm |\bar{Y}_1 - \bar{Y}_2| = \sqrt{\frac{F_{1-\alpha}(1, n_1 + n_2 - 2) \cdot \left( (n_1 - 1) \cdot S_1^2 + (n_2 - 1) \cdot S_2^2 \right)}{\frac{n_1 \cdot n_2}{n_1 + n_2} \cdot (n_1 + n_2 - 2)}} \quad (4.79)$$

Diese Gleichung liefert ein Toleranzband innerhalb dessen Unterschiede nicht zuverlässig detektiert werden können. Weicht die Vergleichsgruppe jedoch stärker als das Toleranzband ab, so kann dies unter einer statistischen Fehlerwahrscheinlichkeit festgestellt werden. Mit dieser Gleichung wird somit die Schärfe gegenüber einem Vergleichsexperiment abgeschätzt. Zeigt sich bei den tatsächlichen Messungen eine deutlich abweichende Varianz oder die Zurückweisung einer anderen Voraussetzung der ANOVA, wird diese Abschätzung gegebenenfalls hinfällig.

Die Datenbasis dieser Abschätzung ist die Varianz der Referenzversuche entsprechend den Ausführungen in Kapitel 4.4.2. Ungewiss bleibt, ob bei den Referenzuntersuchungen alle Störeinflüsse, die in einer Prüfung vorkommen können, tatsächlich wirken konnten. Dieser Unsicherheit wird mit einer Reduzierung des Signifikanzniveaus  $\alpha$  Rechnung getragen. An dieser Stelle wird das Ergebnis des Kruskal-Wallis-Tests relevant. Zeigten sich innerhalb der Referenzexperimente keine signifikanten Abweichungen zwischen den Reihen, so wird die ermittelte Varianz als verhältnismäßig zuverlässig eingestuft und die F-Verteilung mit einem Fünftel des gewünschten Signifikanzniveaus ermittelt. Bei signifikanten Abweichungen innerhalb eines Experiments ist diese Ungewissheit deutlich höher, weswegen nur ein Fünfzigstel von  $\alpha$  bei der Berechnung des F-Wertes berücksichtigt wird.

Durch die Reduzierung des Signifikanzniveaus verringert sich der Fehler 1. Art, eine Abweichung irrtümlich als Unterschied zu erkennen. Gleichzeitig erhöht sich jedoch auch der Fehler 2. Art, eine tatsächlich vorhandene Differenz zu übersehen. Die Manipulation von  $\alpha$  anstelle der Varianz selbst bietet den Vorteil, dass unter Annahme einer korrekt ermittelten Streuung weiterhin Aussagen zur Fehlerwahrscheinlichkeit gemacht werden können. Mit dem so angepassten Signifikanzniveau und der Summe an geplanten Versuchen pro Experiment  $n_1$  und  $n_2$  wird aus einer statistischen Tafel oder mithilfe eines Rechenprogramms das  $(1 - \alpha)$ -Quantil der  $F(1, n_1 + n_2 - 2)$ -Verteilung ermittelt. In MATLAB dient hierzu die Funktion *finv* gemäß Gleichung 4.80.

$$F = \text{finv}(1-\alpha, 1, n_1 + n_2 - 2) \quad (4.80)$$

Nachdem mit mehreren Referenzexperimenten verhältnismäßig viel Aufwand betrieben wurde, um die Streuung des Prüfprozesses möglichst genau zu ermitteln, sollte diese Varianz sowohl in der Abschätzung der Untersuchungsgenauigkeit als auch bei der Auswertung verwendet werden. Wurde während einer Untersuchung über den gesamten Prüfzeitraum regelmäßig ein Referenzexperiment gemäß dem Prüfprozess wiederholt, sollte die Varianz dieser Versuche für die Auswertung von Wiederholungsvergleichen verwendet werden. In diesem Fall ist eine Anpassung des Signifikanzniveaus nicht erforderlich. Bei der Reproduktion eines Experiments sollte die Streuung, die unter Variation von Maschine, Bediener, Messsystem, Örtlichkeit, Umweltbedingungen usw. ermittelt wurde, eingesetzt werden. Falls einzelne Einflussfaktoren zwischen den zu vergleichenden Experimenten nicht verändert wurden, kann die Varianz um die entsprechenden Streuungsanteile dieser Faktoren reduziert werden. Dazu wird die Abweichungsquadratsumme des konstanten Einflussfaktors von der Gesamtabweichungsquadratsumme der Referenzexperimente abgezogen und die Varianz neu berechnet. Die Freiheitsgrade werden dabei auf die Versuchszahl abzüglich der Faktorstufen des herausgerechneten Faktors reduziert. Die Berechnung der Abweichungsquadratsumme zwischen den Beobachtungswerten und den Mittelwerten der nicht variierten Faktorstufen führt

wahlweise zum selben Ergebnis. Aufgrund der verbleibenden Unsicherheit sollte jedoch weiterhin mit einem angepassten Signifikanzniveau gearbeitet werden.

Wird neben einer übereinstimmenden Varianz davon ausgegangen, dass für jedes Vergleichsexperiment gleich viele Wiederholungen zur Verfügung stehen, vereinfacht sich die Schätzung für die mögliche Schärfe zu:

$$\pm |\bar{Y}_1 - \bar{Y}_2| = \sqrt{\frac{F_{1-\alpha}(1, 2(n_1-1))}{n_1} \cdot 2 S_{Ref}^2} \quad (4.81)$$

Ist die erzielbare Messschärfe kleiner als der zu untersuchende Effekt, so wird der Prüfprozess als geeignet eingestuft. Den Gleichungen 4.79 und 4.81 ist zu entnehmen, dass sich die mögliche Untersuchungs-genauigkeit insbesondere durch die Erhöhung des Versuchsumfanges verbessern lässt. Der Zusammenhang zwischen Versuchsanzahl und erzielbaren Schärfe ist in Abbildung 4.5 veranschaulicht. In diesem Beispiel könnte bei  $n_1 = 25$  Versuchen eine Abweichung von 1 auf einem Signifikanzniveau von  $\alpha = 1\%$  festgestellt werden. Zur Reduzierung der Varianz beziehungsweise der Standardabweichung sind grundlegende Änderungen in der Versuchsführung erforderlich.

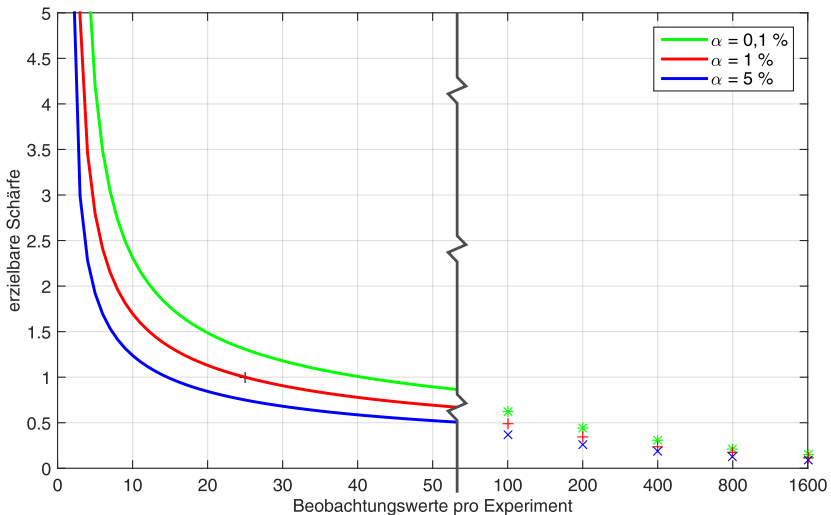


Bild 4.5: Erzielbare Schärfe in Abhängigkeit der Versuchsanzahl

In der Statistik-Literatur wird für Post-Hoc-Analysen häufig der Tukey HSD-Test oder der Scheffé-Test verwendet. Während letzterer sich eher konservativ und relativ robust gegenüber Verletzungen der Testvoraussetzungen erweist [BS10, S.232-235], bietet der Tukey HSD-Test eine sehr hohe Teststärke [RFHN06, S.46ff.]. Diese Tests schätzen alle auf Basis der Residualvarianz eine kritische Differenz, die die Mittelwerte der Untersuchungsgruppen überschreiten müssen, um einen signifikanten Unterschied zu erkennen. Auf eine Abschätzung der Residualvarianz zukünftiger Messungen wird in den Quellen nicht eingegangen. Mit dem Ziel die Anzahl und Komplexität sowie die Nachvollziehbarkeit der verwendeten statistischen Verfahren in dieser Arbeit in einem überschaubaren Rahmen zu halten, soll im Weiteren die Abschätzung der erzielbaren Schärfe auf Basis des F-Tests erfolgen.

#### 4.4.4 Auswerteverfahren für Direktvergleiche

In einem realitätsnahen und damit praxisrelevanten Prüfprozess für eine mobile Arbeitsmaschine wirken, wie bereits in den vorangegangenen Kapiteln detailliert ausgeführt, eine Vielzahl teilweise schwer erfassbarer Störeinflüsse. Eine engere Beschränkung der Prüfbedingungen führt zur Reduktion dieser Streuungen, begrenzt jedoch auch die zeitliche Verfügbarkeit des Prüfprozesses. Der hier erforderliche Kompromiss bedingt eine Reststreuung, mit der im besten Fall eine Differenzierung im Prozentbereich möglich ist. Eine höhere Trennschärfe kann durch die Verlagerung der Prüfung in eine synthetische Prüfumgebung oder einen Direktvergleich erzielt werden.

Bei einem Direktvergleich wird die Annahme getroffen, dass unmittelbar aufeinander folgende Versuchsreihen unter nahezu gleichen Prüfbedingungen durchgeführt werden. Für die Auswertung wird ein weiterer Einflussfaktor für die Prüfbedingungen eingeführt. Dies führt zu einer Reduktion der unsystematischen Streuungsanteile und damit zu einer maßgeblich besseren Differenzierbarkeit der Ergebnisse. Dem steht gegenüber, dass lediglich ein relatives Ergebnis in Bezug zu einem Referenzfahrzeug oder zum Referenzzustand des Fahrzeuges gewonnen werden kann. Für eine absolute Einordnung verschiedener Prüflinge müssen die Versuche mit dem

Referenzfahrzeug bei den jeweiligen aktuellen Prüfbedingungen wiederholt werden.

Dieses Verfahren ist besonders für die Optimierung eines Fahrzeuges zu empfehlen. Beispiele hierfür sind Untersuchungen verschiedener Applikationsdatensätze für die Maschinensteuerung oder die Gegenüberstellung eines Eco-Betriebsmodus im Vergleich zum Normalbetrieb. Voraussetzung für dieses Vorgehen ist, dass der untersuchte Einflussfaktor sehr kurzfristig gewechselt werden kann. Dieses Konzept ist auch im unmittelbaren Vergleich verschiedener Fahrzeuge in ihrer natürlichen Arbeitsumgebung nutzbar. Dazu müssen die Versuchsfahrzeuge einen Prüfprozess in räumlich sowie zeitlich minimalen Abstand durchführen. In diesem Fall können die Veränderungen der Umgebungsbedingungen wie Fahrbahneigenschaften, Pflanzenbestand oder Schüttguteigenschaften zumindest teilweise von den Fahrzeugeinflüssen getrennt werden. Die Streuung aufgrund des Fahrzeugzustands bleibt hier bei den unsystematischen Anteilen. Grundsätzlich dient diese Methode einem Relativvergleich der Wirkung eines Einflussfaktors unter verschiedenen Rahmenbedingungen. Zur Vermeidung von sequencing Effekten ist dabei auf eine Randomisierung der Versuchsreihenfolge zu achten.

Durch die Gruppierung der Beobachtungswerte in Versuchsreihen mit nahezu gleichen Prüfbedingungen ergibt sich neben dem zu untersuchenden Einfluss  $A$  ein weiterer Einflussfaktor  $B$ . Folglich ist eine Auswertung mit Hilfe einer **zweifaktorielle Varianzanalyse** naheliegend. Die Wiederholungen innerhalb einer Versuchsreihe werden als unabhängige Stichproben behandelt. Ein signifikantes Ergebnis für  $A$  belegt die Wirkung des Einflussfaktors  $A$ . Wird die Nullhypothese für den Haupteffekt  $B$  verworfen, so bestätigt dies Unterschiede bei den Prüfbedingungen zwischen den verschiedenen Versuchsabschnitten. Eine signifikante Interaktion zwischen  $A$  und  $B$  bedeutet, dass der Faktor  $A$  je nach Prüfbedingungen die Zielgröße unterschiedlich stark beeinflusst.

Bei Verwendung der zweifaktoriellen Varianzanalyse bleibt jedoch unberücksichtigt, dass einer der beiden Einflussgrößen nicht gezielt sondern zufällig variiert wurde. BORTZ spricht in diesem Fall von einem *gemischten Modell*. Abweichend von einer Varianzanalyse mit gezielt variierten Ein-

flussfaktoren, wird hier der F-Bruch der Haupteffekte mit der mittleren Quadratsumme der Interaktion  $A \times B$  statt mit der mittleren Quadratsumme der Zellen im Nenner gebildet [BS10, S.248f.].

Dieses Auswerteverfahren entspricht der in Kapitel 4.3.7 erläuterten **Varianzanalyse mit Messwiederholungen**. Anstelle der individuellen Eigenschaften der Versuchsperson rücken in diesem Anwendungsfall die spezifischen Umweltbedingungen eines Untersuchungsabschnitts  $B_m$ . Dieser beinhaltet jeweils eine Versuchsreihe für jede Stufe des untersuchten Einflussfaktors  $A$ . Die Wiederholungen innerhalb einer Versuchsreihe können als Faktor  $C$  ebenfalls in der Analyse berücksichtigt werden. Das Auswerteverfahren entspricht somit einer zweifaktoriellen Varianzanalyse mit Messwiederholungen auf beiden Faktoren.

Das Hauptinteresse bei der Auswertung gilt dem Einflussfaktor  $A$ . Ein signifikantes Ergebnis für den Faktor  $C$  weist auf abweichende Entwicklungstrends der Beobachtungswerte innerhalb der Reihen hin. Mögliche Ursachen sind beispielsweise ein Trainingseffekt während der ersten Reihen oder starke Abweichungen der Systemtemperaturen in Relation zur Beharrungstemperatur. Die Interaktion  $A \times C$  würde auf eine Abhängigkeit zwischen dem Einflussfaktor  $A$  und der Anzahl an Wiederholungen hinweisen.

Voraussetzung dieser Varianzanalyse ist die Varianzhomogenität bei den Differenzen von jeweils zwei Treatmentstufen, die Sphärizitätsannahme. Wird der untersuchte Einflussfaktor  $A$  in nur zwei Stufen variiert, kann die Sphärizitätsbedingung nicht verletzt werden und es ist keine Korrektur der Freiheitsgrade erforderlich [BS10, S.300f.].

Eine alternative Betrachtungsweise eines Direktvergleichs ist mithilfe der **Kovarianzanalyse** möglich. Mittels linearer Regression sollen die Störeinflüsse, bedingt durch die Kovariate, entfernt werden. Allerdings ist die messtechnische Erfassung der zahlreichen Störeinflüsse mit einem Mehraufwand verknüpft und in vollem Umfang für jeden Versuch praktisch nicht möglich. Werden jedoch bei einem Direktvergleich die Beobachtungswerte der Referenzversuche gleichzeitig als Kovariate interpretiert, steht eine Messgröße zur Verfügung, welche die Wirkung aller Störeinflüsse des jeweiligen Untersuchungsabschnitts enthält. Die Voraussetzungen der Kova-

riananalyse sind weitestgehend erfüllt. Bei einer Interaktion zwischen den Einflussfaktoren und den Beobachtungswerten ist mit einer Ergebnisverzerrung zu rechnen. Grundsätzlich beschreibt BORTZ die Kovarianzanalyse bei homogenem Stichprobenumfang und normalverteilten Residuen jedoch als ausgesprochen robust gegen die Verletzung einzelner Bedingungen [BS10, S.311f.].

Dass aufgrund der zufällig variierten Prüfbedingungen ein gemischtes Modell vorliegt, kann in diesem Anwendungsfall nicht berücksichtigt werden. Da die Kovariate in jedem Direktvergleich nur einmalig erhoben wird, bleibt sie bei einer Kovarianzanalyse mit Messwiederholungen ohne Wirkung für den wiederholten Faktor. Die Untersuchungsergebnisse würden einer Varianzanalyse mit Messwiederholungen entsprechen [BS10, S.318]. Deshalb wird von einem Modell mit festen Effekten ausgegangen.

Ein weiteres Auswerteverfahren ergibt sich, wenn die Unterschiede aufgrund der Versuchsbedingungen  $B$  vollständig nivelliert werden. Dazu wird die Abweichungsquadratsumme zwischen den Gruppen  $A_l$  nicht bezüglich des Gesamtmittelwerts, sondern bezogen auf den Mittelwerten der jeweiligen Gruppe  $B_m$  gebildet.  $QS_{zw}$  wird folglich nur auf Ebene gleicher Prüfbedingungen ermittelt und berechnet sich gemäß Gleichung 4.82.

$$QS_{zw} = N \cdot \sum_{l=1}^L \sum_{m=1}^M (\bar{Y}_{lm\cdot} - \bar{Y}_{\cdot m\cdot})^2 \quad (4.82)$$

Dieser Ausdruck lässt sich nicht direkt bei den Termen der mehrfaktoriellen Varianzanalyse in Tabelle 4.6 wiederfinden. Vielmehr scheint Gleichung 4.82 ein Teil der Formel für die Abweichungsquadratsumme der Interaktion zwischen den Einflussfaktoren  $A$  und  $B$  zu sein. Tatsächlich lässt sich die Quadratsumme  $QS_{zw}$  (4.82) aus der Summe von  $QS_A$  und  $QS_{A \times B}$  bilden (Gleichung 4.83).  $QS_{zw}$  beinhaltet demzufolge alle Streuungsanteile in denen der Einflussfaktor  $A$  wirkt - sowohl den Haupteinfluss als auch die Interaktion mit den Versuchsbedingungen.



$$\begin{aligned}
\overbrace{N \cdot \sum_{l=1}^L \sum_{m=1}^M (\bar{Y}_{lm\cdot} - \bar{Y}_{\cdot m\cdot})^2}^{QS_{zw}} &= \overbrace{M \cdot N \cdot \sum_{l=1}^L (\bar{Y}_{l\cdot\cdot} - \bar{Y})^2}^{QS_A} \\
&+ \underbrace{N \cdot \sum_{l=1}^L \sum_{m=1}^M (\bar{Y}_{lm\cdot} - \bar{Y}_{l\cdot\cdot} - \bar{Y}_{\cdot m\cdot} + \bar{Y})^2}_{QS_{A \times B}}
\end{aligned} \tag{4.83}$$

Die Abweichungssumme innerhalb der Gruppen  $QS_{in}$  entspricht dem Streuungsanteil innerhalb der Zellen  $QS_R$ .

$$QS_{in} = QS_R = \sum_{l=1}^L \sum_{m=1}^M \sum_{n=1}^N (Y_{lmn} - \bar{Y}_{lm\cdot})^2 \tag{4.84}$$

Mit der Prüfgröße wird somit innerhalb der jeweiligen Bedingungen  $B_m$  die Varianz zwischen den Experimenten  $A_l$  an der Varianz innerhalb der Experimente relativiert. Diese wird als Quotient aus den mittleren Streuungen der in den Gleichungen 4.82 und 4.84 gebildeten Quadratsummen aufgestellt (Gleichung 4.87). Bei BORTZ taucht eine derartigen Teststatistik bei der Analyse einfacher Haupteffekte auf [BS10, S.259f.]. Die Freiheitsgrade der beiden Terme können aus den Indizes abgeleitet werden und sind in den Gleichungen 4.85 und 4.86 angegeben.

$$df_{zw} = L \cdot M - M = (L - 1) \cdot M \tag{4.85}$$

$$df_{in} = L \cdot M \cdot N - L \cdot M = L \cdot M \cdot (N - 1) \tag{4.86}$$

$$F = \frac{MQS_{zw}}{MQS_{in}} = \frac{\frac{QS_{zw}}{df_{zw}}}{\frac{QS_{in}}{df_{in}}} = \frac{\frac{N \cdot \sum_{l=1}^L \sum_{m=1}^M (\bar{Y}_{lm\cdot} - \bar{Y}_{\cdot m\cdot})^2}{(L - 1) \cdot M}}{\frac{\sum_{l=1}^L \sum_{m=1}^M \sum_{n=1}^N (Y_{lmn} - \bar{Y}_{lm\cdot})^2}{L \cdot M \cdot (N - 1)}} \tag{4.87}$$

Mit dieser Teststatistik kann geprüft werden, ob der Einflussfaktor  $A$  weder einen direkten noch einen indirekten Einfluss auf die Zielgröße hat.

Dementsprechend wird dieses Verfahren im Folgenden als **gekoppelte Varianzanalyse** bezeichnet. Die Erwartungswerte von Zähler und Nenner unterscheiden sich um den Term  $M \cdot N \cdot \sigma_A^2 + N \cdot \sigma_{A \times B}^2$ , der nur Null ist, falls die Nullhypothese gilt. Außerdem wächst dieser Term mit dem Grad der Verletzung der Nullhypothese. Damit sind die bei BORTZ beschriebenen Anforderungen an den Differenzterm zur Bildung der F-Teststatistik erfüllt [BS10, S.209f.].

Die Verwendung zufälliger Faktoren beziehungsweise die Abhängigkeit zwischen den Versuchsreihen innerhalb des Untersuchungsabschnitts wird auch bei diesem Auswerteverfahren nicht berücksichtigt. Formal ist damit eine der Voraussetzungen der Varianzanalyse nicht erfüllt. Inwieweit diese Verletzung die Gültigkeit der Wahrscheinlichkeiten und damit die Aussagekraft des Verfahrens beeinträchtigt, wurde nicht detailliert untersucht.

### Vergleich der Auswerteverfahren

Nachdem nun vier unterschiedliche Auswerteverfahren zur Analyse eines Direktvergleichs vorgestellt wurden, sollen diese anhand beispielhafter Datensätze mit der einfaktoriellen Varianzanalyse sowie untereinander verglichen werden. Die Basis dieser Untersuchungen bilden zwei Experimente mit jeweils sechs Versuchsreihen, bestehend aus je fünf Einzelversuchen. Zur Untersuchung einiger grundlegender Messergebniskonstellationen, wurde dieser Datensatz in mehreren Schritten manipuliert. Die Grafiken oben rechts in den Diagrammen von Abbildung 4.6 veranschaulichen die Art der Datenmanipulation.

Die Beobachtungswerte der Referenzmessung  $Y_{1mn}$  bleiben in allen Datenbeispielen unverändert. Für Abbildung 4.6a wurden die arithmetischen Mittel der Beobachtungswerte paarweise auf etwa den gleichen Wert gesetzt. Der Einflussfaktor  $A$  sollte folglich keinen Einfluss auf die Zielgröße haben. Die genauen Werte dieses Beispieldatensatzes sind im Anhang, Tabelle A.3 zu finden.

In 4.6b ist die Wirkung des Einflussfaktors  $A$  unterschiedlich stark ausgeprägt aber durchgehend positiv. Hier wurden die Daten der Vergleichsmessung mit der Rechenvorschrift  $l \cdot 0,6$  manipuliert.

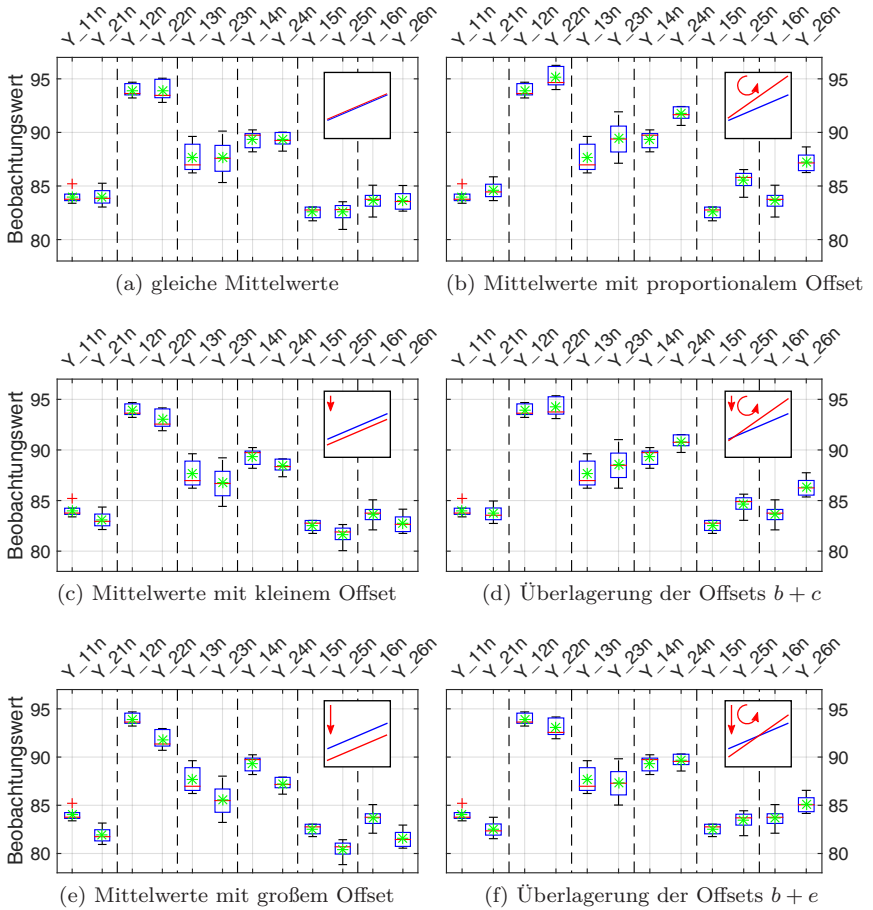


Bild 4.6: Beispieldaten zum Vergleich der Auswerteverfahren

Für die Datenbeispiele in 4.6c und 4.6e wurden die Werte der Vergleichsmessung um ein Offset von  $-0,9$  beziehungsweise  $-2,1$  verschoben. Folglich hat  $A$  hier unter allen Versuchsbedingungen  $B_m$  einen konstanten Einfluss. In den Datensätzen (d) und (f) wurden die beiden Manipulationsarten überlagert. Während in 4.6d die Ausprägung der Abweichung lediglich abgeschwächt wird, zeigt die Einflussgröße  $A$  im Beispiel 4.6f sowohl positive als auch negative Einflüsse auf die Beobachtungswerte.

Tabelle 4.9 gibt die Analyseergebnisse dieser Beispieldatensätze wider. Da im vorliegenden Anwendungsfall die Prüfbedingungen als zufälliger Einflussfaktor eingehen, liegt laut Literatur eindeutig ein gemischtes Modell vor, das mit einer Varianzanalyse mit Messwiederholungen auszuwerten ist. Die Ergebnisse dieses Verfahrens gelten folglich als Referenz.

Tabelle 4.9: Fehlerwahrscheinlichkeiten für die Ablehnung der Nullhypothese bei verschiedenen Auswerteverfahren

Auswerteverfahren	Bsp.(a)	Bsp.(b)	Bsp.(c)	Bsp.(d)	Bsp.(e)	Bsp.(f)
<b>einfaktorielle Varianzanalyse</b>	99,99 %	4,53 %	40,18 %	24,71 %	5,35 %	99,99 %
<b>einfaktorielle Varianzanalyse mit Messwiederholungen</b>	93,58 %	0,59 %	0 %	4,73 %	0 %	99,98 %
<b>zweifaktorielle Varianzanalyse <math>A</math></b>	99,96 %	0 %	0,13 %	0 %	0 %	99,96 %
<b>zweifaktorielle Varianzanalyse <math>A \times B</math></b>	100 %	1,89 %	100 %	1,89 %	100 %	1,89 %
<b>einfaktorielle Kovarianzanalyse</b>	93,5 %	0,14 %	0 %	2,86 %	0 %	99,98 %
<b>gekoppelte Varianzanalyse</b>	100 %	0 %	9,32 %	0,01 %	0 %	3,36 %

Alle Auswerteverfahren bestätigen bei einem Signifikanzniveau von  $\alpha = 5\%$  für das Datenbeispiel (a) die Nullhypothese. Wie erwartet bewirkt der Einflussfaktor  $A$  hier keine signifikanten Unterschiede. Ebenso wird von allen Verfahren im Beispiel (b) ein signifikanter Einfluss durch  $A$  erkannt. Der einfaktoriellen Varianzanalyse gelingt das jedoch nur knapp. Bei einer größeren Streuung zwischen den Versuchsbedingungen respektive Faktorstufen  $B_m$  würde die ANOVA zu einem anderen Ergebnis kommen. Eine derartige Abhängigkeit besteht bei allen Datensätzen und ist für die Analyseergebnisse unerwünscht.

Für die Datensätze (c) bis (f) kommen die verschiedenen Auswerteverfahren zu abweichenden Ergebnissen. Ein kleiner Offset über alle Stufen des Einflussfaktors  $B$  wird von der Varianzanalyse mit Messwiederholungen, dem Haupteffekt der zweifaktoriellen Varianzanalyse sowie der Kovarianzanalyse als hoch signifikant bewertet. Die anderen Verfahren lehnen die Nullhypothese dagegen nicht ab. Ein Blick auf Abbildung 4.6 zeigt, dass die Mittelwerte des manipulierten Experiments ausnahmslos eine ähnliche, negative Abweichung haben. Gleichzeitig sind diese Abweichungen aber noch innerhalb der Quartile der Referenzbeobachtungen. Subjektiv betrachtet, könnte die beobachtete Abweichung auch lediglich auf der Streuung innerhalb einer Faktorstufe  $B_m$  beruhen. Bei größeren und damit eindeutigen Offsets lehnt auch die gekoppelte Varianzanalyse die Nullhypothese zuverlässig ab. Aufgrund der exakt parallelen Verschiebung der Beobachtungswerte, kann für die Datensätze (a), (c) und (e) keine Interaktion zwischen  $A$  und  $B$  durch die zweifaktorielle Varianzanalyse festgestellt werden.

Die Überlagerung von konstant negativer Abweichung und ansteigendem Offset in den Beispielen (d) und (f) führt zu einer Reduktion der Abweichungsquadratsummen. Die Varianzanalyse mit Messwiederholungen sowie die Kovarianzanalyse verlieren erkennbar an Trennschärfe. Obwohl in Abbildung 4.6d der Einfluss des Faktors  $A$  offensichtlich ist, wird die Nullhypothese nur knapp abgelehnt. Hier urteilen diese Verfahren recht konservativ. Im Datenbeispiel 4.6f sind positive und negative Abweichungen etwa ausgeglichen. Die Mittelwerte der Faktorstufen  $A$  sind dadurch nahezu identisch. Ein signifikanter Einfluss von  $A$  kann nur mit der gekoppelten Varianzanalyse oder durch die gezielte Analyse der Interaktion mit der zweifaktoriellen Varianzanalyse festgestellt werden. Bei der Interpretation dieser Ergebnisse ist zu klären, inwieweit die Variation des Faktors  $B$  tatsächlich eine Umkehrung des Effektes von  $A$  bewirkt. Hier könnten auch Fehler in der Versuchsdurchführung oder Sequencingeffekte das Ergebnis beeinflussen.

Mithilfe dieser Beispieldatensätze wurden die Stärken und Schwächen der verschiedenen Auswerteverfahren veranschaulicht. Aufgrund der starken Unterschiede zwischen den Faktorstufen  $B_m$  erwies sich die einfaktoriel-

le Varianzanalyse bei diesen Beispielen als weitestgehend ungeeignet. Die zweifaktorielle Varianzanalyse bietet durch die Aufschlüsselung in Haupteffekt und Interaktion den höchsten Informationsgehalt. Da die Abhängigkeit der Beobachtungswerte nicht berücksichtigt wurde, wird die Nullhypothese teilweise zu Unrecht abgelehnt ( $\alpha$ -Fehler). Die Kombination beider Untersuchungen in der gekoppelten Varianzanalyse besitzt eine schwächere Trennschärfe und liefert mitunter zu konservative Urteile ( $\beta$ -Fehler). Dafür gibt es ein zusammengefasstes Analyseergebnis, mit dem sowohl direkte als auch indirekte Einflüsse von  $A$  identifiziert werden können. Diese Eigenschaft wird sich noch bei der Analyse von Datenverläufen bewähren. Die Ergebnisse der Kovarianzanalyse weichen nur geringfügig von den Wahrscheinlichkeiten der Varianzanalyse mit Messwiederholungen ab. Weil die Bedingungen der ANCOVA unter Umständen verletzt werden und diese keinen deutlichen Mehrwert liefert, ist eine derartige Verwendung nicht zielführend.

Die zufällige Variation des Einflussfaktors  $B$  wird nur in der Varianzanalyse mit Messwiederholungen berücksichtigt. Aus Sicht der Statistik ist dieser Sachverhalt unbedingt mit einzubeziehen [BS10, S.245-249]. Bei einer quantitativ einheitlichen Wirkung des Einflussfaktors zeigt das Verfahren eine besonders hohe Trennschärfe. Ist der Einfluss von  $A$  dagegen mal mehr, mal weniger ausgeprägt oder wechselt sogar zeitweise das Vorzeichen urteilt die Varianzanalyse mit Messwiederholungen wesentlich konservativer. Dessen ungeachtet ist die Varianzanalyse mit Messwiederholungen das geeignetste Verfahren für die Auswertung von Direktvergleichen.

Die beschriebene Charakteristik der Varianzanalyse mit Messwiederholungen ist unter weiteren Untersuchungsaspekten sehr nützlich. Bis jetzt war die oberste Prämisse, Prüfprozess und Prüfbedingungen möglichst nicht zu verändern. Damit sind Aussagen zur Wirkung eines Einflussfaktors genau genommen auf diesen Prozess und diese Bedingungen beschränkt. Zur Erhöhung der Aussagekraft ist es demzufolge erstrebenswert, die Wirksamkeit einer Veränderung unter möglichst vielen Arbeitsprozessen und Arbeitsbedingungen in Direktvergleichen zu untersuchen. Ein signifikantes Ergebnis ist nur zu erwarten, falls der Einflussfaktor bei allen Prüfungen und damit über ein breites Einsatzspektrum eine ähnliche Wirkung zeigt.

Erzeugt die Veränderung dagegen sowohl positive als auch negative Effekte auf die Zielgröße, bleibt das Ergebnis nicht signifikant und der Nutzen der Veränderung auf die Zielgröße ist in Frage zu stellen. Können diese Bedingungen gezielt wiederholt eingestellt werden, so kann die Auswertung mithilfe der zweifaktoriellen Varianzanalyse erfolgen. Falls sich diese Bedingungen jedoch zufällig ergeben, liegt ein gemischtes Modell vor, wodurch eine Auswertung als Direktvergleich erforderlich ist.

#### 4.4.5 Erzielbare Schärfe bei Direktvergleichen

Zur Beurteilung der Prüfprozesseignung soll Anhand einiger Referenzversuche die erzielbare Trennschärfe für weitere Direktvergleiche abgeschätzt werden. Grundsätzlich ist ein Vorgehen analog zu Kapitel 4.4.3 möglich. Die Teststatistik 4.88 wird für  $L = 2$  nach dem Abstand zwischen den Experimenten aufgelöst, wobei der Nenner des F-Bruchs auf Basis von Referenzversuchen geschätzt wird. Zur Abschätzung der erzielbaren Schärfe bei Direktvergleichen ergibt sich somit Gleichung 4.89.

$$F_A = \frac{MQS_A}{MQS_{A \times B}} = \frac{M \cdot N \cdot \sum_{l=1}^L (\bar{Y}_{l..} - \bar{Y})^2}{L - 1} \cdot \frac{N \cdot \sum_{l=1}^L \sum_{m=1}^M (\bar{Y}_{lm.} - \bar{Y}_{l..} - \bar{Y}_{.m.} + \bar{Y})^2}{(L - 1) \cdot (M - 1)} \quad (4.88)$$

$$\pm |\bar{Y}_{1..} - \bar{Y}_{2..}| = \sqrt{\frac{F_{1-\alpha}(1, M-1)}{M \cdot N} \cdot 2 \cdot MQS_{A \times B}} \quad (4.89)$$

Als Referenzversuche für die Abschätzung sollten halb so viele Versuchsreihen mit jeweils doppelt so vielen Einzelversuchen als im eigentlichen Prüfprozess durchgeführt werden. Aus der Forderung nach mindestens 25 Einzelversuchen pro Experiment ergeben sich beispielsweise  $3 \times 10$  Einzelversuche für eine Abschätzung und jeweils mindestens  $6 \times 5$  Versuche pro Faktorstufe für den Prüfprozess. Um tatsächlich die verschiedensten Störeinflüsse wirken zu lassen, sollten diese drei Versuchsreihen nicht unmittelbar hintereinander, sondern beispielsweise an verschiedenen Tagen durchgeführt werden. Für diese Referenzversuche wird die mittlere Abwei-

chungsquadratsumme der Interaktion  $A \times B$  berechnet. Dazu wird jeweils die erste Hälfte innerhalb der Versuchsreihen der Faktorstufe  $A_1$ , der Rest  $A_2$  zugeordnet. Zur Berücksichtigung möglicher Sequencingeffekte werden alle Beobachtungswerte mit umgekehrter Zuordnung zu den Stufen von  $A$  gedoppelt. Für  $B$  werden dabei neue Faktorstufen angenommen, da sonst gar keine Interaktion auftreten würde. Mit dieser Stufenzuordnung, veranschaulicht in Tabelle 4.10, ergeben sich 12 Versuchsreihen mit jeweils  $N = 5$  Einzelversuchen.

Tabelle 4.10: Faktorzuordnung zur Abschätzung der erzielbaren Schärfe von Direktvergleichen

Beobachtungswert	Faktor $A$	Faktor $B$	Faktor $C$
$Y_{111}$	1	1	1
$Y_{112}$	1	1	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Y_{115}$	1	1	5
$Y_{116}$	$1 \Rightarrow 2$	1	$6 \Rightarrow 1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Y_{1110}$	$1 \Rightarrow 2$	1	$10 \Rightarrow 5$
$Y_{121}$	1	2	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Y_{1310}$	$1 \Rightarrow 2$	3	$10 \Rightarrow 5$
.....			
$Y_{111}$	$1 \Rightarrow 2$	$1 \Rightarrow 4$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Y_{1310}$	1	$3 \Rightarrow 6$	$10 \Rightarrow 5$

Beispielhaft wird die erzielbare Schärfe für die Datenbeispiele in Abbildung 4.6 berechnet. Dabei wird davon ausgegangen, dass der Einflussfaktor  $A$  für keinen Versuch variiert wurde und die Abweichungen innerhalb einer Prüfbedingung durch Störeinflüsse verursacht werden. Wie zuvor erläu-



tert, werden für dieses Beispiel nur die jeweils ersten 30 Versuche eines Datensatzes berücksichtigt.

Tabelle 4.11: Abschätzung der erzielbaren Schärfe für die Beispieldaten mit  $M = 1$  bis 3 unter der Annahme  $A_1 = A_2$

Datensatz	Bsp.(a)	Bsp.(b)	Bsp.(c)	Bsp.(d)	Bsp.(e)	Bsp.(f)
erzielbare Schärfe	$\pm 0,0096$	$\pm 3,97$	$\pm 1,63$	$\pm 1,76$	$\pm 3,79$	$\pm 1,85$

Eine Abschätzung der erzielbaren Schärfe anhand bestehender Daten eines Direktvergleichs ist ebenfalls möglich. Dazu muss sichergestellt sein, dass die Versuchsreihenfolge bei der Hälfte der Reihen umgekehrt wurde. Für die vorliegenden Datensätze aus Abbildung 4.6 wird die erzielbare Schärfe in Tabelle 4.12 abgeschätzt. Diesen Ergebnissen kann entnommen werden, dass die erzielbare Schärfe unabhängig vom Haupteffekt lediglich von der Streuung der Differenzen, der Interaktion, abhängt. Bei Verwendung eines Referenzdirektvergleichs mit variiertem Einflussfaktor  $A$  bleibt allerdings unklar, ob die Streuung durch die Faktorstufen von  $A$  oder die Varianz innerhalb der Reihen bewirkt wird.

Tabelle 4.12: Abschätzung der erzielbaren Schärfe für die gesamten Beispieldaten aus Abbildung 4.6 mit  $A_1 \neq A_2$

Datensatz	Bsp.(a)	Bsp.(b)	Bsp.(c)	Bsp.(d)	Bsp.(e)	Bsp.(f)
erzielbare Schärfe	$\pm 0,0099$	$\pm 3,15$	$\pm 0,0058$	$\pm 3,15$	$\pm 0,0058$	$\pm 1,85$

Umgekehrt wird bei der Abschätzung der Schärfe mit Hilfe von  $3 \cdot 10$  Referenzversuchen angenommen, dass es keine Interaktion zwischen dem Einflussfaktor  $A$  und den Prüfbedingungen gibt. Allgemein beruht die Abschätzung also auf der Annahme, dass Interaktionsverhalten bleibe bei den folgenden Direktvergleichen gleich oder werde sogar kleiner. Damit kleine Veränderungen bei der Interaktion nicht unmittelbar die Beurteilung der Prüfprozesseignung widerlegen, wird wie zuvor mit einem Sicherheitsfaktor beim Signifikanzniveau  $\alpha$  gearbeitet. Standardmäßig wird hier empfohlen  $\alpha$  für die Abschätzung auf ein Fünftel zu reduzieren. Varianzinhomogenität zwischen den Reihen ist ein Indiz dafür, dass der Einflussfaktor  $A$  oder

die Prüfbedingungen  $B$  unter einzelnen Faktorkombinationen eine abweichende Wirkung zeigen. Aufgrund der größeren Unsicherheit sollte hier ein Fünfzigstel von  $\alpha$  eingesetzt werden. Für die Berechnung der erzielbaren Schärfe in den Tabellen 4.11 und 4.12 wurde die Varianzhomogenität bereits geprüft und ein entsprechender Sicherheitsfaktor verwendet. Dieser Sicherheitsfaktor bedingt eine konservativere Abschätzung der Beispiele  $a$ ,  $b$  und  $d$ .

Resultiert aus der Analyse eines konkreten Direktvergleiches ein nicht signifikantes Ergebnis, sollte geprüft werden, ob bei der vorliegenden Interaktion die geforderte Trennschärfe erreicht werden kann. Ist dies nicht der Fall, muss der Prüfprozess für diese Versuchskonstellation nachträglich als ungeeignet beurteilt werden.

Zusammenfassend gilt: Je geringer die Interaktion, desto höher die erreichbare Schärfe des Prüfprozesses. Ist die Wirkung des Einflussfaktors dagegen mal mehr, mal weniger ausgeprägt oder wechselt sogar das Vorzeichen, bleibt das Verfahren sehr liberal.

#### **4.4.6 Auswerteverfahren für transiente Versuche**

In Kapitel 3.2 ist beschrieben, dass bei einem transienten Prüfprozess lediglich Anfangs- und Endzustand oder aber ein fester Bewegungs- oder Geschwindigkeitsverlauf vorgegeben sein kann. Da ein derartiger Verlauf nicht immer automatisiert, sondern von einem Maschinenbediener vorgegeben wird, unterliegt auch diese Vorgabe einer Streuung. Um nun sicherzustellen, dass ein signifikanter Unterschied bei der Zielgröße nicht nur aus einer Abweichung der Eingangswerte resultiert, sollte auch deren Werteverlauf analysiert werden. Im besten Fall zeigen die Eingangswerte nachweislich keine signifikanten Unterschiede zwischen den Experimenten. Somit lässt sich eine Abweichung bei der Zielgröße klarer dem gezielt manipulierten Einflussfaktor zuordnen. Wird bei der Definition des Prüfprozesses ein Toleranzband um den Sollwerteverlauf festgelegt, können sich die Eingangswerte dennoch signifikant unterscheiden. Zweckmäßiger ist es deshalb eine zulässige Varianz zum Sollwerteverlauf zu definieren.

Ein weiterer Anwendungsfall für eine Analyse des Messdatenverlaufs auf Streuung und signifikante Abweichungen besteht für die Übertragung eines

Prüfzyklus auf einen Prüfstand. Hier sollte überprüft werden, ob sowohl Einfluss- als auch Zielgrößen dem ursprünglichen Verlauf entsprechen und demzufolge die realen Lasten korrekt durch den Prüfstand simuliert werden.

Zur Auswahl eines geeigneten Auswerteverfahrens für die statistische Auswertung von Messdatenverläufen, werden die Randbedingungen aus statistischer Sicht zusammengetragen. Der zu untersuchende Einflussfaktor  $A$  wird gezielt in  $L \geq 2$  Stufen variiert. Zum Vergleich der Messdatenverläufe, müssen diese in gleicher zeitlicher Diskretisierung vorliegen und an einem charakteristischen Punkt des Versuchs synchronisiert werden. Eine Synchronisation auf den minimalen Kurvenabstand ist ebenfalls denkbar, wird jedoch bei mehr als zwei Kurven schnell sehr komplex, wenn Abhängigkeiten der Reihenfolge oder der Bezugskurve vermieden werden sollen. Die Zeit wird bei der Analyse als Einflussfaktor  $T$  verstanden. Dabei gilt jeder Messzeitpunkt seit Versuchsstart als Faktorstufe  $T_p$ . Für jeden Einzelversuch stehen Beobachtungswerte aus  $P$  Messzeitpunkten zur Verfügung. Diese Beobachtungswerte zeigen Abhängigkeiten zu vorangegangenen Werten des Einzelversuchs. Derartige Wechselbeziehungen werden als Autokorrelation bezeichnet [Sto18, Kap. 9]. Wiederholungen eines Versuchs mit gleicher Faktorstufe  $A_l$  werden als  $n$ -te Stufe des zufälligen Einflusses  $C$  berücksichtigt. Die etwaige Zuordnung zu Versuchsreihen wird auf Betrachtungsebene der Experimente ignoriert.

Demzufolge besitzen die Beobachtungswerte eines Messdatenverlaufs sowohl feste als auch zufällige Faktoren. Es wird somit von einem gemischten Modell gesprochen [BS10, A.245]. Insofern wäre es naheliegend wie zuvor bei den Direktvergleichen die Varianzanalyse mit Messwiederholungen als Auswerteverfahren zu nutzen. Anstelle der Versuchsbedingungen der Reihen rücken hier die individuellen Bedingungen infolge des Versuchsfortschritts. Die Varianz aufgrund der verschiedenen Rahmenbedingungen zwischen den Reihen wird damit wieder Teil der unsystematischen Einflüsse.

Allerdings wird in der Anwendung hier der untersuchte Einflussfaktor nicht zwischen den abhängigen Beobachtungswerten variiert. Außerdem wurde in den Beispielen aus Abbildung 4.6 festgestellt, dass die Varianzanalyse

mit Messwiederholungen bei Wechselwirkungen zwischen dem Einflussfaktor und den Bedingungen erheblich an Trennschärfe verliert. Derartige Interaktionen sind jedoch bei einem Messdatenverlauf sehr wahrscheinlich. Zur Verdeutlichung werden in Abbildung 4.7 die Datenbeispiele 4.6d und 4.6f als kurze Messsequenzen mit jeweils 6 Messpunkten betrachtet. Obwohl der Faktor  $A$  offensichtlich einen abweichenden Verlauf der Zielgröße verursacht, liefert die Varianzanalyse mit Messwiederholungen nur knapp ein beziehungsweise kein signifikantes Ergebnis (vgl. Tabelle 4.9). Lediglich die Auswertung bezüglich Interaktionen mit der zweifaktoriellen Varianzanalyse sowie die gekoppelte Varianzanalyse führt zu einer korrekten Bewertung beider Datenverläufe in Abbildung 4.7.

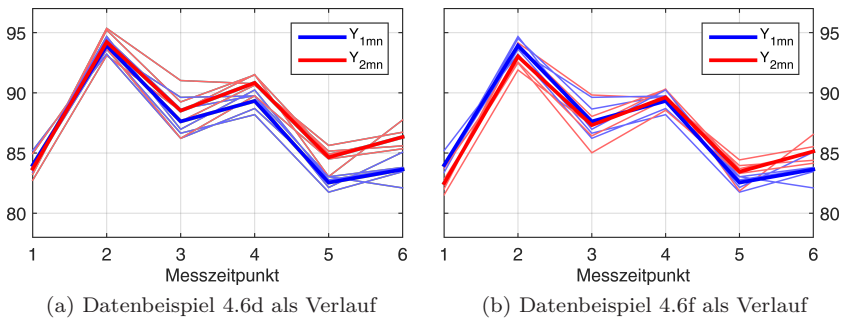


Bild 4.7: Datenverläufe zum Vergleich der Auswerteverfahren

Mit der zweifaktoriellen Varianzanalyse werden Haupteffekt und Interaktion separat betrachtet und ausgewertet. Die gekoppelte Varianzanalyse verbindet diese beiden Auswertungen zumeist auf Kosten der Trennschärfe. Wird die Autokorrelation zwischen den Beobachtungswerten des Datenverlaufs ignoriert, liefern beide Verfahren aufgrund der zahlreichen Stichproben und entsprechenden Freiheitsgrade eines Messdatenverlaufs eine offensichtlich zu hohe Trennschärfe. Es würden praktisch immer signifikante Unterschiede zwischen den Datenverläufen herauskommen. Zur Korrektur der Ergebnisverzerrung der Varianzanalyse müssen demzufolge die Zähler- und Nennerfreiheitsgrade der Prüfstatistik auf die Anzahl unabhängiger Werte reduziert werden. Dies soll mit dem Faktor  $\varepsilon$  erfolgen.

Die Ausprägung der Autokorrelation eines Zeitverlaufs lässt sich mithilfe der empirischen Autokorrelationsfunktion in Gleichung 4.90 ermitteln [KN06, S.31f.].

$$\hat{\varrho}_{ln}(h) = \frac{\hat{\gamma}_{ln}(h)}{\hat{\gamma}_{ln}(0)} = \frac{\frac{1}{P-h} \sum_{p=1}^{P-h} (Y_{ln\ p+h} - \bar{Y}_{ln.})(Y_{ln\ p} - \bar{Y}_{ln.})}{\frac{1}{P} \sum_{p=1}^{P-h} (Y_{ln\ p+0} - \bar{Y}_{ln.})(Y_{ln\ p} - \bar{Y}_{ln.})} \quad (4.90)$$

für  $h = 1, 2, \dots$

Hier müsste nun für jede Versuchsreihe und jede Zeitverzögerung  $h$  die Autokorrelation berechnet und der größte Wert  $h$  bestimmt werden, bei dem noch Autokorrelation auftritt.  $\varepsilon_T$  ergibt sich anschließend als Kehrwert von  $\max(h)$ . Zugunsten des Berechnungsaufwandes, der Komplexität und der Robustheit des Auswerteverfahrens wurde dieser Ansatz jedoch nicht weiter verfolgt. Stattdessen wird  $h$  direkt auf den Maximalwert  $P$  gesetzt und damit sämtliche Freiheitsgrade aus dem Datenverlauf entfernt (Gleichung 4.91).

$$\varepsilon_T = \frac{1}{P} \quad (4.91)$$

Ein ähnliches Korrekturverfahren wird von GREENHOUSE und GEISSER für die Varianzanalyse mit Messwiederholungen vorgeschlagen, falls die Zirkularitätsannahme verletzt ist [RFHN06, S.110f.]. Bei Verwendung der adjustierten Freiheitsgrade ergeben sich größere F-Werte beziehungsweise höhere Fehlerwahrscheinlichkeiten. Die Nullhypothese wird häufiger beibehalten.

In der Anwendung mit der zweifaktoriellen Varianzanalyse zeigt sich, dass sich für  $F_A$  mit diesem Korrekturverfahren Zählerfreiheitsgrade weit unter eins, nahe Null ergeben. Die Berechnung des F-Wertes in diesem Bereich erweist sich als unsinnig, das Korrekturverfahren somit als ungeeignet. Bei der gekoppelten Varianzanalyse gibt es dagegen genug Freiheitsgrade um mit dieser Korrektur zu arbeiten.

Weitere Voraussetzungen der Varianzanalyse sind Normalverteilung und Varianzhomogenität der Eingangswerte. Da die Einflussfaktoren nicht nur den Erwartungswert sondern teilweise auch die Varianz beeinflussen, liegt

bei Fahrzeugversuchen nicht immer Varianzhomogenität vor. Ebenso ist aufgrund des Werteverlaufs eine Normalverteilung aller Beobachtungswerte unwahrscheinlich. Für die gekoppelte Varianzanalyse ist das von untergeordneter Bedeutung. Da die Varianzen bezüglich des durchschnittlichen Datenverlaufs anstatt des Gesamtmittelwertes gebildet werden. Bildlich werden die Beobachtungswerte um den Mittelwert des jeweiligen Untersuchungsabschnitts verschoben und somit die Niveauabweichungen aufgrund des Versuchsfortschritts entfernt. Anstatt um die Durchschnittswerte jedes Untersuchungsabschnitts streuen die Messwerte jetzt um die Abszisse. Aufgrund des gemeinsamen Mittelwertes weichen diese Residuen meistens nur wenig von einer Normalverteilung ab. Selbst bei Verletzung dieser beiden Voraussetzungen zeigt sich die Varianzanalyse laut BRAY und MAXWELL [Bra85, S.34] sowie PERREAULT und DARDEN [PD75, S.334] bei umfangreichen, ähnlich großen Testgruppen als sehr robust.

Auf Basis dieser Überlegungen erweist sich die gekoppelte Varianzanalyse als geeignetes Auswerteverfahren für die Analyse von Messdatenverläufen. Nur hier können alle Effekte, sowohl Haupteffekte als auch Wechselwirkungen, erkannt und gleichzeitig die Abhängigkeit zwischen den Beobachtungswerten berücksichtigt werden. Weiter rückt durch die Adjustierung der Freiheitsgrade die Signifikanzgrenze in einen angemessenen Bereich. Einzelne sowie kleinste Abweichungen werden noch akzeptiert, während bedeutsame Effekte zuverlässig erkannt werden. Die Genauigkeit der ermittelten Fehlerwahrscheinlichkeit wurde nicht weiter untersucht. Aufgrund des konservativen Korrekturverfahrens mit der Entfernung aller Freiheitsgrade des Datenverlaufs ist die tatsächliche Fehlerwahrscheinlichkeit kleiner als berechnet. Signifikante Ergebnisse sind somit sehr zuverlässig.

In Gleichung 4.92 ist die Prüfgröße sowie die Ablehnungsbedingung der gekoppelten Varianzanalyse für die Auswertung von Datenverläufen dokumentiert. Dieses Auswerteverfahren lässt sich sowohl auf Ebene der Versuchsreihen als auch im Vergleich verschiedener Experimente anwenden. Konsequenz bei signifikanten Abweichungen zwischen den Reihen ist die Reduzierung des Signifikanzniveaus  $\alpha$  bei der Abschätzung der erzielbaren Trennschärfe des Prüfprozesses.

$$\frac{N \cdot \sum_{l=1}^L \sum_{p=1}^P (\bar{Y}_{l \cdot p} - \bar{Y}_{\cdot p})^2}{(L-1) \cdot P} > F_{1-\alpha}((L-1) \cdot \mathcal{R}, L \cdot \mathcal{R} \cdot (N-1)) \quad (4.92)$$

$$\frac{\sum_{l=1}^L \sum_{n=1}^N \sum_{p=1}^P (Y_{lnp} - \bar{Y}_{l \cdot p})^2}{L \cdot P \cdot (N-1)}$$

Aufgrund der großen Anzahl an Beobachtungswerten ergeben sich relativ zuverlässige Varianzen für jeden Versuch. Die Anwendung des Kruskal-Walis-Test ist somit nicht erforderlich. Von einer Reduzierung der Versuchsanzahl wird dennoch abgeraten. Zum einen erhöht sich sonst das Risiko die Einflüsse durch Störgrößen nicht hinreichend zu erfassen. Zum anderen wird durch die geringere Versuchsanzahl die Prüfgröße geschwächt. Bei der Bestimmung des  $(1 - \alpha)$ -Quantil der F-Verteilung bleibt die große Anzahl an Beobachtungswerten wegen ihrer Abhängigkeit unberücksichtigt.

#### 4.4.7 Erzielbare Schärfe bei transienten Versuchen

Wie zuvor bei der Auswertung mit nur einem Beobachtungswert pro Versuch soll auch hier die erzielbare Schärfe abgeschätzt werden. Das Ergebnis ist der mittlere Kurvenabstand zwischen zwei Experimenten ab dem eine statistisch signifikante Unterscheidung möglich ist. Die Annahmen der Abschätzung sind eine identische Versuchsanzahl und Länge sowie eine übereinstimmende Streuung der Datenverläufe innerhalb eines Experiments beziehungsweise des Prüfprozesses.

Analog zur Abschätzung in Kapitel 4.4.3 wird hierfür die Streuung des Prüfprozesses benötigt. Dazu wird die mittlere Abweichungsquadratsumme zwischen den Messwerten und dem durchschnittlichen Datenverlauf innerhalb der Referenzexperimente gemäß Gleichung 4.93 ermittelt. Für Aussagen zur Wiederholbarkeit wird die größte aufgetretene Experimentstreuung verwendet.

$$MQS_{Ref} = \max \{MQS_1, \dots, MQS_L\} = \frac{\sum_{n=1}^N \sum_{p=1}^P (Y_{lnp} - \bar{Y}_{..p})^2}{P \cdot (N - 1)} \quad (4.93)$$

Wird mit der Abschätzung die Reproduzierbarkeit beurteilt, so ist die korrigierte Gesamtstreuung entsprechend Gleichung 4.77 einzusetzen. Hierin sind die Streuungen aufgrund der zulässigen Versuchsbedingungen sowie die maximal gemessene Experimentstreuung berücksichtigt. In beiden Fällen fließt die Streuung der Referenzexperimente als mittlere Abweichungsquadratsumme  $MQS_{Ref}$  in Gleichung 4.94 ein. Zur Herleitung dieser Gleichung wird wie bereits in Kapitel 4.4.3 die zugehörige Teststatistik für zwei Experimente aufgestellt und nach dem Abstand der Mittelwerte aufgelöst. Die Annahmen einheitlicher Experimentumfänge und Streuungen führt zu weiteren Vereinfachungen. Die resultierende Gleichung ist nahezu identisch mit 4.81 aus der einfaktoriellen Varianzanalyse. Allerdings fließen hier wesentlich mehr Messdaten bei der Berechnung der Streuung ein.

$$\pm |\bar{Y}_{1..} - \bar{Y}_{2..}| = \sqrt{\frac{F_{1-\alpha}(1, 2(N-1))}{N} \cdot 2 MQS_{Ref}} \quad (4.94)$$

Den Unsicherheiten bei der Erfassung der möglichen Störeinflüsse wird, wie bereits in Kapitel 4.4.3 erläutert, mit einem Faktor auf das Signifikanzniveau  $\alpha$  Rechnung getragen. Bei signifikanten Abweichungen zwischen den Versuchsreihen innerhalb eines Experiments wird  $\alpha$  auf ein Fünfstel reduziert, andernfalls nur auf ein Fünftel.

#### 4.4.8 Direktvergleiche mit transienten Versuchen

Das Konzept des Direktvergleichs zur Berücksichtigung der Störeinflüsse bei den Versuchsbedingungen lässt sich auch auf transiente Versuche anwenden. Dazu werden die Varianzen innerhalb jeweils gleicher Versuchsbedingungen berechnet. Anstatt an der Varianz innerhalb einer Zelle wird die mittlere Abweichungsquadratsumme der Messzeitpunkte an der Interaktion mit den Versuchsbedingungen relativiert. Die Teststatistik wird folglich entsprechend Gleichung 4.95 gebildet.



$$F = \frac{M \cdot N \cdot \sum_{l=1}^L \sum_{p=1}^P (\bar{Y}_{l..p} - \bar{Y}_{...p})^2}{(L-1) \cdot P} \quad (4.95)$$

$$\frac{N \cdot \sum_{l=1}^L \sum_{m=1}^M \sum_{p=1}^P (\bar{Y}_{lm..p} - \bar{Y}_{l..p} - \bar{Y}_{..m..p} + \bar{Y}_{...p})^2}{(L-1) \cdot (M-1) \cdot P}$$

$$F > F_{1-\alpha} [(L-1) \cdot \mathcal{R}, (L-1) \cdot (M-1) \cdot \mathcal{R}] \quad (4.96)$$

Bewirken die verschiedenen Versuchsbedingungen nur kleine Störeinflüsse, bringt dieses Auswerteverfahren gegenüber der gekoppelten Varianzanalyse ohne Wiederholungen keine Vorteile. Die Abschätzung der erzielbaren Schärfe erfolgt analog zur Beschreibung in Kapitel 4.4.5.  $MQS_{A \times B}$  wird lediglich durch den angepassten Nenner  $MQS_{in}$  der Prüfgröße 4.95 ersetzt.

$$\pm |\bar{Y}_{1..p} - \bar{Y}_{2..p}| = \sqrt{\frac{F_{1-\alpha}(1, M-1)}{M \cdot N} \cdot 2 \cdot MQS_{in}} \quad (4.97)$$

Für  $\alpha$  wird wie zuvor ein Fünftel des gewählten Signifikanzniveaus eingesetzt. Wird zwischen den transformierten Versuchsreihen Varianzheterogenität festgestellt, sollte  $\alpha$  auf ein Fünfzigstel reduziert werden.

### Kapitelzusammenfassung

In diesem Kapitel wurden verschiedene statistische Auswerteverfahren in Bezug auf Fahrzeugversuche vorgestellt. Je nach Fragestellung und Rahmenbedingungen der Untersuchung ist ein anders Auswerteverfahren aus Abbildung 4.8 am zweckmäßigsten. Die Skizzen veranschaulichen zu jedem Auswerteverfahren eine typische Kombination aus Untersuchungsgröße und Störeinflüssen. Für Sonderfälle finden sich in der Literatur gegebenenfalls weitere Verfahren.

Durch Auflösen der Teststatistik nach den mittleren Experimentunterschieden kann anhand von Referenzversuchen der kleinstmögliche, differenzierbare Effekt abgeschätzt und damit die Eignung des Prüfprozesses beurteilt werden. Den Gleichungen 4.81 und 4.94 ist zu entnehmen, dass

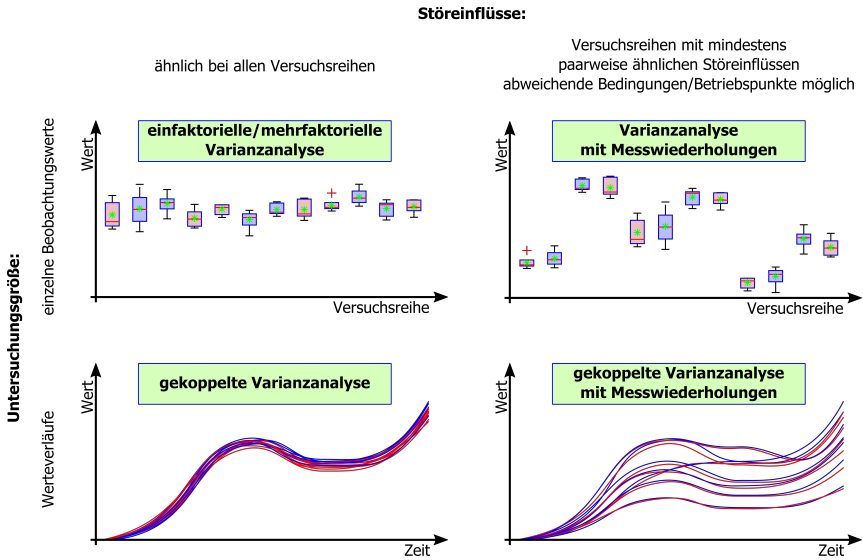


Bild 4.8: Auswahl des Auswerteverfahren in Abhängigkeit von Untersuchungsgröße und Störeinflüssen

die Fehlervarianz des Prüfprozesses sowie die Versuchsanzahl die entscheidenden Einflussgrößen auf die erzielbare Schärfe sind. Während sich die Experimentumfänge relativ leicht erhöhen lassen, kann die Streuung des Prüfprozesses nur durch eine Veränderung der Prozessvorgaben verbessert werden. Beispiele sind die engere Eingrenzung der Rahmenbedingungen, Entfernen von Störeinflüssen oder die Reduzierung der Komplexität.

Außerdem wurde gezeigt, dass sich durch die Auswertung als Direktvergleiche die Störeinflüsse durch die Versuchsbedingungen teilweise herausrechnen lassen. Diese Auswerteverfahren eignen sich zudem für eine Beurteilung über mehrere Betriebspunkte und verschiedene Umweltbedingungen. Wie in den Gleichungen 4.89 und 4.97 zu sehen, ist die erzielbare Schärfe dann neben der Versuchsanzahl von der Interaktion zwischen Einflussgröße und Prüfbedingungen abhängig.

Der Einfluss der Versuchslänge kann nicht allgemein vorhergesagt werden. Bei Werteverläufen wächst in den meisten Fällen durch eine Ausdehnung der Länge auch die Abweichungsquadratsumme zwischen den Kurvenver-

läufen überproportional. Somit erhöht sich die Varianz und die erzielbare Schärfe verschlechtert sich.

Bei der Auswertung stationärer oder integrierter Prüfgrößen macht sich eine Verlängerung der Einzelversuche dagegen häufig positiv bemerkbar. Schwingungen können länger abklingen und träge Störgrößen nähern sich besser ihrem Zielwert. Damit reduziert sich die Varianz, die Beobachtungswerte werden robuster. Prüfgrößen mit integrierendem Charakter, wie die zurückgelegte Strecke oder die verbrauchte Kraftstoffmasse, erreichen höhere Absolutwerte. Damit ergeben sich auch größere Unterschiede und eine differenziertere Auswertung wird möglich. Gleichzeitig sinken die Anforderungen an die Messauflösung, sofern die Zielgröße nicht durch Integration einer Messgröße berechnet wird.

Folglich kann durch die Verlängerung des Prüfzyklus zwar die Differenzierbarkeit der Zielgröße verbessert werden, gleichzeitig müssen dafür jedoch häufig auch größere Streuungen bei den Eingangswerten akzeptiert werden. Dies führt wiederum zu einer größeren Streuung bei der Zielgröße. Inwieweit sich diese Effekte kompensieren, ist vor allem von der Ausprägung der Störgrößen abhängig und muss individuell überprüft werden.

Derartige und weitere Analysen sind mit den hier vorgestellten statistischen Auswerteverfahren möglich. Dank der hinterlegten Statistik sind dabei die Fehlerwahrscheinlichkeiten der Ergebnisse unter den gemachten Annahmen bekannt. Mithilfe der jeweils geeigneten Varianzanalysen können sowohl einfache Versuchsergebnisse als auch Werteverläufe hinsichtlich signifikanter Unterschiede untersucht werden. Ein entscheidender Schritt ist dabei nicht nur die Streuung des einzelnen Experiments zu kennen, sondern die Varianz von umfangreichen Referenzmessungen zu verwenden. Dadurch wird die Wiederholbarkeit beziehungsweise die Reproduzierbarkeit abgesichert. Die Wahrscheinlichkeit Abweichungen aufgrund von Streuungen im Prüfprozess irrtümlich als Effekt der Einflussgröße zu interpretieren sinkt. Allerdings können in Folge dessen kleinere Effekte übersehen werden. In Kapitel 5 ist Anwendung und Nutzen dieser Analyseverfahren in konkreten Beispielen veranschaulicht.



## 5 Die Prüfprozessbeurteilung in der Anwendung

Nachdem in Kapitel 4 Vorgehensweise sowie Mathematik zur Beurteilung der Prüfprozesseignung für eine Messaufgabe beschrieben wurde, steht im folgenden die Anwendung dieser Verfahren im Vordergrund.

In bewusst einfach gehaltenen Prüfprozessen wird die Prüfprozessbeurteilung bei Anwendung verschiedener Auswerteverfahren durchgeführt. Sowohl stationäre als auch transiente Daten werden hinsichtlich ihrer Wiederholpräzision wie auch bezüglich ihrer Präzision in einem Direktvergleich untersucht. Hierbei offenbaren sich einige Grenzen der Präzisionsabschätzung, die in Kapitel 5.3 zusammengefasst sind.

### 5.1 Wiederholpräzision auf verschiedenen Untergründen

Die Fahrleistung ist je nach Fahrzeugart eine der entscheidenden Fahrzeugcharakteristika einer mobilen Arbeitsmaschine. Dementsprechend sind in Prüfprozessen zur Beurteilung der Leistungsfähigkeit häufig auch Fahranteile zu finden. Bei der Definition eines Prüfprozesses für ein Offroadfahrzeug besteht bei der Festlegung der Fahrbahn in vielen Fällen ein Konflikt zwischen Praxisnähe und Wiederholpräzision. Zur Abschätzung des Fahrbahneinflusses auf die Prüfprozessstreuung wurden mit einem Teleskopladler Claas Scorpion 7040 Referenzversuche auf einem Feld, einer Asphaltfläche und einem Rollenprüfstand durchgeführt (Abbildung 5.1). Eine tabellarische Zusammenfassung der Informationen zu Entstehung und Durchführung der Prüfprozesse, wie in Kapitel 4.1 vorgeschlagen, befindet sich im Anhang A.4. Die textuelle Beschreibung in den folgenden Abschnitten erläutert Details und Hintergründe zur Gestaltung der Prüfprozesse. Nä-

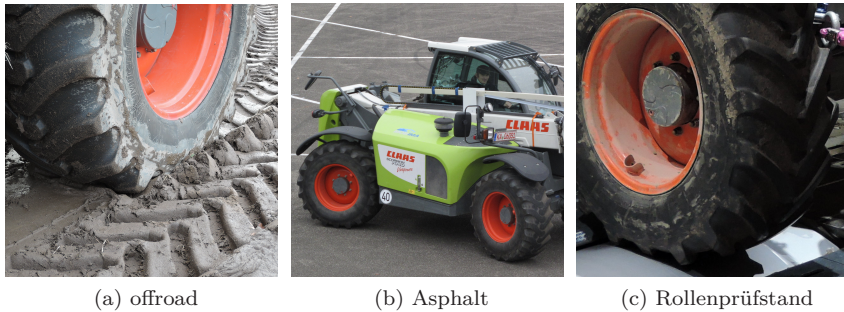


Bild 5.1: Untersuchte Untergründe zur Abschätzung des Fahrbahneinflusses

here Informationen zum Versuchsfahrzeug sind im Anhang Kapitel A.1 zu finden.

Ziel der im Folgenden vorgestellten Untersuchungen war es den Fahrbahneinfluss auf die Prüfprozessstreuung zu analysieren. Um keine fahrzeugartspezifischen Streuungsanteile mitzumessen und eine Übertragbarkeit auf andere mobile Arbeitsmaschinen zu ermöglichen wurde der Arbeitsantrieb während dieser Versuche nicht aktiv betätigt. Die Schaufel des Teleskopladers wurde vollständig angekippt und etwa 20 cm über dem Boden positioniert. Der Laststabilisator des Teleskoparms ist in dieser Position deaktiviert. Der Fahrzeugbediener steuert das Fahrzeug ausschließlich mit stationären Vorgaben. Damit schließt die Systemgrenze der Untersuchung sowohl den Arbeitsantrieb als auch das Zusammenspiel zwischen Fahrzeug und Fahrer aus. Dementsprechend repräsentieren die Prüfprozesse nicht das Einsatzprofil eines Teleskopladers.

Zur Minimierung von Störeinflüssen wurde der Teilprozess *Fahren* nicht in einem repräsentativen Fahrzyklus, sondern in zwei sehr einfach gehaltenen Prüfzenarien untersucht. Bei einer Fahrt mit geringer, konstanter Geschwindigkeit in der Ebene müssen im Wesentlichen nur die Radwiderstände überwunden werden. Entsprechend treten bei diesen Versuchen eher kleine Kräfte und Leistungen auf. Die Untersuchungsgrößen sind einzelne Beobachtungswerte, die mit den Auswerteverfahren *einfaktorische Varianzanalyse* und *Varianzanalyse mit Messwiederholungen* ausgewertet werden.

Im zweiten Prüfprozess, einer Beschleunigung des Fahrzeuges aus dem Stand, müssen aufgrund der Fahrzeugträchtigkeit deutlich größere Leistungen aufgebracht werden. Außerdem haben die Messgrößen einen transienten Verlauf. Dementsprechend kommen hier die Auswerteverfahren mit *gekoppelter Varianzanalyse* zur Anwendung. In Abhängigkeit von Geschwindigkeit und Beschleunigung ergeben sich verschiedene Leistungsniveaus. Allerdings sind die Fahranteile bei größeren Geschwindigkeiten oder hohen Lasten sehr klein. Ebenso sind Verzögerungsphasen nicht Teil der Tests. Um mögliche Wettereinflüsse nachvollziehen zu können, wurden Windrichtung und Stärke sowie die Umgebungstemperatur während der Versuche mithilfe einer Wetterstation lokal erfasst und aufgezeichnet. Weitere Informationen zum Wetterverlauf liefert die Datenbank einer nahegelegenen, fest installierten Wetterstation.

In Tabelle 5.1 sind die Wertebereiche ausgewählter Rahmenbedingungen dokumentiert. Die Umgebungstemperaturen der Außenexperimente sind saisonbedingt deutlich niedriger und besitzen eine größere Spannweite als im Prüfstandsgebäude. Die Versuche auf der Rolle wurden über mehrere Tage verteilt gefahren, weshalb der Luftdruck hier die größte Spannweite aufweist. Der Wind wehte bei den Außenversuchen schwach bis mäßig in Fahrtrichtung mit einer deutlichen Seitenkomponente. Bei den Asphaltversuchen kam der Wind teilweise auch quer zur Fahrtrichtung. Die Luftanströmung auf dem Rollenprüfstand sorgt für die Fahrzeugkühlung an Stelle des Fahrtwindes und trägt nicht zum Fahrwiderstand bei. Der Niederschlag in der Nacht zwischen den beiden offroad Versuchstagen wirkte sich beträchtlich auf die Bodeneigenschaften und damit auch auf Betrag sowie Streuung der Zielgrößen aus.

Zur besseren Einschätzung der Systemtemperaturen ist die Spannweite der Getriebeöltemperatur angegeben. Die Werte hierfür werden mithilfe zweier Temperatursensoren in der Hoch- und Niederdruckseite im hydrostatischen Getriebe erfasst und gemittelt. Allen Versuchsreihen geht eine mindestens 20 minütige Warmlaufphase mit nicht spezifiziertem Fahrzeugbetrieb voran. Dennoch schwankt das Temperaturniveau des Getriebeöls in Abhängigkeit der unmittelbar vorangegangenen Belastungsintensität erheblich.

Tabelle 5.1: Spannweite ausgewählter Versuchsbedingungen bei den Versuchen auf verschiedenen Untergründen

Messgröße	offroad Tag 1	offroad Tag 2	offroad gesamt	Asphalt	Rollen- prüfstand
Umgebungs- temperatur	4 – 16 °C	5 – 10 °C	4 – 16 °C	7 – 9 °C	19 – 21 °C
Luftdruck	1008 hPa bis 1010 hPa	1011 hPa bis 1015 hPa	1008 hPa bis 1015 hPa	1019 hPa bis 1020 hPa	1004 hPa bis 1025 hPa
rel. Luftfeuchte	28 – 44 %	67 – 82 %	28 – 82 %	50 – 56 %	25 – 39 %
Wind- geschwindigkeit	0 – 3,3 $\frac{m}{s}$	0 – 3,0 $\frac{m}{s}$	0 – 3,3 $\frac{m}{s}$	0 – 2,2 $\frac{m}{s}$	5,0 – 6,0 $\frac{m}{s}$
Windrichtung rel. zur Fahrrichtung	22,5 – 45°	22,5 – 45°	22,5 – 45°	67,5 – 90°	180°
Niederschlag in den letzten 48h	0 $\frac{1}{m^2}$	6,3 $\frac{1}{m^2}$	6,3 $\frac{1}{m^2}$	0 $\frac{1}{m^2}$	–
Fahrzeug- Getriebetemperatur	22 – 30 °C	13 – 23 °C	13 – 30 °C	13 – 24 °C	27 – 35 °C

Zielgröße der Untersuchungen ist die Fahrleistung des Untersuchungsfahrzeuges, die sich aus Fahrzeuggeschwindigkeit und Antriebsmoment zusammensetzt. Die **Messgenauigkeiten** der verwendeten Messtechnik sind in Tabelle 5.2 dokumentiert. Durch Verschmutzungen kam es auf dem Feld zeitweise zu Unterbrechungen bei der Drehzahlerfassung der Gelenkwelle. Zusätzlich wurde bei den Außenversuchen die Fahrgeschwindigkeit über Grund mithilfe eines RTK GNSS Systems erfasst. Das Antriebsmoment wird an den beiden Gelenkwellen zu den Achsen mittels Dehnmessstreifen ermittelt. Bei den Prüfstandversuchen steht darüber hinaus Rollendrehzahl und -moment zur Verfügung.

Für die Beurteilung der Genauigkeit muss zu Beginn der Prüfprozessentwicklung der kleinste zu differenzierende Effekt  $\Delta Y_{min}$  für die Zielgrößen festgelegt werden. Beispielhaft sei hier die Anforderung definiert, in Wiederholungsversuchen Abweichungen in Höhe von 1 % des erwarteten Maximalwertes zuverlässig unterscheiden zu können. Im Reparaturhandbuch



des Teleskopladlers ist zu lesen, dass die Gelenkwelldrehzahl bei maximaler Fahrzeuggeschwindigkeit  $3460 \frac{1}{\text{min}} \approx 60 \frac{1}{\text{s}}$  erreicht [CLA07, S.5-28]. Aus Schluckvolumen, Maximaldruck und Übersetzungsverhältnis kann ein maximales Drehmoment von 2700 Nm geschätzt werden [CLA07, S.1-8f.]. Bei einem Wirkungsgrad der beiden Hydraulikeinheiten von etwa 90 % ergibt sich  $\approx 80 \text{ kW}$  als maximale Leistung auf der Gelenkwelle. Dementsprechend sollen Drehzahldifferenzen von  $0,6 \frac{1}{\text{s}}$ , Drehmomentabweichungen von 27 Nm und Leistungsunterschiede von 800 W unterschieden werden können.

Tabelle 5.2: Messgenauigkeit der verwendeten Messtechnik

Messgröße	Sensorbezeichnung	Mess-Auflösung	Messbereich	Genauigkeit
<b>Antriebsmoment Gelenkwelle</b>	Manner Sensortelemetrie DMS Drehmomentsensor	0,07 Nm	-2...2 kNm	$\pm 0,25 \%$ Hrst.- angabe
<b>Drehzahl Gelenkwellen</b>	Motrona Frequenz- Spannungs-Wandler FU 252	$0,01 \frac{1}{\text{s}}$	$0,01 \dots 80 \frac{1}{\text{s}}$	$\pm 0,1 \%$ Hrst.- angabe

Mit den Informationen zur Messtechnik in Tabelle 5.2 werden zunächst die direkt gemessenen Größen, Drehzahl und Drehmoment, beurteilt. Die Auflösung der Sensoren ist wesentlich feiner, als ein Fünftel der erwarteten Abweichungen. Der Einfluss auf die Messunsicherheit bleibt damit vernachlässigbar. Die Genauigkeit der Sensorik ist als maximale Abweichung in Prozent vom Endwert angegeben. Die Absolutwerte sind  $\pm 5 \text{ Nm}$  beziehungsweise  $\pm 0,08 \frac{1}{\text{s}}$ . Auch diese Werte sind kleiner als ein Fünftel der zu unterscheidenden Abweichungen. Beim Drehmoment ist allerdings zu beachten, dass das Gesamtmoment über je einen Sensor für Vorder- und Hinterachse gemessen und summiert wird. Bei linearer Fehlerfortpflanzung ergibt sich somit  $\pm 10 \text{ Nm}$  als Bias des Gesamtmoments. Die zuvor definierte Anforderung an die Genauigkeit der Momentenerfassung kann demnach formal nicht erfüllt werden.

Gleichzeitig beschränken sich die Untersuchungen auf Wiederholungsversuche mit derselben Messtechnik in demselben Fahrzeug. Dabei liegt der Fokus nicht auf den absoluten Erwartungswerten sondern auf den Streu-

ungen der Messwerte. Weil jeweils ähnliche Betriebspunkte verglichen werden, ähneln sich auch die systematischen Messabweichungen zum tatsächlichen Wert. Die Unterschiede zwischen diesen Werten sind folglich in einem begrenzten Messbereich beträchtlich genauer. Deshalb können bei Wiederholungsversuchen größere Abweichungen bei der absoluten Genauigkeit akzeptiert werden. Wichtig ist dabei jedoch, dass auf die Sensorik jeweils vergleichbare Einflüsse wirken. Auftretende Abweichungen werden als unbekannter Bruchteil der Streuung des Prüfprozesses erfasst. Auf Basis dieser Überlegungen sind die erwarteten systematischen Messabweichungen für die Bestimmung der Prüfprozessstreuung akzeptabel. Da die Experimente auf verschiedenen Untergründen unter ungleichen Bedingungen für die Sensorik durchgeführt wurden, ist für einen Vergleich der Absolutwerte der Experimente die Anforderung nicht erfüllt.

Zur Beurteilung der berechneten Zielgröße *Leistung* ist das Bias mit linearer Fehlerfortpflanzung entsprechend Gleichung 4.7 zu ermitteln. Angewendet auf Gleichung 5.1 ergibt sich Gleichung 5.2.

$$P = 2\pi \cdot (M_{VA} + M_{HA}) \cdot n \quad (5.1)$$

$$\Delta P = 2\pi \cdot \left( (\Delta M_{VA} + \Delta M_{HA}) \cdot n + (M_{VA} + M_{HA}) \cdot \Delta n \right) \quad (5.2)$$

Da das Bias vom gemessenen Wertebereich abhängt, wird die Messunsicherheit unter Berücksichtigung der erwarteten Maximalwerte in Diagramm 5.2 visualisiert. Weil im konkreten Fall die Drehzahl den ausgeprägteren Einfluss besitzt wurde hier das Bias der Gelenkwellenleistung über die Drehzahl aufgetragen. Der Einfluss des Drehmoments ist durch jeweils vier Leistungsniveaus abgebildet.

Zur Berechnung der blauen Kurven wurden die Werte der Messauflösung in Gleichung 5.2 eingesetzt. Die Messunsicherheit aufgrund der Auflösung liegt bei kleinen Drehzahlen knapp über 160 W, sonst deutlich darunter. Die Anforderungen an die Auflösung der Leistungswerte, ein Fünftel der zu differenzierenden Abweichungen, ist somit erfüllt.

Die Genauigkeit absoluter Messwerte, rot dargestellt, steigt über die Drehzahl stark an. Die gestellten Anforderungen an die Genauigkeit werden bei weitem nicht eingehalten. Wie weit sich das bei reinen Wiederholungsver-

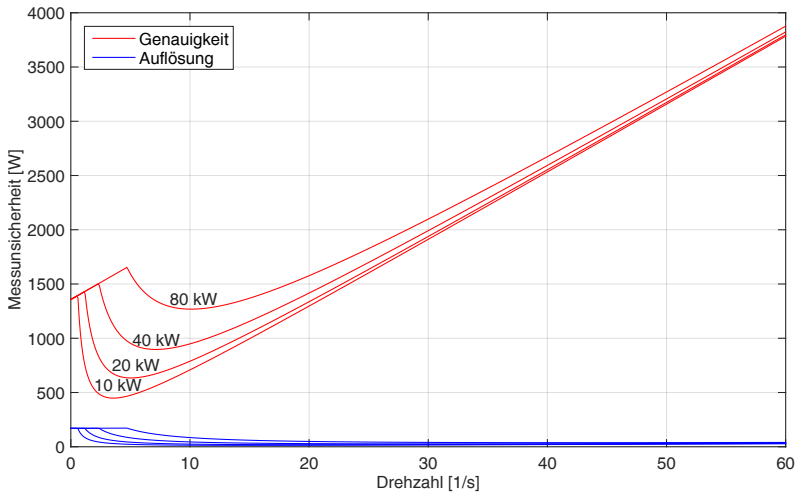


Bild 5.2: Genauigkeit und Auflösung des Messsystems für die Leistung

suchen verbessert, ist vorab kaum abzuschätzen. Einige Sensordatenblätter enthalten zusätzlich Angaben zu Linearität, Hysterese, Wiederholbarkeit, Temperaturdrift und sonstigen Einflussgrößen auf den Sensor. Die tatsächliche Unsicherheit der Messtechnik lässt sich jedoch am besten mithilfe von Referenzexperimenten abschätzen. Sofern möglich, sollte bei den bestehenden Anforderungen ein genaueres Messsystem eingesetzt werden.

### 5.1.1 Wiederholpräzision stationärer Versuche

Gegenüber einer dynamischen Versuchsführung sind bei stationären Versuchen die Störeinflüsse auf ein Minimum begrenzt. Alle Eingangsgrößen werden konstant gehalten und dem System genügend Zeit gegeben, um einen stationären Betriebspunkt zu erreichen. Bei einem Fahrversuch auf einer ebenen Freifläche bedeutet das Gaspedal- und Lenkradstellung gleich zu halten.

Für die Untersuchungen des Streuungseinflusses der Fahrbahn betätigt der Fahrer das Gaspedal gegen einen mechanischen Anschlag. Die Fahrtrichtung ist geradeaus und wird nur falls erforderlich korrigiert. Bei dieser unveränderten Bedienvorgabe beschleunigt das Fahrzeug, bis sich die zur Verfügung gestellte Zugkraft im Gleichgewicht mit den in Kapitel 2.4 er-

läuterten Fahrwiderständen befindet. Die Zielgröße Antriebsleistung wird aus den beiden Messgrößen Drehzahl und Drehmoment der Kardanwellen berechnet. Diese werden als gewichteter Durchschnitt eines 20 s langen Messfensters am Ende eines Versuches ermittelt.

Die Abbildungen 5.3 und 5.4 geben einen sehr guten optischen Eindruck zu Lage und Streuung der Messungen. Die Daten stammen aus Fahrversuchen mit einem Teleskoplader auf verschiedenen Untergründen in der Fahrstufe *Arbeiten*. Auf jedem Untergrund wurden sechs Versuchsreihen mit jeweils fünf Einzelversuchen durchgeführt. Für die Drehzahl der Versuchsreihe  $O_9$  ist kein Boxplot zu sehen, da hier der Drehzahlsensor aufgrund starker Verschmutzungen ausgefallen ist.

Wie erwartet müssen auf dem nachgiebigen Boden eines Feldes größere Fahrwiderstände überwunden werden. Das Drehmoment steigt, die Drehzahl wird gedrückt. Ebenso streuen die Beobachtungswerte der Offroadfahrten insbesondere am 2. Versuchstag deutlich stärker als auf Asphalt oder auf dem Rollenprüfstand. Diese größere Streuung ist, wie der erhöhte Fahrwiderstand, auf die Verdrängungen der inhomogenen Erde zurück-

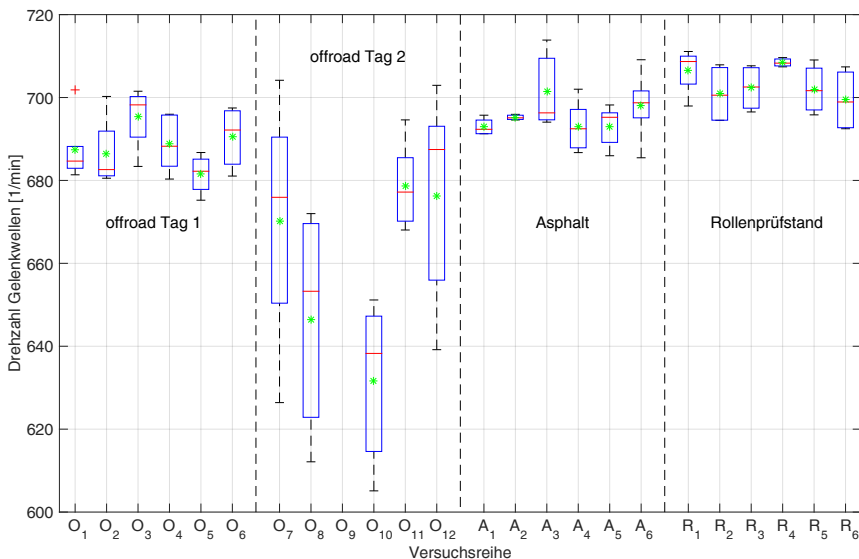


Bild 5.3: Gelenkwellendrehzahl der stationären Versuchsreihen auf verschiedenen Untergründen

zuführen. Dieser Effekt wurde durch die höhere Bodenfeuchtigkeit am 2. Versuchstag verstärkt. Eine Reproduktion der Messergebnisse des Vortages war nicht möglich.

Die Boxplots aus dem Asphalt- und dem Rollenprüfstand-Experiment liegen auf ähnlichem Niveau und besitzen vergleichbare Streuungen. Dieses Ergebnis betätigt die Aussage des Prüfstandsherstellers, nach dem der gesamte „... Reifen-Rolle-Kontakt [...] mit einer trockenen Fahrbahnoberfläche (Straßenbelag) vergleichbar...“ sei [MAH12, S. 34]. Da bei diesen beiden Experimenten keine wesentlichen Unterschiede zwischen den Versuchstagen festgestellt werden konnten, sind diese nicht explizit unterschieden.

Die statistische Auswertung dieser als unabhängig angenommenen Stichproben erfolgt mit einer einfaktoriellen Varianzanalyse entsprechend Kapitel 4.4. Die Analyseergebnisse zur Beurteilung der Reproduzierbarkeit der Gelenkwendendrehzahl sind in Tabelle 5.3 eingetragen. Die Standardabweichungen auf Asphalt und dem Rollenprüfstand sind dicht beieinander. Selbst der Wert des offroad Experiments am ersten Tag bewegt sich in der gleichen Größenordnung. Nur die Streuung der Versuche am zweiten

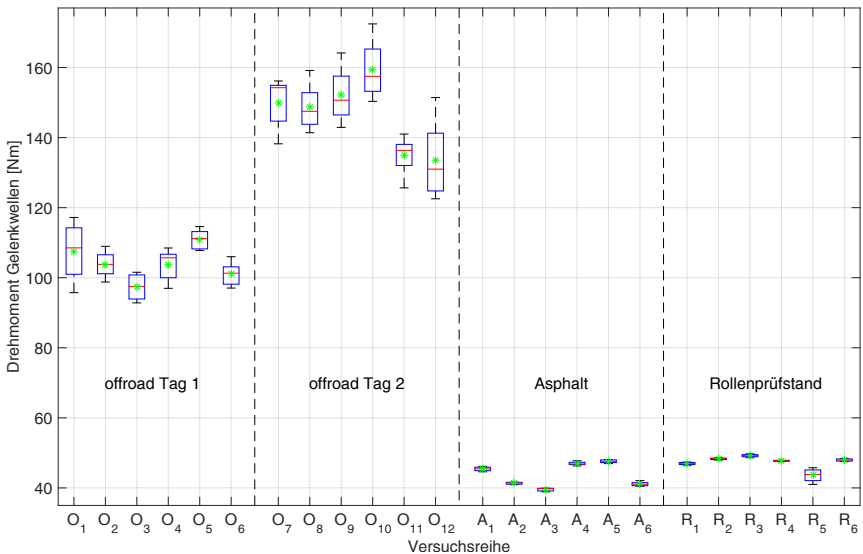


Bild 5.4: Gesamtdrehmomente der Gelenkwellen der stationären Versuchsreihen auf verschiedenen Untergründen

Tabelle 5.3: Analyseergebnisse der Gelenkwelldrehzahl bei stationären Versuchen auf verschiedenen Untergründen

Auswertekriterium	offroad Tag 1	offroad Tag 2	offroad gesamt	Asphalt	Rollen- prüfstand
<b>Anzahl Referenzversuche</b>	6 × 5	5 × 5	11 × 5	6 × 5	6 × 5
<b>Mittelwert Referenzversuche</b>	688 $\frac{1}{\text{min}}$	661 $\frac{1}{\text{min}}$	675 $\frac{1}{\text{min}}$	695 $\frac{1}{\text{min}}$	703 $\frac{1}{\text{min}}$
<b>Standardabweichung</b>	7,77 $\frac{1}{\text{min}}$	28,4 $\frac{1}{\text{min}}$	24,2 $\frac{1}{\text{min}}$	6,47 $\frac{1}{\text{min}}$	6,31 $\frac{1}{\text{min}}$
<b>vierfacher Interquartilabstand</b>	56,1 $\frac{1}{\text{min}}$	177 $\frac{1}{\text{min}}$	83,9 $\frac{1}{\text{min}}$	27,8 $\frac{1}{\text{min}}$	45,2 $\frac{1}{\text{min}}$
<b>Wahrscheinlichkeit Kruskal-Wallis-Test</b>	24 %	3,1 %	0,23 %	40 %	14 %
<b>Normalverteilung des Experiments</b>	KST <sup>2</sup> ✓ SWT <sup>1</sup> X	KST ✓ SWT ✓	KST X SWT X	KST ✓ SWT X	KST ✓ SWT X
<b>Schärfe bei 5 × 5 Versuchen pro Experiment</b>	± 5,90 $\frac{1}{\text{min}}$	± 28,1 $\frac{1}{\text{min}}$	± 24,0 $\frac{1}{\text{min}}$	± 4,91 $\frac{1}{\text{min}}$	± 4,79 $\frac{1}{\text{min}}$

Tag im Gelände ist wesentlich stärker ausgeprägt. Der vierfache Interquartilabstand der Experimente liefert als robuste Schätzung der Spannweite ein ähnliches Bild. Mit dem Kruskal-Wallis-Test lassen sich nur bei den Feldversuchen signifikante Unterschiede zwischen den Reihen feststellen. Bei trockenen Bedingungen ergeben sich auch auf dem Feld keine signifikanten Unterschiede zwischen den Versuchsreihen. Der Shapiro-Wilk-Test<sup>1</sup> zeigt, dass lediglich die offroad Versuche des zweiten Tags eine Normalverteilung besitzen. Mit dem etwas schwächeren Kolmogorov-Smirnov-Test<sup>2</sup> wird die Normalverteilung nur für die zusammengefassten offroad Versuche abgelehnt. Unter geeigneten Rahmenbedingungen kann bei den Referenzmessungen auf allen drei Untergründen die geforderte Schärfe von  $0,6 \frac{1}{\text{s}} = 36 \frac{1}{\text{min}}$  erreicht werden. Wenn die beiden offroad Versuchstage zu einem Experiment zusammengefasst werden, können die guten Ergebnisse des

<sup>1</sup> SWT = Shapiro-Wilk-Test auf Normalverteilung<sup>2</sup> KST = Kolmogorov-Smirnov-Test auf Normalverteilung

Tabelle 5.4: Analyseergebnisse des Gelenkwelldrehmomentes bei stationären Versuchen auf verschiedenen Untergründen

Auswertekriterium	offroad Tag 1	offroad Tag 2	offroad gesamt	Asphalt	Rollen- prüfstand
<b>Anzahl Referenzversuche</b>	6 × 5	6 × 5	12 × 5	6 × 5	6 × 5
<b>Mittelwert Referenzversuche</b>	104 Nm	146 Nm	126 Nm	44,1 Nm	47,3 Nm
<b>Standardabweichung</b>	6,47 Nm	12,1 Nm	23,3 Nm	3,19 Nm	2,02 Nm
<b>vierfacher Interquartilabstand</b>	35,6 Nm	66,5 Nm	175 Nm	24,0 Nm	5,84 Nm
<b>Wahrscheinlichkeit Kruskal-Wallis-Test</b>	3,5 %	0,4 %	4,1 E-5 %	4,9 E-2 %	0,1 %
<b>Normalverteilung des Experiments</b>	KST ✓ SWT ✓	KST ✓ SWT ✓	KST ✓ SWT X	KST ✓ SWT X	KST ✓ SWT X
<b>Schärfe bei 5 × 5 Versuchen pro Experiment</b>	± 6,41 Nm	± 12,0 Nm	± 23,1 Nm	± 3,16 Nm	± 2,01 Nm

ersten Tages die große Streuung des zweiten Tages ein wenig ausgleichen. Für die Reproduktion eines offroad Experiments müssten die zulässigen Rahmenbedingungen wie beispielsweise die Bodenfeuchte und Zusammensetzung deutlich eingeschränkt werden.

Für das Antriebsmoment am Getriebeausgang liefert die Analyse größere Unterschiede zwischen den Untergründen. Streuungen, Interquartilabstände und erzielbare Schärfe zeigen jeweils etwa eine Halbierung von feuchter Erde zu trockener Erde zu Asphalt und schließlich zum Rollenprüfstand. In allen Experimenten ist die Streuung innerhalb der Reihen deutlich geringer als zwischen den Reihen. Folglich gibt es zwischen den Reihen nicht erfasste signifikante Veränderungen im Prüfprozess. Bei den Feldversuchen könnte das auf Änderungen der Bodeneigenschaften zurückzuführen sein. Auf Asphalt und dem Rollenprüfstand werden abweichende Reifen- und Öltemperaturen als Störeinflüsse angenommen. Tatsächlich wurden bei der

Versuchsreihe  $R_5$  höhere Reifentemperaturen aufgrund einer vorhergehenden halbstündigen Fahrt bei hoher Geschwindigkeit gemessen. Wird die Reihe  $R_5$  bei der Auswertung nicht berücksichtigt, sind die Beobachtungswerte des Experiments normalverteilt und die Schärfe verbessert sich auf  $\pm 0,8 \text{ Nm}$ . Folglich kann durch eine sorgfältigere Vorkonditionierung des Versuchsfahrzeuges die erreichbare Schärfe auf festen Untergründen nochmals wesentlich gesteigert werden.

Auf allen drei untersuchten Fahrbahnen konnten in den Referenzversuchen die geforderte Differenzierungsschärfe von  $27 \text{ Nm}$  bei 25 Versuchen erreicht werden. Die Unsicherheiten auf fester Fahrbahn sowie auf der trockenen Erde des Feldes innerhalb eines Tages waren mit  $\pm 7 \text{ Nm}$  sogar deutlich geringer.

Aus den Beobachtungswerten für Drehzahl und Drehmoment der Gelenkwelle wird die zu den Achsen übertragene Antriebsleistung berechnet. Niveau und Streuung der Beobachtungswerte der Antriebsleistung sind in Abbildung 5.5 dargestellt. Da die relativen Abweichungen bei den Dreh-

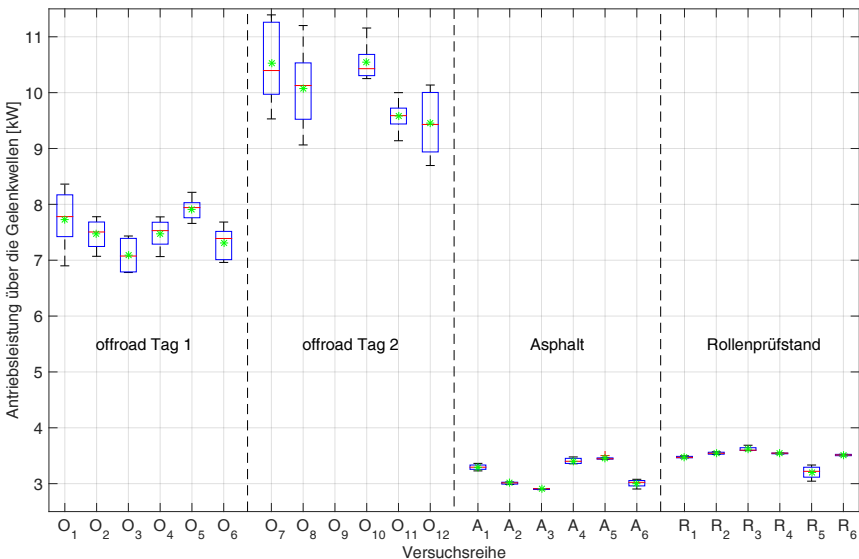


Bild 5.5: Antriebsleistung über die Gelenkwellen der stationären Versuchsreihen auf verschiedenen Untergründen



momenten deutlich ausgeprägter als bei den Drehzahlen sind, zeigen die Grafiken für Leistung und Drehmoment qualitativ gleiche Ergebnisse. Gleichfalls zeigen die Analyseergebnisse dieselben Tendenzen, wie bereits bei den Drehmomenten. Die Anforderung Unterschiede von 800 W nachzuweisen kann bei den offroad Referenzversuchen nur innerhalb eines Tages erreicht werden. Gemäß diesen Referenzmessungen können auf festen Untergründen deutlich höhere Anforderungen an die Messschärfe erfüllt werden. Wird bei dem Experiment auf dem Rollenprüfstand die Reihe  $R_5$  als Ausreißer herausgenommen, sind hier Differenzen bis  $\pm 55$  W mit diesem Prüfprozess messbar.

Die Auswertung zeigt, dass sich bei der erläuterten stationären Versuchsführung die Geschwindigkeit abgesehen von Asphalt und Rollenprüfstand innerhalb eines Versuchstages auch auf trockenem, unbefestigtem Boden mit etwa vergleichbarer Schärfe reproduzieren lässt. Diese Aussage lässt

Tabelle 5.5: Analyseergebnisse der Antriebsleistung über die Gelenkwelle bei stationären Versuchen auf verschiedenen Untergründen

Auswertekriterium	offroad Tag 1	offroad Tag 2	offroad gesamt	Asphalt	Rollen- prüfstand
<b>Anzahl Referenzversuche</b>	6 × 5	5 × 5	11 × 5	6 × 5	6 × 5
<b>Mittelwert Referenzversuche</b>	7,52 kW	10,0 kW	8,68 kW	3,21 kW	3,48 kW
<b>Standardabweichung</b>	0,42 kW	0,72 kW	1,39 kW	0,22 kW	0,15 kW
<b>vierfacher Interquartilabstand</b>	2,01 kW	3,47 kW	10,3 kW	1,72 kW	0,35 kW
<b>Wahrscheinlichkeit Kruskal-Wallis-Test</b>	3,3 %	2,2 %	1,7 E-4 %	7,1 E-2 %	0,11 %
<b>Normalverteilung des Experiments</b>	KST ✓ SWT ✓	KST ✓ SWT ✓	KST ✓ SWT X	KST ✓ SWT X	KST ✓ SWT X
<b>Schärfe bei 5 × 5 Versuchen pro Experiment</b>	± 0,43 kW	± 0,72 kW	± 1,38 kW	± 0,22 kW	± 0,15 kW

sich jedoch nicht direkt auf andere Messgrößen der Untersuchung übertragen. Des Weiteren verdeutlichen die Messergebnisse den großen Einfluss der Bodenfeuchte bei unbefestigten Untergründen auf die Streuung. Ebenso ist bei einer Änderung der Bodenzusammensetzung mit deutlich abweichenden Versuchsergebnissen zu rechnen.

Wie erwartet zeigt der Rollenprüfstand für alle Beobachtungswerte die geringsten Streuungen. Ein befestigter Untergrund wie Asphalt oder Beton fällt dagegen bei der Wiederholpräzision nicht weit ab. Für die Untersuchungen über längere Zeiträume wird sich die erzielbare Schärfe in natürlicher Umgebung jedoch zunehmend verschlechtern.

### **Wiederholpräzision in verschiedenen Betriebspunkten**

Da auf dem Rollenprüfstand offensichtlich die beste Wiederholpräzision erzielt werden kann, wurde in dieser Prüfumgebung der Einfluss des Fahrzeugbetriebspunktes auf die Streuung und damit auf die erzielbare Messschärfe untersucht. Dazu wurde der zuvor beschriebene Prüfprozess mit fünf verschiedenen Gaspedalansschlägen in beiden Fahrstufen durchfahren. Jedes dieser Experimente setzt sich wiederum aus sechs Reihen a fünf Einzelversuchen zusammen, die randomisiert über mehrere Tage aufgezeichnet wurden.

Die Auswertung dieser 300 Einzelversuche erfolgt analog zu dem Vorgehen bei der Analyse der Einflüsse des Untergrundes. Wie zuvor werden die Drehzahl und das Drehmoment der Gelenkwelle sowie die Leistung als Produkt dieser Messgrößen ausgewertet. In Abbildung 5.6 ist die jeweils ermittelte Schärfe für  $5 \times 5$  Versuche über den Mittelwerten der zehn Experimente aufgetragen. Die Diagramme zeigen, dass sich die Ergebnisse eines Betriebspunktes nicht beliebig auf andere Fahrgeschwindigkeiten oder Lasten übertragen lassen. Tendenziell ergeben sich mit zunehmenden Werten auch größere Streuungen. Ein klarer Trend zeichnet sich jedoch nur für die Drehzahl und die Leistung der Gelenkwelle ab.

Insgesamt zeigt sich ein ernüchterndes Ergebnis. Mit dem verwendeten Prüfprozess lassen sich bei  $5 \times 5$  Versuchen pro Experiment lediglich signifikante Abweichungen ab etwa 1,8 % des Messwertes bei der Gelenkwelldrehzahl und ab rund 9 % beim Drehmoment ermitteln. Für die Leistung

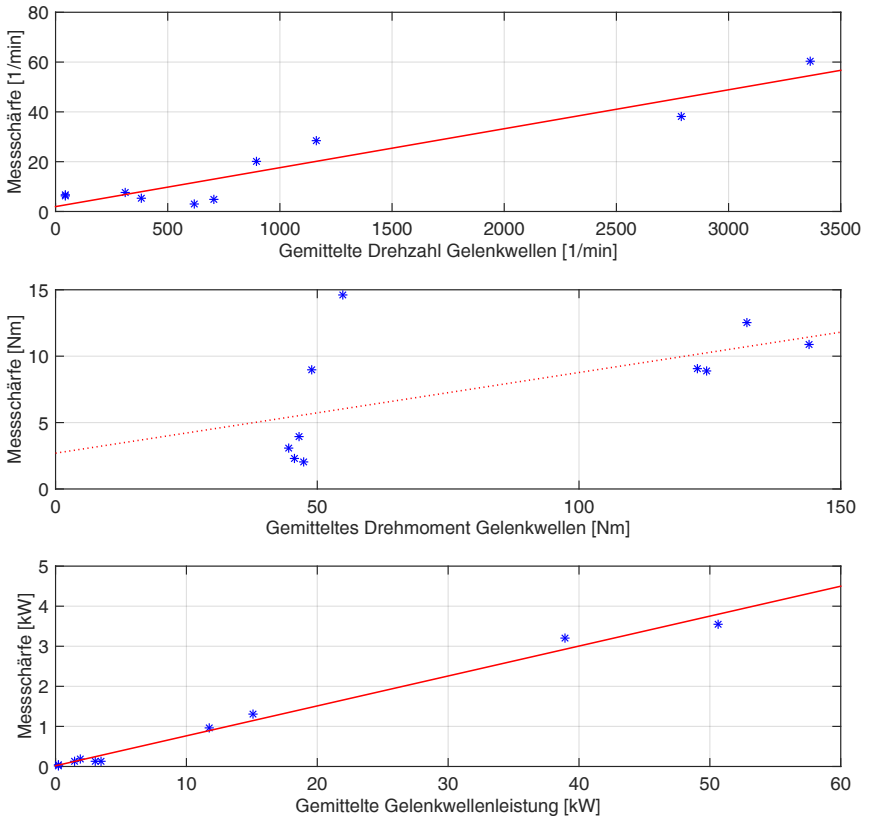


Bild 5.6: Erzielbare Schärfe in verschiedenen Betriebspunkten

indiziert die Trendlinie eine Trennschärfe von ca. 7,5% des Messwertes. Folglich können bei höheren Wertenniveaus die geforderten Trennschärfen nicht erreicht werden.

In der konkreten Anwendung werden diese Werte jedoch teilweise relativiert. Beispielsweise wirkt sich eine Variation der Fahrbahnsteigung um 1% bei geringer Geschwindigkeit und damit geringer Last kaum auf die Drehzahl aus und kann dort nicht signifikant unterschieden werden. Die Unterschiede im Moment und der Leistung sind dagegen sowohl bei kleinen als auch bei großen Geschwindigkeiten signifikant.

Eine Verbesserung der Schärfe lässt sich durch Anheben der Versuchszahl sowie durch eine exakt beschriebene Vorkonditionierung des Prüfsys-

tems erreichen. Ferner wird in Kapitel 5.2 aufgezeigt, welche Trennschärfe bei der Durchführung von Direktvergleichen erreicht werden kann.

### 5.1.2 Wiederholpräzision transientser Versuche

Zur Beurteilung der Wiederholbarkeit transientser Versuche auf verschiedenen Untergründen werden Beschleunigungsvorgänge betrachtet. Fahrzeug, Versuchsort und sonstige Umweltbedingungen sind identisch mit den zuvor beschriebenen stationären Versuchen. Auf dem Rollenprüfstand wird die Massenträgheit des Prüflings durch die Momentenregelung mit einer Fahrsimulation nachgebildet. Die Radwiderstände, die auf dem Prüfstand natürlich entstehen, werden bewusst nicht manipuliert. Der Luftwiderstand wird bei diesen Untersuchungen vernachlässigt, die Fahrbahn waagrecht angenommen. Der Teleskoplader beschleunigt aus dem Stand in der Fahrstufe *Fahren*. Der Fahrzeugführer tritt hierzu das Gaspedal direkt zügig bis zum Endanschlag durch und hält das Fahrzeug lediglich mit minimalen Lenkbewegungen auf Kurs.

Die Synchronisation der Einzelversuche erfolgt anhand der steilen Flanke der Gaspedalposition. Das Überschreiten von 25 % des Pedalwegs gilt als Startpunkt des Prüfprozesses. Das Ende wird bei 8 s gesetzt.

Die Auswertung der so generierten Messdatenverläufe erfolgt mit einer gekoppelten Varianzanalyse entsprechend Kapitel 4.4.6. In Abbildung 5.7 ist zunächst der Drehzahlverlauf der Gelenkwellen bei den Referenzversuchen auf dem Rollenprüfstand dargestellt. Der starke, grüne Verlauf ist das arithmetische Mittel aller 30 aufgezeichneten Einzelversuche. Im zweiten Diagramm dieser Abbildung sind die Abweichungen der einzelnen Messwerte zu diesem durchschnittlichen Verlauf aufgetragen. Eine Glättung des Messdatenverlaufs würde bei diesem Auswerteverfahren die Varianzanalyse vollständig aushebeln. Mit diesem Wissen ist auf jegliche glättende Datenmanipulationen bei der Analyse von Werteverläufen zu verzichten. Zur besseren Vergleichbarkeit des Grundniveau des Verlaufs mit anderen Experimenten, wird entgegen der Erläuterungen in Kapitel 4.4.6 das arithmetische Mittel aller Messwerte des Experiments als Bezugslinie verwendet. Die entsprechende Transformationsvorschrift ist in Gleichung 5.3 beschrieben.

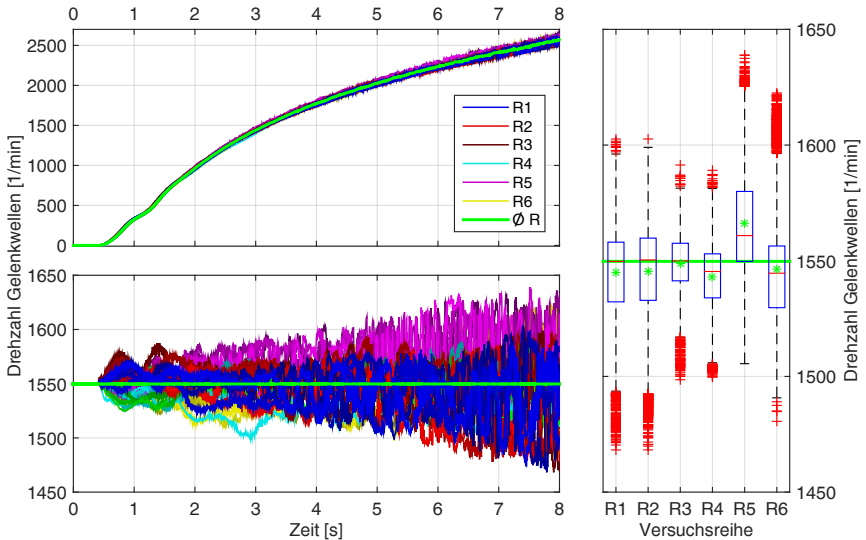


Bild 5.7: Streuung der Gelenkwelldrehzahl bei Beschleunigungsversuchen auf dem Rollenprüfstand

$$Z_{mnp} = Y_{mnp} - \bar{Y}_{..p} + \bar{Y} \quad (5.3)$$

Die Boxplots auf der rechten Seite von Abbildung 5.7 sind aus den transformierten Werten  $Z_{mnp}$  der sechs Versuchsreihen auf dem Prüfstand gebildet. Auf die gleiche Weise wurden auch für die anderen Experimente Boxplots erstellt und in Abbildung 5.8 gegenübergestellt. In den Abbildungen 5.9 und 5.10 wurde dasselbe für die Summe der Drehmomente der beiden Gelenkwellen sowie die berechnete Leistung umgesetzt.

Für die Reihe  $O_3$  gibt es keine Daten für die Gelenkwelldrehzahl und die Leistung aufgrund eines Sensorausfalls. Erwartungsgemäß zeigen diese dynamischen Prüfprozesse qualitativ die gleichen Eigenschaften wie bei den stationären Versuchen. Wegen des höheren Fahrwiderstands auf nachgiebigen Untergründen werden dort im Prüfprozess geringere Geschwindigkeiten bzw. Gelenkwelldrehzahlen erreicht. Der Beschleunigungswiderstand bildet bei diesem Prüfprozess den größten Anteil des Fahrwiderstandes. Dementsprechend sind die Unterschiede bei den Drehmomenten weniger ausgeprägt. Im Vergleich der Gelenkwellenmomente fällt das augenschein-

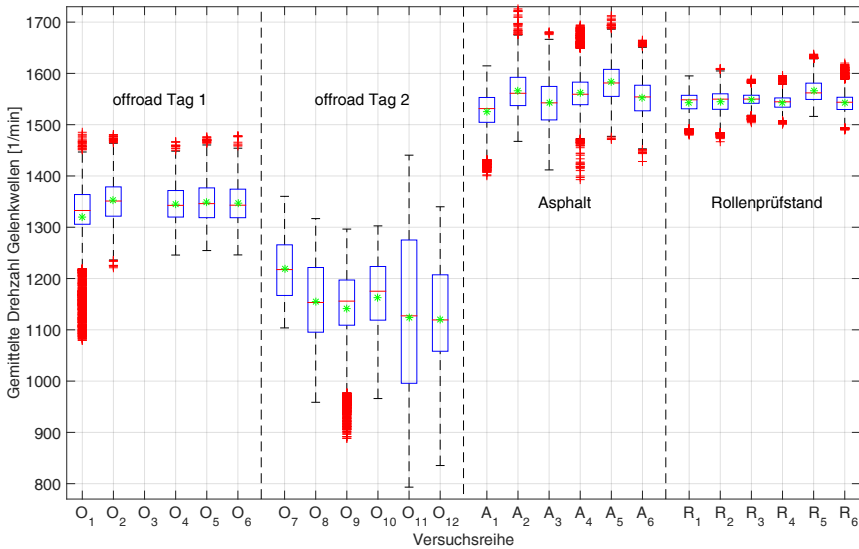


Bild 5.8: Streuung der Gelenkwellendrehzahl bei Beschleunigungsversuchen auf verschiedenen Untergründen

lich geringere Drehmoment auf dem Rollenprüfstand auf. Mögliche Ursachen sind die Vernachlässigung des Luftwiderstandes, Totzeiten und Regelungsabweichungen in der Simulation der Fahrzeugmassenträgheit sowie geringere Verluste aufgrund der höheren Umgebungstemperatur. Ferner ist ein Temperaturdrift der Drehmomentenerfassung denkbar. Die Versuchsreihen  $O_{11}$  und  $O_{12}$  sowie  $A_4$  und  $A_5$  zeigen im Vergleich zu den anderen Reihen desselben Experiments deutlich größere Streuungen. Hier liegt die Vermutung nahe, dass in einzelnen Versuchen dieser Reihen der Prüfprozess nicht korrekt durchgeführt wurde oder weitere Störeinflüsse gewirkt haben. Die tatsächliche Ursache lässt sich im Nachhinein kaum noch rekonstruieren. Die Auswirkungen auf die statistische Auswertung in den Tabellen 5.6 bis 5.8 sind jedoch erheblich.

Wie zuvor bei den stationären Untersuchungen entsprechen sowohl die relativen Unterschiede zwischen den Reihen als auch die Streuungen der Leistung qualitativ weitestgehend den Boxplots der Gelenkwellendrehmomente. Die Multiplikation mit der Drehzahl verstärkt die Streuung, wodurch weniger Werte als Ausreißer betrachtet werden.

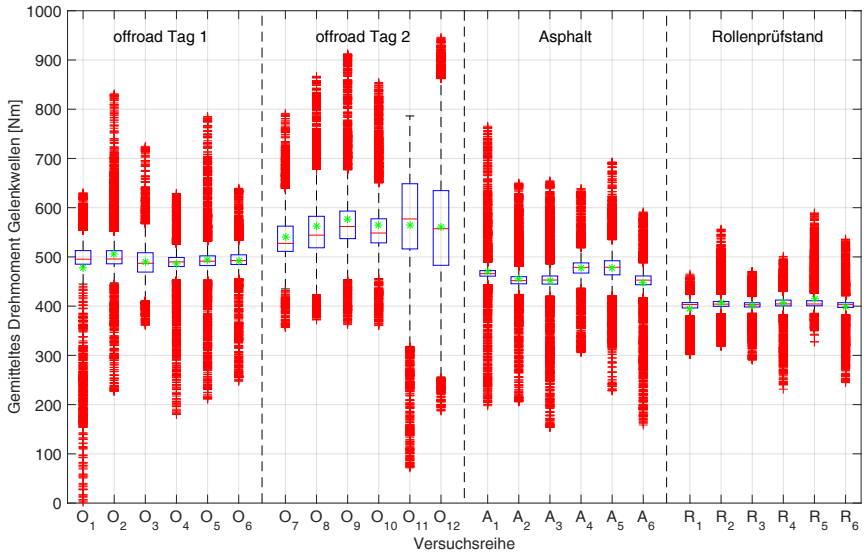


Bild 5.9: Streuung der Gelenkwellendrehmomente bei Beschleunigungsversuchen auf verschiedenen Untergründen

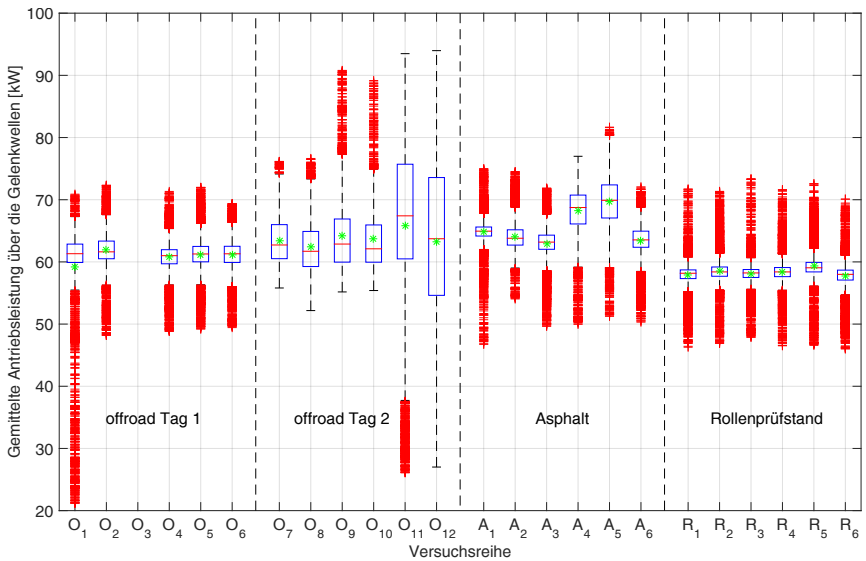


Bild 5.10: Streuung der Leistung über die Gelenkwellen bei Beschleunigungsversuchen auf verschiedenen Untergründen

Bei der Auswertung der Analyseergebnisse in den Tabellen 5.6 bis 5.8 fällt zunächst auf, dass Standardabweichung und vierfacher Interquartilabstand wesentlich größere Werte als bei den stationären Experimenten aufweisen. Außerdem sind die Unterschiede in der Streuung zwischen den verschiedenen Untergründen deutlich ausgeprägter. Gleichzeitig treten während der Beschleunigungsphase wesentlich höhere Absolutwerte auf, was den Mittelwerten nur bedingt anzusehen ist.

Tabelle 5.6: Analyseergebnisse der Gelenkwellendrehzahl bei dynamischen Versuchen auf verschiedenen Untergründen

Auswertekriterium	offroad Tag 1	offroad Tag 2	offroad gesamt	Asphalt	Rollen- prüfstand
<b>Anzahl Referenzversuche</b>	$5 \times 5$	$6 \times 5$	$11 \times 5$	$6 \times 5$	$6 \times 5$
<b>Mittelwert Referenzversuche</b>	$1344 \frac{1}{\text{min}}$	$1153 \frac{1}{\text{min}}$	$1246 \frac{1}{\text{min}}$	$1555 \frac{1}{\text{min}}$	$1549 \frac{1}{\text{min}}$
<b>Standardabweichung</b>	$45,7 \frac{1}{\text{min}}$	$105 \frac{1}{\text{min}}$	$146 \frac{1}{\text{min}}$	$43,3 \frac{1}{\text{min}}$	$21,7 \frac{1}{\text{min}}$
<b>vierfacher Interquartilabstand</b>	$178 \frac{1}{\text{min}}$	$512 \frac{1}{\text{min}}$	$801 \frac{1}{\text{min}}$	$221 \frac{1}{\text{min}}$	$94,2 \frac{1}{\text{min}}$
<b>Wahrscheinlichkeit Varianzanalyse</b>	25 %	18 %	1,1 E-6 %	6,2 %	1,1 %
<b>Normalverteilung des Experiments</b>	KST X	KST X	KST X	KST X	KST X
<b>Schärfe bei <math>5 \times 5</math> Versuchen pro Experiment</b>	$\pm 35 \frac{1}{\text{min}}$	$\pm 80 \frac{1}{\text{min}}$	$\pm 144 \frac{1}{\text{min}}$	$\pm 33 \frac{1}{\text{min}}$	$\pm 22 \frac{1}{\text{min}}$

In der Varianzanalyse können die Versuchsreihen des Rollenprüfstandes trotz der geringeren Streuung keiner gemeinsamen Grundgesamtheit zugeordnet werden. Das Experiment *offroad Tag 2* liefert dagegen für alle drei untersuchten Beobachtungsgrößen relativ hohe Wahrscheinlichkeiten. Zur Erläuterung dieser scheinbar widersprüchlichen Ergebnisse sei an den Ansatz der Varianzanalyse, wie in Kapitel 4.3.4 beschrieben, erinnert. Bei dieser statistischen Methode wird die Summe der Varianzen innerhalb



der Versuchsreihen, der Summe der Varianzen zwischen dem Reihenmittel und dem Gesamtmittel gegenüber gestellt. Wenn der Quotient dieser beiden Summen, dividiert durch ihre jeweiligen Freiheitsgrade, die Prüfgröße überschreitet muss die Nullhypothese verworfen werden. Im konkreten Fall tritt genau eine derartige Überschreitung der F-Verteilung ein. Auf dem Rollenprüfstand sind die Abweichungen innerhalb einer Versuchsreihe zwischen den Versuchen und damit die Summe der Varianzen minimal. Relativ kleine Störeinflüsse und Änderungen der Systemeigenschaften führen dagegen zu verhältnismäßig großen Unterschieden zwischen den Versuchsreihen, weshalb die Nullhypothese abgelehnt werden muss. Demzufolge können auf dem Rollenprüfstand selbst kleinste Veränderungen mithilfe der Varianzanalyse festgestellt werden. Da in dem bewerteten Prüfprozess die Vorkonditionierung des Prüfsystems nicht vorgegeben ist, gibt es zwischen den Reihen beispielsweise in den Öl- und Reifentemperaturen Unterschiede, die zu signifikanten Abweichungen führen. Auf dem Feld sind dagegen die Streuungen innerhalb der Reihe durch den Untergrund bereits so stark, dass derartig feine Einflüsse zwischen den Reihen nicht mehr differenziert werden können. Demzufolge sind signifikante Abweichungen zwischen den Versuchsreihen eines Experiments ein Indiz für Verbesserungspotentiale im Prüfprozess.

Eine Normalverteilung ist für keine Beobachtungsgröße der transienten Experimente feststellbar. Dank der großen Gruppengröße ist dies jedoch unkritisch. Weil der Shapiro-Wilk-Test für so große Gruppen ungeeignet ist, wird lediglich der Kolmogorov-Smirnov-Test verwendet.

Bereits bei den Boxplots ist aufgefallen, dass die Versuchsreihen auf Asphalt teilweise deutliche Abweichungen in der Streuung wie auch im Gesamtniveau aufweisen. Dies führt bei der Auswertung der Antriebsleistung zu signifikanten Unterschieden und damit zu einer konservativen Abschätzung der erzielbaren Schärfe auf einem Signifikanzniveau von 0,1 %. Deshalb fällt das Asphalt-Experiment hier leicht zurück. Bei Untersuchungen in weiteren Betriebspunkten liegen die Werte von Asphalt erwartungsgemäß zwischen den Ergebnissen des Rollenprüfstandes und der Feldversuche des ersten Tages.

Tabelle 5.7: Analyseergebnisse des Gelenkwellendrehmomentes bei dynamischen Versuchen auf verschiedenen Untergründen

Auswertekriterium	offroad Tag 1	offroad Tag 2	offroad gesamt	Asphalt	Rollen- prüfstand
Anzahl Referenzversuche	6 × 5	6 × 5	12 × 5	6 × 5	6 × 5
Mittelwert Referenzversuche	492 Nm	563 Nm	532 Nm	432 Nm	404 Nm
Standardabweichung	60 Nm	105 Nm	94 Nm	51 Nm	28 Nm
vierfacher Interquartilabstand	100 Nm	297 Nm	357 Nm	112 Nm	38 Nm
Wahrscheinlichkeit Varianzanalyse	31 %	27 %	1,0 %	28 %	8,4 E-2 %
Normalverteilung des Experiments	KST X	KST X	KST X	KST X	KST X
Schärfe bei 5 × 5 Versuchen pro Experiment	± 46 Nm	± 80 Nm	± 94 Nm	± 39 Nm	± 28 Nm

Bei den transienten Versuchen können die Anforderungen an die Schärfe aus Kapitel 5.1 weitestgehend nicht erreicht werden. Dies ist angesichts der beträchtlich höheren Werte kaum verwunderlich. Bereits bei den stationären Experimenten zeichnete sich ein Anstieg der Streuung mit größeren Messwerten ab. Die hier ermittelten Trennschärfen für den Rollenprüfstand liegen dabei noch unterhalb der Trendlinien aus Abbildung 5.6. Obwohl beide Prüfprozesse sehr einfach gehalten wurden, wirken hier noch zu viele Störeinflüsse. Auf den Fahrzeugzustand rückführbare Anteile können durch eine einheitliche Vorkonditionierung der Maschine reduziert werden. So könnten zumindest auf dem Rollenprüfstand die geforderten Trennschärfen bei Drehzahl und Drehmoment erreicht werden. Für die Leistungsermittlung dürfte diese Maßnahme allerdings nicht ausreichen. Kapitel 5.2 zeigt auf, wie weit sich die Trennschärfe durch eine Auswertung von Direktvergleichen verbessern lässt.

Tabelle 5.8: Analyseergebnisse der Antriebsleistung über die Gelenkwelle bei dynamischen Versuchen auf verschiedenen Untergründen

Auswertekriterium	offroad Tag 1	offroad Tag 2	offroad gesamt	Asphalt	Rollen- prüfstand
<b>Anzahl Referenzversuche</b>	5 × 5	6 × 5	11 × 5	6 × 5	6 × 5
<b>Mittelwert Referenzversuche</b>	61,3 kW	63,7 kW	62,4 kW	65,3 kW	58,4 kW
<b>Standardabweichung</b>	4,4 kW	9,2 kW	7,2 kW	4,2 kW	2,5 kW
<b>vierfacher Interquartilabstand</b>	11 kW	32 kW	19 kW	17 kW	6,3 kW
<b>Wahrscheinlichkeit Varianzanalyse</b>	23 %	6,7 %	0,75 %	0,14 %	1,9 %
<b>Normalverteilung des Experiments</b>	KST X	KST X	KST X	KST X	KST X
<b>Schärfe bei 5 × 5 Versuchen pro Experiment</b>	± 3,3 kW	± 6,9 kW	± 7,1 kW	± 4,1 kW	± 2,5 kW

Insgesamt kann mit dem ausgeführten Vergleich der verschiedenen Untergründe sehr gut auf die Störeinflüsse bei Fahrversuchen aufgrund der Umgebungsbedingungen geschlossen werden. Auf dem Rollenprüfstand sind die Störeinflüsse der Umgebung minimal und können sehr gut reproduziert werden. Folglich muss der Großteil der Varianz auf das Versuchsfahrzeug sowie den Maschinenbediener zurückzuführen sein. Auf Asphalt kommen die Unregelmäßigkeiten der Fahrbahneigung und -struktur hinzu. Der entscheidende Faktor sind hier jedoch meistens die direkten Wittereinflüsse. Diese bedingen einen Streuungsanteil, der 50 bis 100 % der Streuung aufgrund des Fahrzeugzustandes entspricht. In Summe ergibt sich dadurch nahezu eine Verdoppelung der Standardabweichung.

Indessen sind auf unbefestigten Untergründen die Störeinflüsse der Fahrbahn ausgeprägter. Außerdem hat hier das Wetter über die Bodenfeuchtigkeit einen zusätzlichen, indirekten Einfluss. Je nach Untersuchungsgröße

bewirkt der Wechsel von Asphalt auf ein trockenes Feld einen weiteren Anstieg der Standardabweichung um 20 bis 200 % der Fahrzeugstreuungen. Ist der nachgiebige Untergrund dagegen feucht, steigt die Streuung im Fahrversuch so weit, dass die übrigen Störeinflüsse kaum noch eine Rolle spielen. Im Beispiel verursachte die höhere Bodenfeuchtigkeit einen zusätzlichen Streuungsanteil von 200 bis 300 % der Standardabweichung infolge des Fahrzeugzustandes. Zusammen genommen ergibt sich auf dem feuchten Feld eine bis zu sechsmal höhere Streuung im Vergleich zum Rollenprüfstand.

## **5.2 Präzision bei Direktvergleichen auf verschiedenen Untergründen**

Da die Anforderungen an die Trennschärfe nur von einem Teil der Prüfprozesse erreicht werden konnte, wird in diesem Kapitel dargelegt, welche Verbesserungen eine Versuchsdurchführung und Auswertung von Direktvergleichen in diesem Anwendungsfall bringen. Durch das Herausrechnen eines Teils der Störeinflüsse kann die erzielbare Schärfe mitunter beträchtlich verbessert werden. In einigen Fällen sind diese Störeinflüsse jedoch nicht signifikant und die Auswertung als Direktvergleich führt zu einer verminderten Trennschärfe.

### **5.2.1 Direktvergleiche stationärer Versuche**

Zur Gegenüberstellung der Auswerteverfahren wurden die Beobachtungswerte der stationären Versuche hinsichtlich der erzielbaren Trennschärfe in Direktvergleichen ausgewertet. Die Abschätzung erfolgt gemäß der Beschreibung in Kapitel 4.4.5. Allerdings stehen aus diesen Versuchen nicht, wie vorgeschlagen, je 10 Einzelversuche bei 3 Versuchsbedingungen zur Verfügung. Stattdessen bilden die jeweils 6 Versuchsreihen aus Kapitel 5.1.1 die Basis der Referenzversuche. Da für die Abschätzung eine gerade Versuchsanzahl pro Reihe erforderlich ist, wurden hier nur die ersten 4 Einzelversuche ausgewertet. Die erzielbare Schärfe ist für einen Direktvergleich mit je 5 Einzelversuchen bei 6 Prüfbedingungen abgeschätzt. Zur Gewährleistung der Vergleichbarkeit wurden die Abschätzungen der einfaktoriellen

Varianzanalyse ebenfalls auf einen Versuchsplan mit  $6 \times 5$  Einzelversuchen angepasst.

Im Diagramm 5.11 sind die so gewonnenen Ergebnisse aus einfaktorieller Varianzanalyse und Direktvergleich gegenübergestellt. In den meisten Fällen führt die Versuchsdurchführung und Auswertung in Direktvergleichen zu einer Verbesserung der Trennschärfe. Insbesondere bei signifikanten Unterschieden zwischen den Versuchsreihen kann mit Direktvergleichen häufig eine bessere Schärfe erreicht werden. So wird nun selbst bei der gemeinsamen Auswertung der offroad Versuche die Anforderungen an die Trennschärfe in diesem Betriebspunkt erfüllt.

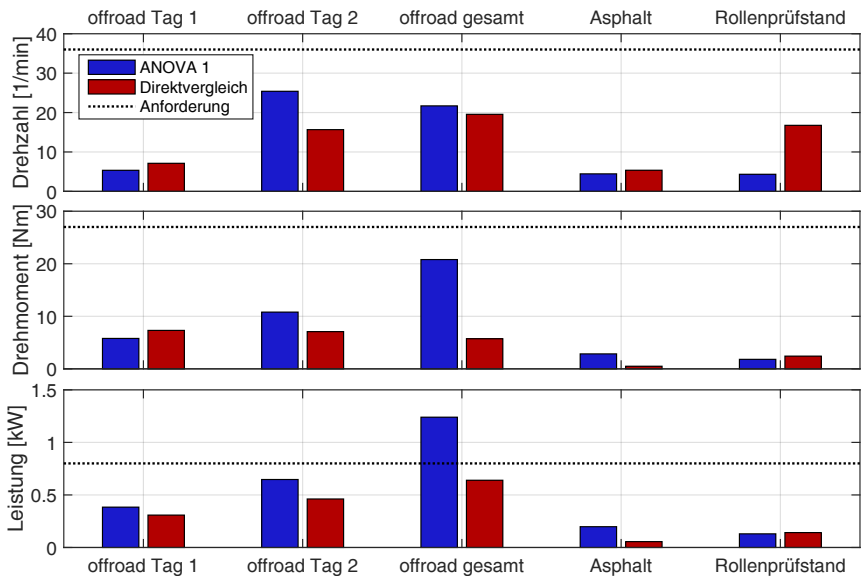


Bild 5.11: Erzielbare Schärfe bei stationären Versuchen mit der einfaktoriellen Varianzanalyse gegenüber Direktvergleichen

In einigen Fällen wird für den Direktvergleich eine geringere Trennschärfe gegenüber der einfaktoriellen Varianzanalyse abgeschätzt. Aufgrund der wenigen Beobachtungswerte pro Versuchsreihe können Interaktionen und Varianzinhomogenitäten überschätzt werden. Die Auswertung eines umfangreicheren Referenzexperimentes mit  $4 \times 12$  Versuchen kommt für den gleichen Betriebspunkt auf dem Rollenprüfstand zu höheren Trennschär-

fen. Diese Abschätzung liefert für  $6 \times 5$  Einzelversuche als kleinste unterscheidbare Differenzen  $\Delta n = 3,5 \frac{1}{\text{min}}$ ,  $\Delta M = 0,8 \text{ Nm}$  und  $\Delta P = 60 \text{ W}$ . An dieser Stelle wird ein weiteres Mal deutlich, wie wichtig eine umfangreiche Datengrundlage ist. Mit 20 bis 30 Beobachtungswerten pro Gruppe ist aus statistischer Sicht lediglich das Minimum erfüllt.

Im vorgestellten Beispiel sind die Faktorstufen des Einflussfaktors  $B$  ausschließlich durch unbeabsichtigte Störeinflüsse bedingt. Darüber hinaus kommt das Auswerteverfahren *Direktvergleich* mit wesentlich ausgeprägteren, zufälligen Faktorstufen zurecht. So kann beispielsweise die Wirkung einer Fahrzeugoptimierung mit verschiedenen Umschlagsmaterialien, Fahrern oder Fahrbahnbedingungen ausgewertet werden. Ein so gewonnenes Ergebnis bildet einen größeren Einsatzbereich des Untersuchungsfahrzeuges ab und gewinnt so an Aussagekraft. Wenn die zweite Einflussgröße gezielt manipuliert werden kann, sollte diese als unabhängiger Faktor in einer zweifaktoriellen Varianzanalyse ausgewertet werden. Aber auch hier besteht die Möglichkeit, weitere, zufällige Störeinflüsse der Versuchsbedingungen als abhängige Größe in einem Direktvergleich zu berücksichtigen.

## 5.2.2 Direktvergleiche transienter Versuche

Transiente Versuche sollten gemäß Kapitel 4.4.6 mit einer gekoppelten Varianzanalyse ausgewertet werden. Auch hier lässt sich das Konzept des Direktvergleichs anwenden. Analog zur Auswertung für stationäre Versuche wurde die erzielbare Schärfe auf der vorhandenen Datenbasis abgeschätzt und die Ergebnisse für  $6 \times 5$  Einzelversuche im Balkendiagramm 5.12 gegenübergestellt.

Bei den verwendeten Datensätzen führt ein Direktvergleich überwiegend zu einer deutlichen Verschlechterung der Trennschärfe. Die auftretenden Verbesserungen sind dagegen verschwindend. Die Wahl einer Datenbasis mit  $3 \times 10$  Einzelversuchen bringt kaum eine Veränderung der abgeschätzten Schärfe. Dessen ungeachtet ist eine Abschätzung basierend auf größeren Reihen zuverlässiger.

Im vorgestellten Beispiel bringt ein Direktvergleich der transienten Versuche keine Vorteile. Bei dieser Auswertemethode reduziert sich die Anzahl der Freiheitsgrade zur Bestimmung des  $(1 - \alpha)$ -Quantil der F-Verteilung

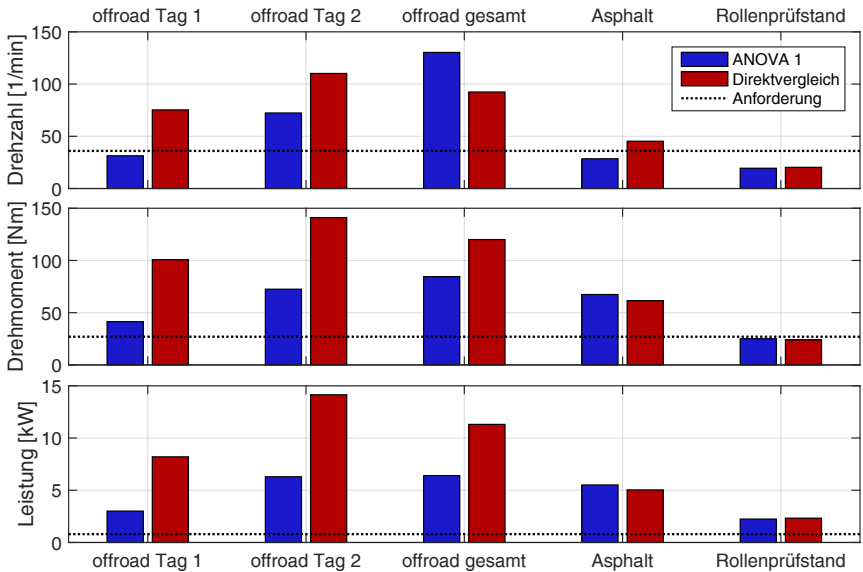


Bild 5.12: Erzielbare Schärfe bei dynamischen Versuchen mit der gekoppelten Varianzanalyse gegenüber Direktvergleichen

beträchtlich. Bei diesem Anschauungsbeispiel müssen 5 anstelle von 58 Freiheitsgraden eingesetzt werden. Die Auswirkung dieses Unterschieds ist mit  $\alpha = \frac{1}{5} \cdot 5\%$  in der Gegenüberstellung der Gleichungen 5.4 und 5.5 zu sehen.

$$\pm |\bar{Y}_{1..} - \bar{Y}_{2..}| = \sqrt{0,47 \cdot MQS_{Ref}} \quad (5.4)$$

$$\pm |\bar{Y}_{1..p} - \bar{Y}_{2..p}| = \sqrt{1,08 \cdot MQS_{in}} \quad (5.5)$$

Demzufolge lohnt sich die Auswertung als Direktvergleich für dieses Beispiel erst, wenn die mittlere Abweichungsquadratsumme zwischen den Messwerten und dem durchschnittlichen Datenverlauf  $MQS_{Ref}$  ca. das 2,3-fache der mittleren Abweichungsquadratsumme der Interaktion  $MQS_{in}$  erreicht. Ein Indiz für ein derartiges Verhältnis der mittleren Abweichungsquadratsummen sind signifikante Abweichungen zwischen den Versuchsreihen. Das Auswerteverfahren Direktvergleich kann seine Stärke vor allem bei Untersuchungen mit zufällige Einflussfaktoren mit ausgeprägter Wir-

kung auf die Zielgröße ausspielen. Ein Anwendungsbeispiel ist die Berücksichtigung unterschiedlicher Bodeneigenschaften oder Pflanzenbestände in verschiedenen Feldbereichen. Ein Boxplot Diagramm entsprechend Abbildung 5.7 unterstützt bei der Entscheidung zwischen den Auswerteverfahren. Bei großen Niveau- oder Varianzunterschieden sowie bei signifikante Abweichungen zwischen den Versuchsreihen sollte die Verwendung eines Direktvergleichs in Betracht gezogen werden.

Zusammenfassend können die eingangs gesetzten Anforderungen an den Prüfprozess bezüglich der Wiederholpräzision nur für kleine Absolutwerte erreicht werden. Ursache ist die zunehmende Streuung mit steigenden Absolutwerten. Unter Annahme eines proportionalen Anstiegs kann mit der gekoppelten Varianzanalyse von Werteverläufen ähnliche Trennschärfen wie bei der Auswertung einzelner Beobachtungswerte erzielt werden. Feste Untergründe wie Asphalt oder ein Rollenprüfstand eignen sich dabei auch für längere Messkampagnen, während ein trockenes Feld nur über einen begrenzten Messzeitraum vergleichbare Rahmenbedingungen bietet. Erwartungsgemäß kann auf dem Rollenprüfstand die höchste Präzision erreicht werden. Mithilfe eines Direktvergleiches kann bei größeren, zufälligen Abweichungen zwischen den Reihen die Trennschärfe weiter verbessert werden.

### **5.3 Grenzen der Präzisionsabschätzung**

Die Abschätzung der Prüfprozesspräzision erfolgt immer unter der Annahme, dass die betrachteten Referenzversuche das gesamte Spektrum der möglichen Störeinflüsse beinhaltet. Aufgrund der begrenzten Versuchszahl kann die Streuung der Zielgröße jedoch nie genau bestimmt werden. Durch die Berücksichtigung der tatsächlichen Freiheitsgrade und die Abschätzung der erzielbaren Schärfe bezüglich eines höheren Signifikanzniveaus, ist das Ergebnis robust gegen kleinere Veränderungen.

Bei einer Modifikation des Prüfprozesses verändert sich jedoch häufig die Streuung innerhalb eines Experiments wodurch die Abschätzung der Trennschärfe hinfällig wird. Dabei ist entscheidend, wie stark sich die Streuung im Verhältnis zur Gesamtstreuung ändert. Welche Veränderun-



gen die Prozessvarianz vergrößern ist primär vom Prüfprozess abhängig und kann deshalb nicht pauschal festgelegt werden. Hierzu werden im folgenden einige Beispiele erörtert.

Die Auswertung der stationären Versuche in Kapitel 5.1.1 hat gezeigt, dass die Streuung der Zielgröße stark mit dem **Betriebspunkt** korreliert. In vielen Fällen verschlechtert sich die erzielbare Schärfe proportional zur Zielgröße. Demnach ist eine Umrechnung der Trennschärfe auf andere Betriebspunkte zwar denkbar aber nicht zuverlässig. Aus diesem Grund sollten die Referenzmessungen über alle relevanten Betriebspunkte durchgeführt werden.

Außerdem wurde mit den Beispielen gezeigt, dass der **Untergrund** die Varianz der Zielgrößen stark beeinflusst. Teilweise sind mit dem Wechsel der Fahrbahn auch Änderungen weiterer Rahmenbedingungen verknüpft. Eine Verlagerung der Versuche von Asphalt auf den Rollenprüfstand zeigte wesentlich kleinere Auswirkungen als ein Wechsel des Feldbereiches oder gar des Feldes.

Bei stationären Versuchen sowie bei den vorgestellten transienten Prüfprozessen ist keine wesentliche Veränderung bei einem **Fahrerwechsel** zu erwarten. Sobald jedoch mehrere Steuervorgaben erfolgen oder der Maschinenbediener als Regler im Prüfprozess wirkt, ist mit erheblichen Abweichungen bei den Varianzen zu rechnen. An dieser Stelle sei nochmals auf die Arbeiten von HOLLÄNDER [Hol98] und FRANK [FSA12] verwiesen, die sich mit dem Einfluss der Fahrer auseinandersetzen. Bei Prüfprozessen mit Bedienerinfluss sollte mithilfe der gekoppelten Varianzanalyse die Werteverläufe der Bedienvorgaben überprüft werden. So kann festgestellt werden, ob signifikante Abweichungen der Zielgröße tatsächlich auf den veränderten Einflussfaktor und nicht nur auf abweichende Bedienvorgaben zurückzuführen sind.

Ferner hat das **Fahrzeug** selbst einen Einfluss auf die Streuungen im Prüfprozess. Dies kann auf dem Rollenprüfstand, auf dem die Varianz im Wesentlichen durch das Fahrzeug bedingt ist, zum Tragen kommen. In einem Feldversuch dürfte dieser Einfluss nicht ins Gewicht fallen. Zu den Unterschieden baugleicher Fahrzeuge oder den Abweichungen zwischen verschie-

denen Fahrzeugen gleicher Fahrzeugart und Leistungsklasse liegen dem Autor keine Informationen vor.

Des Weiteren haben **Messtelle** und **Messsystem** einen erheblichen Einfluss auf die erzielbare Schärfe. Auf dem Rollenprüfstand können das Messsysteme des Fahrzeuges mit dem des Prüfstands verglichen werden. Dazu wird die Rollendrehzahl auf die Gelenkwelldrehzahl umgerechnet. Ein Vergleich der Drehmomente ist aufgrund der Momentenregelung des Prüfstandes nicht sinnvoll. Das Drehzahlsignal des Rollenprüfstandes zeigt wesentlich weniger Schwingungen. Hier wirkt das Trägheitsmoment der Rollen als Tiefpassfilter. Daraus resultiert eine geringere Streuung und eine bessere Trennschärfe. Im Beispiel wurde die erzielbare Schärfe auf 50 bis 70 % verfeinert.

Auf ähnliche Art würde eine **Glättung der Beobachtungswerte** das Ergebnis manipulieren. Die gezielte Glättung eines Messdatenverlaufs einer transienten Zielgröße hebt die Varianzanalyse aus und verfälscht das Ergebnis. Demzufolge ist für die Auswertung von Datenverläufen auf jegliche glättende Datenmanipulationen zu verzichten. Die Mittelwertbildung über die Messdaten eines stationären Versuches zur Bildung eines Beobachtungswertes ist dagegen zulässig.

Eine **Übertragbarkeit der Ergebnisse** einer Varianzanalyse oder einer Präzisionsabschätzung auf andere Zielgrößen ist in der Regel nicht gegeben. Störeinflüsse haben auf verschiedene Beobachtungsgrößen ganz unterschiedliche Wirkungen. Große Streuungen oder signifikante Unterschiede im Drehmoment führen nicht zwangsläufig zu einem vergleichbaren Verhalten bei der Drehzahl, der Öltemperatur oder einer beliebigen weiteren Beobachtungsgröße.

Ebenso lassen die Ergebnisse eines Prüfprozesses nur bedingt Schlüsse auf andere Prüfprozesse zu. Können in einem sehr einfachen Prüfprozess die Anforderungen an die Präzision nicht erfüllt werden, ist für einen komplexeren Prüfprozess mit demselben Ergebnis zu rechnen. Umgekehrt wird durch eine Reduktion der Komplexität und damit der Störeinflüsse die Präzision des Prüfprozesses im Allgemeinen verbessert. Quantitative Aussagen sind hier jedoch kaum möglich.

Die aufgezeigten Grenzen der Präzisionsabschätzung verdeutlichen, dass der Prüfprozess gegebenenfalls sehr empfindlich auf zusätzliche Störeinflüsse reagiert. Aus diesem Grund sieht der dritte Schritt der Prüfprozessbeurteilung in Kapitel 4.2.2 vor, die Präzision für die tatsächlich wirkenden Prüfbedingungen zu bestätigen. Auf diese Weise wird sichergestellt, dass unerwarteten Störeinflüsse entdeckt werden und die Zuverlässigkeit der Ergebnisse nicht beeinträchtigen. Möglicherweise ergibt sich innerhalb der betrachteten Messkampagne sogar eine bessere Trennschärfe.

Die Idee einer Datenbank mit spezifischen Prüfprozesselementen und den daraus resultierenden Streuungsanteilen musste im Verlauf der Auseinandersetzung mit der Thematik verworfen werden. Die ermittelten Werte wären aufgrund der zahlreichen Abhängigkeiten und der damit verbundenen schlechten Übertragbarkeit kaum für weitere Anwendungen nutzbar. Sinnvoll scheint dagegen eine Zusammenstellung existierender Prüfprozesse in Kombination mit der erreichten Trennschärfe aller verfügbaren Messgrößen. Mithilfe einer derartigen Sammlung ließe sich zügig abschätzen, ob die Präzision des angedachten Prüfprozesses in der richtigen Größenordnung liegen kann. Eine solche Datenbank sollte zumindest innerhalb der Organisation angelegt werden. Transparente, hochwertige Veröffentlichungen können zur Ergänzung einer derartigen Datenbank hinzugezogen werden.



## 6 Zusammenfassung

Selbst wenn heutzutage viele Entwicklungen im Bereich der mobilen Arbeitsmaschinen rechnergestützt mithilfe von Simulationen durchgeführt werden, bleibt die Validierung und Feinjustierung des entwickelten Systems im Gesamtfahrzeug ein essentieller Teil des Entwicklungsprozesses. Wie in Kapitel 2.5 verdeutlicht, existieren allerdings nur für wenige mobile Arbeitsmaschinen standardisierte und allgemein anerkannte Prüfprozesse. Deshalb werden in vielen Projekten neue Prüfprozesse generiert, um die Parametrierung des Gesamtsystems zu optimieren und den Mehrwert der Entwicklung herauszustellen. Eine Analyse veröffentlichter Gesamtfahrzeuguntersuchungen zeigt, dass ein Nachweis der Eignung und der Qualität des kreierte Prüfprozesses nur unvollständig oder gar nicht mit den Messergebnissen dokumentiert ist. Das legt die Vermutung nahe, dass bei diesen Untersuchungen nur unzureichende qualitätssichernde Maßnahmen durchgeführt wurden. So bleibt bei den Messergebnissen zumindest teilweise ungewiss, in wie weit sie sich auf einen individuellen Fahrzeugeinsatz übertragen lassen, unter welchen Bedingungen diese entstanden sind, ob geeignete Messtechnik eingesetzt wurde, die beobachteten Unterschiede tatsächlich signifikant sind und mit welcher Unsicherheit das Ergebnis belegt ist.

Demzufolge bedarf es einer einheitlichen Vorgehensweise die Qualität von Prüfprozesse für mobile Arbeitsmaschinen festzustellen und damit ihre Eignung für Untersuchungen nachzuweisen. Diese Forderung erfüllt die in dieser Arbeit erläuterten Methode zur Beurteilung der Prüfprozesseignung von Fahrzeugversuchen mit mobilen Arbeitsmaschinen.

Der Ablauf zur Beurteilung der Prüfprozesseignung ist auf Seite 53 in Abbildung 4.1 in einem Flussdiagramm veranschaulicht. Die Bewertung des Prüfprozesses erfolgt bei dieser Methode hinsichtlich der Kriterien

- Aussagekraft
- Genauigkeit
- Präzision

Eine allgemeingültige, objektive Einstufung der Aussagekraft ist in Anbetracht der zahlreichen Nutzungsmöglichkeiten, Einsatz- und Betriebsbedingungen einer mobile Arbeitsmaschine jedoch kaum möglich. Deshalb sind in Kapitel 4.1 eine Reihe von Aspekten genannt, die zusammen mit den Versuchsergebnissen dokumentiert werden sollen. Diese Informationen ermöglichen es dem Interessenten die Aussagekraft für seinen Anwendungsfall individuell abzuschätzen.

Die Beurteilung der Genauigkeit erfolgt anhand der Auflösung sowie der systematischen Messabweichungen des Messsystems. Bezugsgröße ist die kleinste festzustellende Abweichung, die mindestens das Fünffache der Auflösung und des Bias betragen sollte. Bei berechneten Größen ist die maximale Abweichung mithilfe der linearen Fehlerfortpflanzung zu ermitteln. Zur Bewertung der Präzision werden Versuchsdurchführung sowie Auswerteverfahren analysiert. Aus statistischen Gründen sollte jeder Versuch mindestens 25-mal wiederholt werden. Bei der Auswertung der Versuchsergebnisse mithilfe einer Varianzanalyse können so bereits relativ kleine Abweichungen zuverlässig als signifikante Unterschiede erkannt werden. Zur Verbesserung der erzielbaren Schärfe ist ein Vorgehen für Direktvergleiche beschrieben. Hier wird die Eignung verschiedener Varianten der Varianzanalyse diskutiert. Ebenso wird ein Auswerteverfahren zur Analyse von Datenverläufen auf Basis einer Varianzanalyse entwickelt und vorgestellt. Die erzielbare Trennschärfe des Prüfprozesses wird letztendlich durch Umstellung der jeweils geeignete Varianzanalyse aus Referenzversuchen abgeschätzt. Die wesentlichen Einflussgrößen an dieser Stelle sind die Versuchsanzahl sowie die Streuung der Einzelversuche.

Die Anwendung der Methode auf reale Beispiele zeigt, dass eine sorgfältige Vorkonditionierung des Prüflings und der Prüfumgebung die Präzision erheblich verbessert. In einem Beispiel auf dem Rollenprüfstand konnte

so die Trennschärfe beim Drehmoment und der Leistung auf ein Drittel verbessert werden. Ein Vergleich von Fahrversuchen auf verschiedenen Untergründen verdeutlicht den Einfluss der Rahmenbedingungen auf die Prüfprozesspräzision. Außerdem kann die Ergebnisstreuung durch diesen Vergleich in Anteile des Versuchsfahrzeuges, der Wettereinflüsse sowie in Anteile des Reifen-Boden Kontakts differenziert werden. Bei der Durchführung und Auswertung eines Direktvergleichs lassen sich diese Störeinflüsse teilweise herausrechnen. Dieses Verfahren verbessert jedoch nicht immer die Trennschärfe. Bei nur geringen Abweichungen zwischen den Einzelversuchen beziehungsweise zwischen den Messdatenverläufen ergibt sich bei einem Direktvergleich häufig eine schlechtere Differenzierbarkeit.

Die Präzision des Prüfprozesses sollte nicht nur vorab abgeschätzt, sondern vor allem während der eigentlichen Versuchsdurchführung überprüft werden, um Zuverlässigkeit und Schärfe der konkreten Messkampagne zu evaluieren.

Die dargelegte Methode zur Beurteilung der Prüfprozesseignung ist direkt auf beliebige praktische Untersuchungen in anderen Anwendungsfeldern übertragbar. Denn für alle praktischen Untersuchungen gilt, dass die Genauigkeit der Messtechnik alleine keine Rückschlüsse auf die Differenzierbarkeit der Messergebnisse zulässt. Stattdessen sollte die erzielbare Trennschärfe des jeweiligen Prüfprozesses separat untersucht und angegeben werden. Sicher gibt es Anwendungen, bei denen die Ergebnisstreuung ignoriert werden kann. Im Bereich der mobilen Arbeitsmaschinen ist die auftretende Messwertstreuung in der Regel jedoch ein Vielfaches der Genauigkeit und keinesfalls vernachlässigbar.





## 7 Ausblick

Detaillierte Erläuterungen zur Entstehung von Versuchsergebnissen sowie zur erreichten Trennschärfe im Prüfprozess fördern die Transparenz und Glaubwürdigkeit von Entwicklungsergebnissen. Durch die kritische Auseinandersetzung mit dem Untersuchungsprozess steigt die Qualität einer Veröffentlichung. Darüber hinaus schärfen entsprechende Publikationen das Bewusstsein der Leser für die Qualität eigener Versuche. Gerade unabhängige Einrichtungen, wie die Hochschulen oder das DLG-Testzentrum sollten hier vorangehen und die Qualität ihrer veröffentlichten Messergebnisse nachweisen. Eine Methode zum Beleg der Prüfprozesseignung ist in dieser Arbeit dargelegt.

Eine der größten Herausforderungen bleibt jedoch die Ausarbeitung adäquater Prüfprozesse, die für ein breites Spektrum realer Fahrzeugeinsätze repräsentativ sind und gleichzeitig die erforderliche Trennschärfe aufweisen. Entwicklung und Validierung eines aussagekräftigen Prüfprozesses gestalten sich jedoch verhältnismäßig aufwendig und können in vielen Projekten nicht angemessen geleistet werden. Hier bedarf es tadelloser Veröffentlichungen zur Orientierung. Publikationen mit sorgfältig erarbeiteten Prüfprozessen werden sich zügiger verbreiten und gegebenenfalls als Standard etablieren.

Bei besonders hohen Anforderungen an die Präzision des Prüfprozesses empfiehlt sich eine Verlagerung der Untersuchung auf einen Prüfstand. Hier können Fahrzeugzustand und Rahmenbedingungen sehr genau und reproduzierbar eingestellt werden. Die Lasten aus dem Arbeitsprozess werden in einer Simulation berechnet und mithilfe von Belastungsaggregaten auf das reale Fahrzeug aufgeprägt. Gute Modelle bilden dabei Abhängigkeiten von Geschwindigkeiten sowie weiteren Einflussparametern ab und stellen auch Lastschwingungen reproduzierbar dar. In diesem Fall unterstützen Erfahrung und Kompetenz des Prüfstandsbetreibers bei der Ent-

wicklung eines geeigneten Prüfprozesses sowie bei der Beurteilung der substituierten Schnittstellen. Diese Unterstützung wird mit einem methodischen Nachweis der Prüfprozesseignung vervollständigt. Mithilfe aktueller Prüfstandstechnik lässt sich der Zielkonflikt zwischen Praxisbezug und Reproduzierbarkeit beträchtlich entschärfen.

# A Anhang

## A.1 Versuchsfahrzeug Teleskoplader Claas Scorpion 7040

Die dargestellten Messungen gehen aus Versuchen mit einem Teleskoplader der Firma Claas, Model Scorpion 7040 hervor. Arbeits- und Fahrtrieb werden in dieser Maschine von einem 4 Zylinder Dieselmotor mit 100 kW angetrieben. Die Arbeitshydraulik ist als konventionelles hydraulisch-mechanisches closed center load sensing System ausgeführt. Der permanente Allradantrieb ist über ein hydrostatisches Getriebe mit dem Verbrennungsmotor verbunden. Während der Fahrt gibt die Getriebesteuerung automatisch die Übersetzung durch die stufenlose Verstellung der Schluckvolumina von Fahrpumpe und Fahrmotor vor. Mit einer Fahrstufenschaltung kann die Steuerverhalten der Verstelleinheiten geändert werden. Die Lenkung lässt sich zwischen Allradlenkung und Vorderachslenkung umschalten.

Tabelle A.1: Charakteristische Daten Versuchsfahrzeug Claas Scorpion 7040

<b>Eigenschaft</b>	<b>Wert</b>
Leistung (nach ISO 9249)	100 kW
Leergewicht	9445 kg
Volumen Standardschaufel gestrichen/gehäuft	1,25/1,6 m <sup>3</sup>
Überladehöhe	7050 mm
Fahrtrieb	Allradantrieb
zul. Höchstgeschwindigkeit	40 $\frac{\text{km}}{\text{h}}$
Reifengröße VA/HA	440/70 R24
Reifendruck VA/HA	3,5 bar/3,5 bar

Verwendung finden Teleskoplader sowohl in der Baubranche wie in der Landwirtschaft. Durch verschiedene Anbaugeräte kann diese Fahrzeugart Schüttgüter, wie auch Paletten oder andere Gebinde verladen und über kurze Distanzen transportieren. Während beim Einsatz in einer Biogasanlage vorwiegend innerhalb des Betriebsgeländes auf festem Grund gefahren wird, dominiert auf Baustellen oder bei Verladearbeiten auf dem Feld der offroad Anteil. Wie bei HUBER zu sehen, variieren die Anteile verschiedener Arbeitstätigkeiten von Betrieb zu Betrieb sehr stark [Hub10, S.48-51]. Dementsprechend ist ein Referenzzyklus lediglich für eine spezielle Arbeitstätigkeit oder einen ausgewählten Betrieb repräsentativ.

Aufgrund der Nutzung des Fahrzeuges in mehreren Forschungsprojekten [Hub10, NR14] sind die beiden Antriebsstränge des Teleskopladers mit diversen Sensoren ausgestattet. Diese erfassen unter anderem die Hydraulikdrücke der Zylinder und des hydrostatischen Getriebes sowie die Drehzahlen des Motors, der Gelenkwellen und der Räder, die Position der Arbeitskinematik und die Drehmomente in den Gelenkwellen zur Vorder- und Hinterachse. Außerdem besitzt der Teleskoplader nach diversen Modifikationen neben einem höheren Leergewicht von 9445 kg auch einige Zusatzfunktionen. Beispielsweise können durch ein Kurzschlussventil im hydrostatischen Getriebe Hoch- und Niederdruckseite miteinander verbunden und somit eine Art Neutralgang geschaltet werden.

Zur Bestimmung von absoluter Fahrzeuggeschwindigkeit und Position ist bei allen Versuchsfahrten ein GNSS Receiver 2022-S03 der Firma ppm auf dem Fahrzeug montiert. Hierbei handelt es sich um ein Satelliten gestütztes Positionserfassungssystem, das laut Hersteller eine Messgenauigkeit im Bereich  $\pm 10$  cm mit einer zeitlichen Auflösung von 5 Hz erreicht [PPM12, S.3]. Diese Positionsgenauigkeit erzielt das Real Time Kinematik (RTK) System mithilfe von Korrektursignalen von festen Referenzstationen, die über das Mobilfunknetz empfangen werden.

## **A.2 Rollwiderstandsbeiwerte aus der Literatur**

In Tabelle A.2 sind einige Werte für den Rollwiderstandsbeiwert  $f_R$  auf verschiedenen Untergründen aus der Literatur zusammengetragen.

Tabelle A.2: Rollwiderstandsbeiwerte für mobile Arbeitsmaschinen

Art	Reifen Dimension	Fahrbahn	f_R	Quelle
PKW-Reifen	ohne Angabe	Asphalt	0,0095	[Sch13, S.141] [Lei09, S.88f.]
LKW-Reifen	11 R 22,5	Asphalt	0,0065-0,008	[Wal07, S.29]
militärische Gelände- bereifung	14 R 20 mil	Sand	0,04-0,05	[Har02, S.136]
		Lehm I	0,16-0,23	
		Lehm II	0,2-0,26	
AS-Reifen	540/65 R 30	Grünland	0,072	[Sch06, S.75]
		Stoppelfeld	0,0844	
		gegrubbert gepflügt	0,1166 0,1446	
AS-Reifen	520/70 R 34 TM700	Asphalt	0,034	[Bar96, S.125f.]
	19.6 R 34 TM190	Asphalt	0,0165	
	19.6 R 34 TM300S	Asphalt sandiger Lehm	0,03 0,1	
AS-Reifen	19.6 R 34	Stahlband	0,023	[Pie97, S.133ff.]
	480/70 R 34	Stahlband	0,025	
land- wirtschaft- liche Reifen	ohne Angabe	Asphalt	0,015	[Sch06, S.144]
		Feldweg; festgefahrene Fahrspur	0,05	
		Trockener Stoppelacker	0,08	
		Gepflügelter Acker, feucht; Saatbeet	0,12	
		Gepflügter Acker, sehr locker	0,16	
ohne Angabe	ohne Angabe	fester Asphalt, Beton, Kopfsteinpflaster	0,005-0,015	[Wal07, S.39]
		gewalzter, fester Schotter; angefahren, welliger Asphalt	0,02-0,03	
		geteeter, angefahren, welliger Schotter	0,03-0,04	
		sehr gute Erdwege	0,04-0,05	
		Erde	0,05-0,15	
		Sand	0,15-0,35	
ohne Angabe	ohne Angabe	Betonstraßen oder harte, glatte Erdwege	0,02	[Küh56, S.202]
		Glatter trockener fester Erdboden mit Kies, gut instand gehalten und eingeeb- net	0,023	
		Trockener Boden, mit Kies vermischt, mittelfest gelagert	0,03	
		Erdweg, mäßig gepflegt, feucht gehalten. Geringes Eindringen der Oberfläche	0,035	
		Weicher gewachsener Boden mit tragen- der Oberflächendecke (Grasnarbe)	0,04	
		Häufig befahrener Erdweg, schlecht gepflegt	0,043	
		Feuchte schmierige Oberfläche auf festerem Untergrund	0,045	
		Normale Erdfahrbahn, zerfahren, wenig gepflegt	0,05	
		Ausgetrockneter sandiger Boden	0,06	
		Weicher zerfurchter Mutterboden	0,075	
		Schwach verdichtete Kippe	0,08	
		Zerfahrener schlammiger Erdweg	0,09	
		Sandiger aufgeweichter Lehm	0,1	
		Lockerer Kies	0,11	
		Lockerer Sand erdfeucht	0,125	
Lockerer Sand trocken	0,15			
Zerfurchter weicher Boden	0,16			
Weicher schwammiger Untergrund	0,19			

## A.3 Datengrundlage für den Vergleich der Auswerteverfahren

Die Daten in Tabelle A.3 bilden die Basis für den Vergleich der verschiedenen Auswerteverfahren für Direktvergleiche sowie transiente Versuche. In Abbildung 4.6a sind diese Daten in Boxplots dargestellt.

Tabelle A.3: Basisdatensatz für den Vergleich der Auswerteverfahren

Faktor A	Faktor B					
	Stufe $B_1$	Stufe $B_2$	Stufe $B_3$	Stufe $B_4$	Stufe $B_5$	Stufe $B_6$
Stufe $A_1$	83,7613	93,6294	86,6460	88,1885	81,7522	83,4546
	83,3880	94,5120	88,6554	89,7711	83,0379	82,1031
	85,2061	93,2161	86,2223	88,6975	83,0298	83,8046
	83,7181	93,5968	89,6245	89,7384	82,1401	85,0722
	83,9252	94,6830	86,9722	90,2364	82,7637	83,7541
Stufe $A_2$	83,8472	95,0568	90,1219	89,2560	83,0580	82,8957
	83,0369	92,8001	87,5865	90,0184	80,9471	85,0474
	83,5309	94,9211	88,3445	89,9854	82,7974	83,5538
	85,2568	93,4549	86,7247	89,1418	83,5305	84,0302
	84,3351	93,3870	85,3237	88,2528	82,4071	82,6545

## A.4 Aussagekraft der Prüfprozesse der Anwendungsbeispiele

Tabellarische Beschreibung von Entstehung und Durchführung der Prüfprozesse für die Anwendungsbeispiele in Kapitel 5.

Tabelle A.4: Informationen zur Entstehung der Prüfprozesse der Anwendungsbeispiele

- Vorgehensweise zur Entwicklung des Prüfprozesses
  - Auswahl von stationären und transienten Fahranteilen als Gemeinsamkeit einer Vielzahl mobiler Arbeitsmaschinen
- Einsatzprofile
  - Umfang der Datenerhebung keine Datenerhebung zur Festlegung des Prüfprozesses
  - Einsatzbetrieb -
  - Arbeitsaufgaben -
  - Betriebsphasen -

Tabelle A.5: Informationen zur Durchführung der Anwendungsbeispiele

- Versuchsträger
  - Fahrzeug Teleskoplader Claas Scorpion 7040
  - Fahrzeuggröße charakteristische Fahrzeugdaten in Tabelle A.1
- Fahrer Laie, ca. 40 h Erfahrung mit dem Versuchsfahrzeug
- Vorkonditionierung
  - Fahrzeug mindestens 20 minütige Warmlaufphase mit nicht spezifiziertem Fahrzeugbetrieb
  - Umgebung Asphalt trocken, Feld bereits mehrfach überfahren, größerer Unebenheiten werden geplant
- Prüfprozess
  - Zielgrößen Leistung auf der Kardanwelle gemessen über Drehzahl (Hallsensor mit Frequenz-Spannungswandler) und Drehmoment (Dehnmessstreifen)
  - Arbeitsaufgabe Geradeausfahrt mit konstanter Gaspedalstellung (mechanischer Anschlag)

- Vereinfachungen keine Kurvenfahrt, Arbeitshydraulik unbetätigt
- Lastniveau stationäre Fahrt: überwiegend Teillastbetrieb (Radwiderstände), Abhängig vom mechanischen Gaspedalanschlag; Beschleunigung aus dem Stand: hohe Lasten (Beschleunigungswiderstände)
- Betriebsphasen ausschließlich Fahranteile auf verschiedenen Untergründen: Feld, Asphalt, Rollenprüfstand keine Standzeiten, keine Verzögerungsphasen; stationäre Fahrt bei konstanter Gaspedalstellung; Beschleunigung aus dem Stand
- Durchführung pro Experiment 6 Versuchsreihen mit je 5 Einzelversuchen; Betätigung Gaspedal gegen mechanischen Anschlag; geringe Lenkkorrekturen sofern für Geradeausfahrt erforderlich
- Qualität keine Anforderungen an die Arbeitsergebnisse
- Umweltbedingungen (ggf. als Wertebereich)
  - Wetter Wetterbedingungen in Tabelle 5.1
  - Kontaktstellen Reifen-Fahrbahn Kontakt vgl. Abbildung 5.1  
Feld: Podsol-Baunerde; hohe Bodenfeuchte, nicht bestimmt  
Asphalt: trocken, eben, griffig  
Rollenprüfstand: mit trockenem Straßenbelag vergleichbar



# Abbildungsverzeichnis

2.1	Wichtige Einflüsse auf die Messergebnisse . . . . .	6
2.2	Kräfte und Momente am beschleunigten Rad auf nachgiebigen Untergründen . . . . .	17
2.3	Be- und Entlastungsvorgang beim Multipasseffekt . . . . .	17
2.4	Testgespann aus Traktor und Messwagen im DLG PowerMix . . . . .	26
2.5	Motorkennlinien eines John Deere 5125R PowerShift Traktors . . . . .	27
2.6	Arbeitszyklus und Ergebnisse des „ <i>Liebherr-Normtests</i> “ . . . . .	30
2.7	Arbeitszyklen von Baumaschinen . . . . .	32
2.8	TÜV zertifizierte Leistungszyklen für Flurförderzeuge . . . . .	35
2.9	Bedeutung der Streuung bei einem Vergleich mit je fünf Ein- zelversuchen . . . . .	40
4.1	Ablauf zur Beurteilung der Prüfprozesseignung . . . . .	53
4.2	Gruppierungsebenen zur Strukturierung einer Untersuchung . . . . .	64
4.3	Statistische Kennwerte in Histogrammen und Boxplots . . . . .	70
4.4	Beobachtungswerte des Anwendungsbeispiels und deren Vertei- lung . . . . .	76
4.5	Erzielbare Schärfe in Abhängigkeit der Versuchsanzahl . . . . .	102
4.6	Beispieldaten zum Vergleich der Auswerteverfahren . . . . .	109
4.7	Datenverläufe zum Vergleich der Auswerteverfahren . . . . .	118
4.8	Auswahl des Auswerteverfahren in Abhängigkeit von Untersu- chungsgröße und Störeinflüssen . . . . .	124
5.1	Untersuchte Untergründe zur Abschätzung des Fahrbahnein- flusses . . . . .	128
5.2	Genauigkeit und Auflösung des Messsystems für die Leistung . . . . .	133
5.3	Gelenkwellendrehzahl der stationären Versuchsreihen . . . . .	134

5.4	Gesamtdrehmomente der Gelenkwellen der stationären Versuchsreihen . . . . .	135
5.5	Antriebsleistung über die Gelenkwellen der stationären Versuchsreihen . . . . .	138
5.6	Erzielbare Schärfe in verschiedenen Betriebspunkten . . . . .	141
5.7	Streuung der Gelenkwellendrehzahl bei Beschleunigungsversuchen auf dem Rollenprüfstand . . . . .	143
5.8	Streuung der Gelenkwellendrehzahl bei Beschleunigungsversuchen auf verschiedenen Untergründen . . . . .	144
5.9	Streuung der Gelenkwellendrehmomente bei Beschleunigungsversuchen auf verschiedenen Untergründen . . . . .	145
5.10	Streuung der Leistung bei Beschleunigungsversuchen auf verschiedenen Untergründen . . . . .	145
5.11	Erzielbare Schärfe bei stationären Versuchen mit der einfaktoriellen Varianzanalyse gegenüber Direktvergleichen . . . . .	151
5.12	Erzielbare Schärfe bei dynamischen Versuchen mit der gekoppelten Varianzanalyse gegenüber Direktvergleichen . . . . .	153

# Tabellenverzeichnis

4.1	Informationen zur Entstehung des Prüfprozesses . . . . .	54
4.2	Informationen zur Durchführung einer Untersuchung . . . . .	55
4.3	Gewichtungsfaktoren gemäß Position in der Rangwertreihe . . . . .	66
4.4	Beobachtungswerte für ein Anwendungsbeispiel . . . . .	72
4.5	Nomenklatur der Beobachtungswerte einer zweifaktoriellen Varianzanalyse . . . . .	83
4.6	Abweichungsquadratsummen und Freiheitsgrade einer Untersuchung mit 2 Faktoren . . . . .	84
4.7	Kritische Schranken beim Kruskal-Wallis-Test . . . . .	91
4.8	Teilfaktorieller Versuchsplan: 5 Faktoren, 2 Stufen, 8 Experimente	96
4.9	Fehlerwahrscheinlichkeiten bei verschiedenen Auswerteverfahren	110
4.10	Faktorzuordnung zur Abschätzung der erzielbaren Schärfe von Direktvergleichen . . . . .	114
4.11	Abschätzung der erzielbaren Schärfe mit $A_1 = A_2$ . . . . .	115
4.12	Abschätzung der erzielbaren Schärfe mit $A_1 \neq A_2$ . . . . .	115
5.1	Spannweite ausgewählter Versuchsbedingungen . . . . .	130
5.2	Messgenauigkeit der verwendeten Messtechnik . . . . .	131
5.3	Analyseergebnisse der Gelenkwelldrehzahl bei stationären Versuchen auf verschiedenen Untergründen . . . . .	136
5.4	Analyseergebnisse des Gelenkwelldrehmomentes bei stationären Versuchen auf verschiedenen Untergründen . . . . .	137
5.5	Analyseergebnisse der Antriebsleistung über die Gelenkwelle bei stationären Versuchen auf verschiedenen Untergründen . . . . .	139
5.6	Analyseergebnisse der Gelenkwelldrehzahl bei dynamischen Versuchen auf verschiedenen Untergründen . . . . .	146
5.7	Analyseergebnisse des Gelenkwelldrehmomentes bei dynamischen Versuchen auf verschiedenen Untergründen . . . . .	148

5.8 Analyseergebnisse der Antriebsleistung über die Gelenkwelle  
bei dynamischen Versuchen auf verschiedenen Untergründen . . 149

A.1 Charakteristische Daten Versuchsfahrzeug Claas Scorpion 7040 165

A.2 Rollwiderstandsbeiwerte für mobile Arbeitsmaschinen . . . . . 167

A.3 Basisdatensatz für den Vergleich der Auswerteverfahren . . . . 168

A.4 Informationen zur Entstehung der Prüfprozesse der Anwen-  
dungsbeispiele . . . . . 169

A.5 Informationen zur Durchführung der Anwendungsbeispiele . . . 169

# Literaturverzeichnis

- [A.I10] A.I.A.G - CHRYSLER CORP., FORD MOTOR CO., GENERAL MOTORS CORP. (Hrsg.): *Measurement System Analysis*. 4. Auflage. Mich., USA, 2010. – Reference Manual
- [ASA99] ASAE - AMERICAN SOCIETY OF AGRICULTURAL ENGINEERS (Hrsg.): *ASAE D 497.4 Agricultural Machinery Management Data*. St. Joseph, Mich., USA, 1999
- [Bar96] BARRELMAYER, Thomas: *Untersuchung der Kräfte an gelenkten und angetriebenen Ackerschlepperrädern bei Gelände- und Straßenfahrt*, Universität Hohenheim, Dissertation, 1996
- [Ber80] BERGMANN, Hansjürgen: *Grabwiderstand beim Hydraulikbagger*, Technische Universität Berlin, Dissertation, 1980
- [BH72] BERNACKI, Henryk ; HAMAN, Janusz: *Grundlagen der Bodenbearbeitung und Pflugbau*. Reihe Landmaschinentechnik, Theorie und Konstruktion der Landmaschinen. Berlin : Verlag Technik, VEB, 1972
- [Bom17] BOMAG GMBH (Hrsg.): *Tandemwalzen knickgelenkt*. Boppard, 2017 [https://bsa.bomag.com/bsa/boo/pdf\\_files/PRD103017\\_1703.pdf#toolbar=1&statusbar=1&messages=0&pagemode=thumbs&view=Fit](https://bsa.bomag.com/bsa/boo/pdf_files/PRD103017_1703.pdf#toolbar=1&statusbar=1&messages=0&pagemode=thumbs&view=Fit). – Produktbroschüre
- [Bou13] BOURIER, Günther: *Beschreibende Statistik : Praxisorientierte Einführung Mit Aufgaben und Lösungen*. 11. Auflage. Wiesbaden : Springer Gabler, 2013
- [Bra85] BRAY, James H. ; MAXWELL, Scott E. (Hrsg.): *Sage University Paper Series on Quantitative Research Methods*. Bd. 54: *Multivariate analysis of variance*. Newbury Park, Calif., USA : Sage Publications, 1985

- [BS10] BORTZ, Jürgen ; SCHUSTER, Christof: *Statistik für Human- und Sozialwissenschaftler*. Berlin : Springer, 2010
- [Bur12] BURKSCHAT, Marco ; CRAMER, Erhard (Hrsg.) ; KAMPS, Udo (Hrsg.): *Beschreibende Statistik : Grundlegende Methoden der Datenanalyse*. 2. Auflage. Berlin, Heidelberg : Springer Spektrum, 2012
- [BV08] BANKHOFER, Udo ; VOGEL, Jürgen: *Datenanalyse und Statistik : Eine Einführung für Ökonomen im Bachelor*. Wiesbaden : Springer Gabler, 2008
- [CLA07] CLAAS KGAA MBH (Hrsg.): *Reparaturhandbuch Scorpion*. Harsewinkel, 2007
- [Dei09] DEITERS, Henning: *Standardisierung von Lastzyklen zur Beurteilung der Effizienz mobiler Arbeitsmaschinen*, Technische Universität Braunschweig, Dissertation, 2009
- [Der09] DERENBACH, Bernd: Staplercheck 03: Heli CPCD30. In: *Staplerworld* (2009), Nr. 01, S. 14–18
- [DF03] DEGREL, Ottmar ; FEUERSTEIN, Thomas: *DLG-PowerMix - Ein praxisorientierter Traktorentest / DLG Prüfstelle für Landtechnik*. Groß Umstadt, 2003. – Forschungsbericht
- [Die14] DIETRICH, Edgar ; SCHULZE, Alfred (Hrsg.): *Prüfprozesseignung: Prüfmittelfähigkeit und Messunsicherheit im aktuellen Normenumfeld*. 4. Auflage. München : Hanser Verlag, 2014
- [DIN95] DIN - DEUTSCHES INSTITUT FÜR NORMUNG (Hrsg.): *DIN ISO 9245:1995:-01 Erdbaumaschinen: Leistung der Maschinen - Begriffe, Formelzeichen und Einheiten*. Berlin, 1995
- [DIN99] DIN - DEUTSCHES INSTITUT FÜR NORMUNG (Hrsg.): *DIN V ENV 13005:1999: Leitfaden zur Angabe der Unsicherheit beim Messen*. Berlin : Beuth Verlag, 1999. – identisch mit: ISO/IEC Guide 98-3:2008: Guide to the expression of uncertainty in measurement (GUM). Berlin : Beuth Verlag, 2008.

- [DLG17a] DLG - DEUTSCHE LANDWIRTSCHAFTS-GESELLSCHAFT (Hrsg.): *Datenblatt DLG-PowerMix John Deere 5125R PowerShift*. Groß-Umstadt, 2017 [http://www.dlg-test.de/tests/John-Deere\\_5125R\\_PM-de.pdf](http://www.dlg-test.de/tests/John-Deere_5125R_PM-de.pdf)
- [DLG17b] DLG - DEUTSCHE LANDWIRTSCHAFTS-GESELLSCHAFT (Hrsg.): *Neuer DLG-PowerMix-Prüfstand offiziell übergeben*. Groß-Umstadt, 2017 <http://presse.dlg.org/blob/24601>. – Pressemitteilung
- [FD05] FEUERSTEIN, Thomas ; DEGREL, Ottmar: Praxis aus der Konserve: Der DLG-PowerMix. In: *test & technik* (2005), Nr. 2, S. 12–13
- [Fer09] FERHADBEGOVIC, Bojan: *Entwicklung und Applikation eines instationären Reifenmodells zur Fahrdynamiksimulation von Ackerschleppern*, Universität Hohenheim, Dissertation, 2009
- [FGG95] FREY, Michael ; GNADLER, Rolf ; GÜNTER, Frank: Untersuchung der Verlustleistung an Pkw-Reifen. In: *VDI BERICHTE NR. 1224*. Karlsruhe, 1995, S. 101–128
- [FKPT11] FAHRMEIR, Ludwig ; KÜNSTLER, Rita ; PIGEOT, Iris ; TUTZ, Gerhard: *Statistik : Der Weg zur Datenanalyse*. 7. Auflage. Berlin : Springer, 2011
- [FSA12] FRANK, Bobbie ; SKOGH, Lennart ; ALAKÜLA, Mats: On wheel loader fuel efficiency difference due to operator behaviour distribution. / Lund University. 2012. – Forschungsbericht
- [Fuc11] FUCHS, Andreas: Neue Traktoren mit Vierzylindermotoren. In: *ATZ offhighway* Sonderausgabe November (2011), S. 6–9
- [Gir10] GIRMSCHIED, Gerhard: *Leistungsermittlungshandbuch für Baumaschinen und Bauprozesse*. Springer Berlin Heidelberg, 2010
- [Gün94] GÜNTER, Frank: *Experimentelle Untersuchung der Verlustleistung von Pkw-Reifen*, Universität Karlsruhe, Dissertation, 1994
- [GP14] GEIMER, Marcus ; POHLANDT, Christian ; INSTITUT FÜR FAHRZEUGSYSTEMTECHNIK (Hrsg.): *Grundlagen mobiler Arbeitsmaschinen*. Karlsruhe : KIT Scientific Publishing, 2014

- [GS94] GEHBAUER, Fritz ; STEFAN, Beretitsch: Zum Kräftespiel im System Schneidwerkzeug-Boden. In: *BMT 1* (1994), Februar, Nr. 41, S. 35–37
- [Ham10] HAMM AG (Hrsg.): *Hammtronic*. Tirschenreuth, 2010 [https://media.wirtgen-group.com/media/04\\_hamm/products/documents/tech\\_hammtronic\\_de.pdf](https://media.wirtgen-group.com/media/04_hamm/products/documents/tech_hammtronic_de.pdf). – Broschüre
- [Har02] HARNISCH, Carsten: *Dynamische Echtzeitsimulation der Geländefahrt mehrachsiger Radfahrzeuge*, Universität der Bundeswehr Hamburg, Dissertation, 2002
- [Hed16] HEDDERICH, Jürgen ; SACHS, Lothar (Hrsg.): *Angewandte Statistik : Methodensammlung mit R*. 15. Auflage. Berlin, Heidelberg : Springer Spektrum, 2016
- [Heu75] HEUMANN, Curt: *Dynamische Einflüsse bei der Schnittkraftbestimmung in standfesten Böden*, Universität Karlsruhe, Dissertation, 1975
- [Hol98] HOLLÄNDER, Claus: *Untersuchungen zur Beurteilung und Optimierung von Baggerhydrauliksystemen*, Technische Universität Braunschweig, Dissertation, 1998
- [Hol01] HOLST, Christian v.: *Vergleich von Reifenmodellen zur Simulation der Fahrdynamik von Traktoren*, Technische Universität Berlin, Dissertation, 2001
- [Hub10] HUBER, Andreas: *Ermittlung von prozessabhängigen Lastkollektiven eines hydrostatischen Fahrantriebsstrangs am Beispiel eines Teleskopladers*, Karlsruher Institut für Technologie, Dissertation, 2010
- [Hub14] HUBER, Frank ; MEYER, Frederik (Hrsg.) ; LENZEN, Johann M. (Hrsg.): *Grundlagen der Varianzanalyse : Konzeption - Durchführung - Auswertung*. Wiesbaden : Springer Gabler, 2014
- [ISO12] ISO - INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (Hrsg.): *ISO 22514-7:2012-09: Statistische Verfahren im Prozessmanagement - Fähigkeit und Leistung - Teil 7: Fähigkeit von Messprozessen*. Berlin : Beuth Verlag, 2012



- [JCM07a] JCMA - JAPAN CONSTRUCTION MECHANISATION ASSOCIATION (Hrsg.): *JCMAS H 020:2007: Earth-moving machinery - Fuel consumption of the hydraulic excavator - Test procedures*. Japan, 2007
- [JCM07b] JCMA - JAPAN CONSTRUCTION MECHANISATION ASSOCIATION (Hrsg.): *JCMAS H 021:2007: Earth-moving machinery - Fuel consumption of the tractor-dozzer - Test procedures*. Japan, 2007
- [JCM07c] JCMA - JAPAN CONSTRUCTION MECHANISATION ASSOCIATION (Hrsg.): *JCMAS H 022:2007: Earth-moving machinery - Fuel consumption of the wheel loader - Test procedures*. Japan, 2007
- [Ket81] KETTERER, Bertold: *Modelluntersuchungen zur Prognose von Schneid- und Planierkräften im Erdbau*, Universität Karlsruhe, Dissertation, 1981
- [Küh56] KÜHN, Günter: *Der gleislose Erdbau : Anwendung und Einsatz der Geräte, Organisation und Kalkulation des Förderbetriebes*. Berlin : Springer, 1956
- [KMV11] KUNZE, Günter ; MIETH, Sebastian. ; VOIGTH, Sebastian: Bedienereinfluss auf Leistungszyklen Mobiler Arbeitsmaschinen. In: *ATZ offhighway Sonderausgabe* (2011), April, S. 70–78
- [KN06] KREISS, Jens-Peter ; NEUHAUS, Georg: *Einführung in die Zeitreihenanalyse*. Berlin : Springer, 2006
- [Kno75] KNOBLICH, Klaus: Über den Böschungswinkel von Schutthalden. In: *Catena* 2 (1975), S. 1–10
- [Kra14] KRANKE, Andre: Über hundert Jahre Stapler-Wissen. In: *VerkehrsRUNDSCHAU* (2014), Nr. 20
- [Kro14] KRONTHALER, Franz: *Statistik angewandt : Datenanalyse ist (k)eine Kunst*. Berlin, Heidelberg : Springer Spektrum, 2014
- [KW83] KÜHN, Günter ; WARDECKI, Nobert: Anpassung der Ladeschaufel von Hydraulikbaggern an das Felshaufwerk. In: *BMT* 10 (1983), S. 464–470

- [Lei09] LEISTER, Günter: *Fahrzeugreifen und Fahrwerkentwicklung : Strategie, Methoden, Tools*. Wiesbaden : Vieweg + Teubner, 2009
- [Lie17] LIEBHERR-WERK BISCHOFSHOFEN GMBH (Hrsg.): *Radlader L 526 - L 546*. 2017 <https://www.liebherr.com/external/products/products-assets/283872/Bildprospekt%20L%20526%20-%20L%20546.pdf>. – Produktproschüre
- [Lin09] LINDE MATERIAL HANDLING GMBH (Hrsg.): *Mehr Produktivität, geringere Kosten*. Aschaffenburg, 2009 [http://www.linde-mh.de/de/countrysite/news\\_infoservice\\_1/newspressedetails\\_705.html](http://www.linde-mh.de/de/countrysite/news_infoservice_1/newspressedetails_705.html). – Pressemitteilung
- [Lin10] LINDE MATERIAL HANDLING GMBH (Hrsg.): *Das Ganze sehen. Die Leistung macht den Unterschied*. Aschaffenburg, 2010 [http://www.linde-mh.de/de/countrysite/news\\_info\\_service\\_1/newspressedetails\\_705.html](http://www.linde-mh.de/de/countrysite/news_info_service_1/newspressedetails_705.html). – Broschüre
- [Lin13] LINDE MATERIAL HANDLING GMBH (Hrsg.): *Linde-Schwerstapler führen das Feld an*. Aschaffenburg, 2013 [http://www.linde-mh.de/de/countrysite/news\\_infoservice\\_1/newspressedetails\\_6272.html](http://www.linde-mh.de/de/countrysite/news_infoservice_1/newspressedetails_6272.html). – Pressemitteilung
- [MAH12] MAHA-AIP GMBH & Co. KG (Hrsg.): *Rollenprüfstand AIP-CDM 78.7H-4mot NVH*. Haldenwang, 2012. – Betriebsanleitung
- [McK85] MCKYES, E.: *Soil Cutting and Tillage*. Amsterdam, New York : Elsevier, 1985
- [ME14] MURRENHOF, Hubertus ; ECKSTEIN, Lutz: *Fluidtechnik für mobile Anwendungen*. 6. Auflage. Aachen : Shaker, 2014
- [Mel92] MELCHINGER, Uli: *Simulation der Arbeitsbewegungen und Antriebssysteme von Hydraulikbaggern*, Technische Universität Berlin, Dissertation, 1992
- [Mer16] MERLO DEUTSCHLAND GMBH (Hrsg.): *Schwerlast-Turbofarmer 50.8 + 45.11*. Bremen, 2016 <https://www.merlo.com>

- //www.merlo.de/typo3temp/tx\_products/1281/TF50\_8\_T\_CS\_156\_CVTRONIC\_0\_89342600\_1461236035.pdf. – Produktbroschüre
- [Mur12] MURRENHOF, Hubertus: *Grundlagen der Fluidtechnik*. Bd. 1: Hydraulik. 7. Auflage. Aachen : Shaker, 2012
- [MW14] MITSCHKE, Manfred ; WALLENTOWITZ, Henning: *Dynamik der Kraftfahrzeuge*. 5. Auflage. Wiesbaden : Springer Vieweg, 2014
- [NBL07] NAUNHEIMER, Harald ; BERTSCHE, Bernd ; LECHNER, Gisbert: *Fahrzeuggetriebe : Grundlagen, Auswahl, Auslegung und Konstruktion*. 2. Auflage. Springer Berlin Heidelberg, 2007
- [Neu03] NEUNABER, Manfred: Neue Messverfahren im DLG-profi-Test: Zugleistung , Zapfwelle und Hydraulik: PowerMix. In: *Profi Test* (2003), S. 24-25
- [New13] NEW HOLLAND KOBELCO CONSTRUCTION MACHINERY S.P.A (Hrsg.): *E75C SR / E85C MSR*. Florenz, Italien, 2013 <http://pdf.directindustry.com/pdf/new-holland/e75csr-e85cmsr/20169-382913.html>. – Produktbroschüre
- [NR14] NAGEL, Philip ; ROOS, Lennart: Antriebsstrang mit Energierückgewinnung: Entwicklungsmethodik und Betriebsstrategien für mobile Arbeitsmaschinen / Karlsruher Institut für Technologie, Technische Universität Braunschweig. 2014 (IGF-Vorhaben 16770). – Schlussbericht
- [Par02] PARK, Borinara: *Development of a Virtual Reality Excavator Simulator - a Mathematical Model of Excavator Digging and a Calculation Methodology*, Virginia Polytechnic Institute and State University, Dissertation, 2002
- [PD75] PERREAULT, William D. ; DARDEN, William R.: Unequal cell sizes in marketing experiments : use of the general linear hypothesis. In: *Journal of marketing research : JMR / American Marketing Association*. - Chicago, Ill 3 (1975), Nr. 12, S. 333-342

- [PK96] PLESSER, Joachim ; KUTZBACH, Heinz D.: Spring and Damping Parameters of Agricultural Tyres for the Vertical and Longitudinal Direction. In: *Proceedings of the 12th International Conference of the ISTVS*. Peking, October 1996, S. 233–240
- [Ple97] PLESSER, Joachim: *Dynamisches Verhalten von Ackerschlepperreifen in Vertikal- und Längsrichtung auf fester Fahrbahn*, Universität Hohenheim, Dissertation, 1997
- [PPM12] PPM GMBH (Hrsg.): *20xx GNSS Sens or Serie*. 2012 [http://www.ppmgmbh.de/ppm\\_design/ppm-design-pdf/20xx.pdf](http://www.ppmgmbh.de/ppm_design/ppm-design-pdf/20xx.pdf). – Datenblatt
- [Rei10] REICH, Tristan: *Optimierungspotenzial und Vergleichszyklus für den Kraftstoffverbrauch von Forstspezialschleppern*, Karlsruher Institut für Technologie, Diplomarbeit, 2010
- [Rei14] REICH, Tristan: Vergleichszyklen für Mobile Arbeitsmaschinen auf dem Rollenprüfstand. In: *43. VDBUM Seminar Kassel*, 2014, S. 36–40
- [RFHN06] RASCH, Björn (Hrsg.) ; FRIESE, Malte (Hrsg.) ; HOFMANN, Wilhelm (Hrsg.) ; NAUMANN, Ewald (Hrsg.): *Quantitative Methoden*. Bd. 1. 2. Auflage. Heidelberg : Springer Medizin Verl., 2006
- [Sch06] SCHREIBER, Matthias: *Kraftstoffverbrauch beim Einsatz von Ackerschleppern im besonderen Hinblick auf CO<sub>2</sub>-Emissionen*, Universität Hohenheim, Dissertation, 2006
- [Sch09] SCHUMACHER, Andreas: *Ein generisches Reifenmodell für Fahrdynamiksimulationen von Offroad-Fahrzeugen*, Universität Braunschweig, Dissertation, 2009
- [Sch13] SCHÜTZ, Thomas: *Hucho - Aerodynamik des Automobils : Strömungsmechanik, Wärmetechnik, Fahrdynamik, Komfort*. 6. Auflage. Wiesbaden : Springer Vieweg, 2013
- [Sch17a] SCHERER, Philipp: *Methodik zur Bewertung der Energieeffizienz von mobilen Arbeitsmaschinen*, Karlsruher Instituts für Technologie, Dissertation, 2017

- [Sch17b] SCHERER, Philipp: Themenschwerpunkt I: Ermittlung der Energieeffizienz mobiler Arbeitsmaschinen. In: *TEAM - Entwicklung von Technologien für energiesparende Antriebe mobiler Arbeitsmaschinen* (2017), S. 14–69. – Abschlussbericht
- [SG13] SCHERER, Philipp ; GEIMER, Marcus: Verbundprojekt TEAM - Ermittlung der Energieeffizienz. In: *Mobile Maschinen* (2013), Juni, S. 10–13
- [Sie10] SIEBERTZ, Karl ; BEBBER, David v. (Hrsg.) ; HOCHKIRCHEN, Thomas (Hrsg.): *Statistische Versuchsplanung : Design of Experiments (DoE)*. Springer Berlin Heidelberg, 2010
- [SP13] SMITH, William ; PENG, Huei: Modeling of wheel-soil interaction over rough terrain using discrete element method. In: *Journal of Terramechanics* 50 (2013), S. 277–287
- [Sto18] STOCKER, Herbert: *Angewandte Ökonometrie: Methoden der empirischen Wirtschaftsforschung*, Fakultät für Volkswirtschaft und Statistik, Universität Innsbruck, Vorlesungsskript, 2018
- [Stu15] STURM, Christoph: *Bewertung der Energieeffizienz von Antriebssystemen mobiler Arbeitsmaschinen am Beispiel Bagger*, Karlsruher Institut für Technologie, Dissertation, 2015
- [Tes17] TESTO INDUSTRIAL SERVICES (Hrsg.): *Messunsicherheitsfibel*. 2017 [http://download.testotis.de/files/DE/FibelIn/Messunsicherheitsbestimmung\\_nach\\_GUM.pdf](http://download.testotis.de/files/DE/FibelIn/Messunsicherheitsbestimmung_nach_GUM.pdf)
- [VDA11] VDA - VERBAND DER AUTOMOBILINDUSTRIE (Hrsg.): *VDA Band 5: Prüfprozesseignung*. 2. Auflage. Frankfurt, 2011
- [VDI02] VDI-GESELLSCHAFT FÖRDERTECHNIK MATERIALFLUSS LOGISTIK (Hrsg.): *VDI 2198: Typenblätter für Flurförderzeuge*. 2002
- [Vol13] VOLVO CONSTRUCTION EQUIPMENT GERMANY GMBH: Präzision und Leistung - der Volvo-Radfertiger P6870C ABG. In: *Maschinen&Technik* (2013), April

- [Vol14] VOLZ, Frank: New tractor test station. In: *agrifuture* Autumn + Winter (2014), S. 26–27
- [Vos09] VOSS, Andreas: *Statistik II*, Institut für Psychologie; Albert-Ludwigs-Universität Freiburg, Vorlesungsskript, 2009
- [Wal07] WALLENTOWITZ, Henning: *Längsdynamik von Kraftfahrzeugen*. 13. Auflage. Aachen : IKA - Inst. für Kraftfahrwesen, 2007
- [Wei13] WEISSENBÖCK, Michael: *Radlader-Y-Zyklus auf dem Allradrollenprüfstand*, Karlsruher Institut für Technologie, Bachelorthesis, 2013
- [Wüs71] WÜSTHOF, Peter: Das Verhalten der Antriebsräder und ihre Bedeutung für die hydrostatische Übertragung im Fahrtrieb. In: *Industrie-Anzeiger* 80/81 (1971), S. 2020–2024
- [WSK<sup>+</sup>15] WEBER, Jürgen ; SCHERER, Philipp ; KUNZE, Günter ; GÜNTHER, Marco ; JACOBS, Georg: *Der Effizienz auf der Spur - Entwicklung von Technologien für energiesparende Antriebe mobiler Arbeitsmaschinen*. Dresden, 2015. – Infobroschüre