

Doris Oriwol

**Methodologische Aspekte
biomechanischer Messungen unter
Laborbedingungen**

Eine kritische Betrachtung des gängigen Messprotokolls des
Ausdauerlaufens



**TECHNISCHE UNIVERSITÄT
CHEMNITZ**

**Universitätsverlag Chemnitz
2012**

Impressum

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Angaben sind im Internet über <http://dnb.d-nb.de> abrufbar.

Zugl.: Chemnitz, Techn. Univ., Diss., 2012

Technische Universität Chemnitz/Universitätsbibliothek
Universitätsverlag Chemnitz
09107 Chemnitz
<http://www.bibliothek.tu-chemnitz.de/UniVerlag/>

Herstellung und Auslieferung
Verlagshaus Monsenstein und Vannerdat OHG
Am Hawerkamp 31
48155 Münster
<http://www.mv-verlag.de>

ISBN 978-3-941003-56-9
<http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa-84452>

METHODOLOGISCHE ASPEKTE BIOMECHANISCHER MESSUNGEN UNTER LABORBEDINGUNGEN

Eine kritische Betrachtung des gängigen Messprotokolls
des Ausdauerlaufens

D i s s e r t a t i o n

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der Fakultät für Human- und Sozialwissenschaften der
Technischen Universität Chemnitz

von Doris Oriwol, geboren am 21. Januar 1982 in Meerane

Chemnitz, den 02. Dezember 2011

Datum des Promotionskolloquiums: 01. Februar 2012
Vorsitzender der Promotionskommission: Prof. Dr. Johannes Kopp

Erstgutachter: Prof. Dr. Thomas L. Milani
Zweitgutachter: JP. Dr. Christian Maiwald
Stimmberechtigtes Mitglied: Prof. Dr. Peter Sedlmeier

Für Karl-Heinz und Falk Oriwol

Inhaltsverzeichnis

Abbildungsverzeichnis	xiii
Tabellenverzeichnis	xvii
Abkürzungsverzeichnis	xix
1 Einführung	1
1.1 Status Quo wissenschaftlicher Experimente im Bereich der Laufschuhforschung . . .	4
1.2 Variabilität - Fluch und Segen	7
1.3 Berücksichtigung der Versuchswiederholung zur Charakterisierung der Variabilität	10
1.4 Der Mittelwert als charakteristische Kennzahl eines Probanden	14
1.5 Fragestellungen	17
2 Methodische Aspekte	19
2.1 Probandengut	19
2.2 Datenaufnahme	20
2.3 Datenverarbeitung	22
2.3.1 Bodenreaktionskräfte	22
2.3.2 Winkelgeschwindigkeit	24
2.4 Statistische Auswertungen und Berechnungen (Hypothese I & II)	27
2.4.1 Kumulation	28
2.4.2 Konvergenz	29

2.4.3	Bootstrap-Verfahren zur Bildung von Prädiktionsbändern	31
2.4.4	Maßzahlen für die Reliabilität innerhalb und zwischen den Messtagen	33
2.4.5	Auswertung der vier Gruppierungen	36
2.5	Statistische Auswertungen und Berechnungen (Hypothese III)	36
2.5.1	Nichtlineare Regressionsmodelle für die Anpassung der Fläche der Prädik- tionsbänder	36
2.5.2	Standardabweichungskurven	38
3	Resultate zu Hypothese I: Konvergenz des Mittelwertes	41
3.1	Betrachtung diskreter Parameter	41
3.2	Betrachtung der Gesamtkurve	49
3.3	Diskussion der Ergebnisse	54
4	Resultate zu Hypothese II: Charakterisierung der Probandenvariabilität	59
4.1	Betrachtung diskreter Parameter	59
4.2	Betrachtung der Gesamtkurve	63
4.3	Diskussion der Ergebnisse	67
5	Resultate zu Hypothese III: Algorithmen zur Bestimmung einer minimalen Anzahl	71
5.1	Minimale Anzahl unter Verwendung von Prädiktionsbändern	71
5.2	Minimale Anzahl unter Verwendung von Standardabweichungskurven	75
5.3	Diskussion der Ergebnisse	82
6	Diskussion und Ausblick	87
	Literaturverzeichnis	101
A	Ergänzende Informationen zu bewegungswissenschaftlichen Studien	109
B	Ergänzende Informationen zu den Resultaten von Hypothese I	113

C Ergänzende Informationen zu den Resultaten von Hypothese II	121
D Ergänzende Informationen zu den Resultaten von Hypothese III	125
E Ergänzende Informationen zu den Inhalten der Diskussion	129
Wissenschaftlicher Werdegang	133
Publikationen	134
Danksagung	135

Abbildungsverzeichnis

1.4.1	Galton Brett zur Veranschaulichung des zentralen Grenzwertsatzes	16
1.4.2	Verteilung der Werte der maximalen Kraftanstiegsrate der vertikalen Bodenreaktionskraft	16
2.2.1	Labor mit Kraftmessplatte Studie I und II	21
2.2.2	Messgeräte Studie II	21
2.3.1	Beispiel einer Kurve der vertikalen Bodenreaktionskraft F_z	24
2.3.2	100 Signale der vertikalen Bodenreaktionskraft F_z	24
2.3.3	Beispiel einer Kurve der Bodenreaktionskraft in anterior-posterior Richtung F_y	25
2.3.4	100 Signale der Bodenreaktionskraft in anterior-posterior Richtung F_y	25
2.3.5	Beispiel einer Kurve der Bodenreaktionskraft in mediolateraler Richtung F_x	25
2.3.6	100 Signale der Bodenreaktionskraft in mediolateraler Richtung F_x	25
2.3.7	Schrittseparation des Gyrometers bei Laboraufzeichnungen	26
2.3.8	Beispiel einer Kurve der Winkelgeschwindigkeit ω	27
2.3.9	100 Signale der Winkelgeschwindigkeit ω	27
2.3.10	Beispiel einer Kurve der integrierten Winkelgeschwindigkeit φ	27
2.3.11	100 Signale der integrierten Winkelgeschwindigkeit φ	27
2.4.1	Beispiel kumulierter Mittelwertkurven von F_z (drei Datenpunkte)	30
2.4.2	Entsprechende z-transformierte Mittelwertkurven (drei Datenpunkte)	30
2.4.3	Beispiel kumulierter Mittelwertkurven von F_z (100 Datenpunkte)	31
2.4.4	Entsprechende z-transformierte Mittelwertkurven (100 Datenpunkte)	31
2.4.5	Prädiktionsband von Abbildung 2.3.2 mit allen Versuchen	33
2.4.6	Prädiktionsband von Abbildung 2.3.2	33
3.1.1	FR_{z_I} : konvergierender Mittelwert (Stabilisierung ab 70 Werten)	44
3.1.2	FR_{z_I} : konvergierender Mittelwert (Stabilisierung ab 55 Werten)	44
3.1.3	F_{z_I} : divergierender Mittelwert (1)	45

3.1.4	F_{z_I} : divergierender Mittelwert (2)	45
3.1.5	F_{z_A} : z-transformierte Folgen kumulierter Mittelwerte (Messtag 1/2 und Mittelwertkriterium)	45
3.1.6	Typisierung kumulierter Mittelwerte	48
3.2.1	F_z : Differenzenfläche DA_{intra}^{kr} (prozentuale Abweichung 7%)	49
3.2.2	F_x : Differenzenfläche DA_{intra}^{kr} (prozentuale Abweichung 131%)	49
3.2.3	ω : Differenzenfläche DA_{inter}^{kr} (prozentuale Abweichung 64%)	51
3.2.4	ω : Differenzenkurven innerhalb eines Tages für Abbildung 3.2.3	51
3.2.5	F_z : konvergierende punktweise kumulierte Mittelwertkurven (Stabilisierung ab 60 Werten)	52
3.2.6	F_z : divergierende punktweise kumulierte Mittelwertkurven	52
3.2.7	F_z : kumulierte Mittelwertkurven mit Mittelwertkriterium	53
3.2.8	F_z : Ausschnitt aus Abbildung 3.2.7	53
4.1.1	ω_{max} : konvergierende kumulierte Standardabweichungen (Stabilisierung ab 50 Werten)	60
4.1.2	ω_{max} : divergierende kumulierte Standardabweichungen	60
4.1.3	ω_{max} : kumulierte RMSE-Folge	62
4.2.1	F_z : punktweise konvergierende kumulierte SD-Kurven (Stabilisierung ab 50 Werten)	63
4.2.2	F_z : punktweise divergierende kumulierte SD-Kurven	63
4.2.3	ω : teilweise punktweise konvergierende kumulierte SD-Folge (Stabilisierung zwischen 35 und 90 Werten)	64
4.2.4	F_z : kumulierte Standardabweichungskurven	65
4.2.5	ω : kumulierte Standardabweichungskurven	65
4.2.6	F_z : Prädiktionsband (inklusive involvierter Kurven)	65
4.2.7	ω : Prädiktionsband (inklusive involvierter Kurven)	65
4.2.8	F_z : Prädiktionsband für 5 und 100 Läufe	66
4.2.9	ω : Prädiktionsband für 5 und 100 Läufe	66
4.2.10	F_z : Prädiktionsband Messtag 1/2 (100 Läufe)	67
4.2.11	ω : Prädiktionsband Messtag 1/2 (100 Läufe)	67
5.1.1	F_z : Anpassung von APB_N^{kr} mittels nichtlinearer Regression (1)	73
5.1.2	F_z : Anpassung von APB_N^{kr} mittels nichtlinearer Regression (2)	73
5.1.3	F_z : Kumulierte Folgen von $APBZ_N^{kr}$ ab 15 Läufen	74
5.1.4	F_z : Differenzenfolge $DAPB_N^{kr}$ ab 15 Läufen für Abbildung 5.1.2	74

5.1.5	Peak 1: Auswirkungen auf das Prädiktionsband aus Abbildung 5.1.4	74
5.1.6	Peak 2: Auswirkungen auf das Prädiktionsband aus Abbildung 5.1.4	75
5.2.1	F_z : Differenz der Standardabweichungskurven	77
5.2.2	F_z : maximale Differenzenkurven aller Probanden	77
5.2.3	F_z : absolute Differenzenkurven mit Mittelwert	77
5.2.4	F_z : Prädiktionsband der mittleren maximalen Differenzenkurven	77
5.2.5	F_z : Differenz der Standardabweichungskurven mit $PDSD_{max}^k$ und $PDSD_{mw}^k$	78
5.2.6	F_z : punktweise kumulierte, transformierte SD-Kurven aus Abbildung 5.2.5 mit minimaler Anzahl	78
5.2.7	Populationsdifferenzenkurven: $PDSD_{mw I}^k$, $PDSD_{mw II}^k$, $PDSD_{mw III}^k$	79
6.1	F_{z_t} : Einzelwerte mit zeitlicher Dimension	93
6.2	Folge kumulierter Mittelwerte zu Abbildung 6.1	93
6.3	Auswirkung auf den Mikrokosmos der biomechanischen Wissenschaftsgesellschaft . .	97
6.4	Auswirkung auf den Makrokosmos der biomechanischen Wissenschaftsgesellschaft . .	98

Tabellenverzeichnis

1.1.1	Zielsetzungen und Einflussfaktoren ausgewählter biomechanischer Studien im Bereich der Laufschuhforschung	5
1.3.1	Zusammenfassung Methodik der Studien aus Tabelle 1.1.1	11
2.1.1	Probandengut von Studie I und II	19
3.1.1	Differenz D_{intra}^{pr} diskreter Parameter	42
3.1.2	Differenz D_{inter}^{pr} diskreter Parameter	42
3.1.3	Zusammenfassung Konvergenz des Mittelwertes (diskrete Parameter)	43
3.1.4	Zusammenfassung Mittelwertkriterium (diskrete Parameter)	46
3.1.5	Zusammenfassung Abhängigkeit Mittelwertkriterium (diskrete Parameter)	46
3.1.6	Typisierung kumulierter Mittelwertkurven (prozentualer Anteil)	48
3.1.7	Anzahl signifikant unterschiedlicher Konfidenzintervalle diskreter Parameter	49
3.2.1	Differenz DA_{intra}^{kr} gesamter Kurven	50
3.2.2	Differenz DA_{inter}^{kr} gesamter Kurven	50
3.2.3	Zusammenfassung Konvergenz des Mittelwertes (gesamte Kurven)	52
3.2.4	Zusammenfassung Mittelwertkriterium (gesamte Kurve)	53
3.2.5	Zusammenfassung Abhängigkeit Mittelwertkriterium (gesamte Kurve)	54
4.1.1	Zusammenfassung Konvergenz der Standardabweichungen (diskrete Parameter)	60
4.1.2	FR_{z_I} und ω_{max} : kumulierte RMSE	62
4.1.3	Differenz DE_{intra}^p und DE_{inter}^p des RMSE	62
4.2.1	Konvergenz der Standardabweichungen (gesamte Kurve)	63
4.2.2	Prozentuale Abweichung der Fläche des Prädiktionsbandes bezüglich der mittleren Fläche	66
5.1.1	Güte der Anpassung der nichtlinearen Regression	72
5.1.2	Minimal zu messende Anzahl (Krümmungspunkt von APB_N^{kr})	72

5.1.3	Minimal zu messende Anzahl (anhand $DAPB_N^{kr}$)	76
5.2.1	Minimal zu messende Anzahl (anhand Populationsdifferenzenkurve)	79
5.2.2	Kreuzvalidierung der minimal zu messenden Anzahl (Populationsdifferenzenkurve)	80
5.2.3	Differenz minimal zu messende Anzahl ($\#_{PDSD} - \#_{DAPB}$)	81
5.2.4	Differenz von APB_N^{kr} und APB_{100}^{kr} für minimal zu messende Anzahl $\#_{PDSD}$	81
6.1	Darstellung der Ergebnisse zu Hypothese I, II und III	87
A.1	Ergänzende Informationen zu biomechanischen Studien	112
B.1	Konvergenz des Mittelwertes (diskrete Parameter)	113
B.2	Mittelwertkriterium für $q = 0.1, 0.25, 0.5$ (diskrete Parameter)	114
B.3	Abhängigkeit des Mittelwertkriteriums (diskrete Parameter)	115
B.4	Differenz D_{intra}^{pr} (konvergente diskrete Parameter)	115
B.5	Absolute Differenz von MW_{100}^{pr} und MW_5^{pr} (diskrete Parameter)	116
B.6	Konvergenz des Mittelwertes (gesamte Kurve)	117
B.7	Mittelwertkriterium für $q = 0.1, 0.25, 0.5$ (gesamte Kurve)	117
B.8	Abhängigkeit des Mittelwertkriteriums (gesamte Kurve)	118
B.9	Differenz DA_{intra}^{kr} (konvergente gesamte Kurven)	118
B.10	Differenz DA_{inter}^{kr} (konvergente gesamte Kurve)	119
C.1	Konvergenz der Standardabweichungen (diskrete Parameter)	121
C.2	Kumulierter RMSE	122
C.3	Differenz DE_{intra}^p (diskrete Parameter)	123
C.4	Kreuztabelle Konvergenz/Divergenz von SD-Kurven und MW-Kurven	123
C.5	Prozentuale Abweichung $DAPB_N^{kr}$ für 10 und 100 Läufe	124
C.6	Prozentuale Abweichung $DAPB_N^{kr}$ für 25 und 100 Läufe	124
D.1	Differenz minimale Anzahl zwischen den Messtagen ($DAPB_N^{kr}$)	125
D.2	Differenz minimale Anzahl zwischen den Messtagen (PDSD-Kurve)	126
D.3	Anzahl an Probanden mit annähernd gleicher Differenz am Schwellenwert $s = 5$; Mittlere Differenzen am Schwellenwert $s = 5$	126
D.4	Schwankung der mittleren Anzahl innerhalb eines Probanden	127
E.1	Fehlerarten	131

Abkürzungsverzeichnis

↗	siehe
#	Anzahl
Abb.	Abbildung
BK	Bodenkontakt
Bsp.	Beispiel
BuWo	Butterworth-Filter
EEG	Elektroenzephalogramm
f.	folgende
Frequ.	Frequenz
FU	fester Untergrund
Glg.	Gleichung
indiv.	individuell
LB	Laufband
MW	Mittelwert
RMSE	Root mean square error
S.	Seite
SD	Standardabweichung
u. a.	unter anderem
v	Geschwindigkeit
v. Chr.	vor Christus
vs.	versus
z. B.	zum Beispiel
z. T.	zum Teil

Kapitel 1

Einführung

Antrieb wissenschaftlicher Forschung und Erkenntnisfindung ist der uns eigene Entdeckergeist, welcher sich schon in früher Kindheit bemerkbar machte, indem wir Dinge bis auf den Grund erforschten. Man erinnere sich an die vielen ausgeräumten Küchenschränke, Mehl- oder Nudeltüten verstreut auf dem Boden oder die Klänge der ersten Trommel. All diese Dinge boten einen erheblichen Fundus an Entdeckungen. Nach Piaget liegt in diesen Aktivitäten der Beginn des wissenschaftlichen Experimentierens (Siegler, DeLoache & Eisenberg, 2008, S. 189). Vor allem im späteren Jugendalter entwickeln Kinder die Fähigkeit, tiefgründiger zu forschen - primär in Bezug auf „das Aufstellen von Hypothesen, das Durchführen von Experimenten und das Ziehen von Schlussfolgerungen“ (Siegler et al., 2008, S. 181).

Bei einer späteren Tätigkeit in Wissenschaft und Forschung ist dieser Entdeckergeist vermutlich reichlich erhalten geblieben und man begegnet u. a. in den Naturwissenschaften obigen Aktivitäten wieder. Aus erkenntnis- und wissenschaftstheoretischer Sicht können Lösungen eines Problems auf differenziertem Weg gefunden werden. Bei naturwissenschaftlichen Fragestellungen wird die Beantwortung der aufgestellten Hypothesen eher auf induktiv-empirischem Weg erfolgen, wobei ein Forscher mit mathematischem Hintergrund rigoros einen deduktiv-dogmatischen Weg einschlagen wird (Eberhard, 1999, S. 29 ff., 32 ff.)^{1,2}. Dabei sei angemerkt, dass zumeist das Forschungsgebiet und bereits etablierte Theorien den Weg der Lösung vorgeben. Die Mathematik als Beispiel einer dogmatischen Wissenschaft greift auf ein Arsenal an über viele Jahrhunderte hinweg entwickelter und aufgestellter Theorien zurück und kann daher die Forschungsfrage meist in sich logisch beweisen oder widerlegen. Wissenschaften, die einen derartigen Fundus bereits bewährter Theorien nicht aufweisen können, müssen sich den Phänomenen auf induktiv-empirischem Weg nähern (Wellhöfer, 1984, S. 6). Dies trifft zumeist auf junge Wissenschaften zu wie die Sozialwissenschaft oder Sportwissenschaft.

Im Fall des deduktiv-dogmatischen Erkenntnisweges wird zunächst die Theorie formuliert und mittels logischer bereits vorhandener Axiome und Postulate im Allgemeinen bewiesen. Der Vorteil besteht darin, dass nach erfolgreicher Beweisführung diese als gegeben angenommen werden kann und keiner weiteren Beweise bedarf. Diese Theorien müssen keinen Bezug zur Realität aufweisen

¹ Ein Mathematiker wird immer versuchen, den deduktiven Weg einzuschlagen, da die mathematischen Problemstellungen von jeher eine Satz-Beweis-Struktur aufzeigen und er während des Studiums der universellen mathematischen Wissenschaften wie Analysis und Algebra keine anderen Erkenntniswege kennenlernt. Eine empirische Überprüfung von Lehrsätzen ist für einen Mathematiker, beheimatet in der universellen Mathematik, keine Alternative. Angewandte Mathematiker stehen dem Konzept der Empirie aufgeschlossener gegenüber.

² Das Spektrum der Erkenntniswege in der Wissenschafts- und Erkenntnistheorie umfasst wesentlich mehr als Deduktion und Induktion (Eberhard, 1999; Wellhöfer, 1984). Diese beiden seien lediglich einführend als Basis mathematischer und naturwissenschaftlicher Forschung erwähnt.

(Wellhöfer, 1984, S. 6), können aber in praktisch orientierten Wissenschaften Anwendung finden. Der empirische Weg hingegen betrachtet meist Spezialfälle und versucht über eine große Stichprobenanzahl und/oder viele Studien eine Theorie aufzustellen, zu hinterfragen oder zu bestätigen.

Beide Wege haben ihre Vorteile, wobei der deduktive Weg eher der menschlichen Denkstruktur gleicht (Eberhard, 1999, S. 32). Die Empirie birgt den Vorteil, dass sie sich viel schneller auf eine sich ändernde Realität einstellen kann, was gleichzeitig den Nachteil des deduktiven Weges bedeutet. Dieser verharrt in seinem starren Korsett und kann nicht so schnell auf eine Änderung der Gegebenheiten reagieren wie die Empirie (Eberhard, 1999, S. 32, 34). Dies äußert sich u. a. darin, dass viele historische mathematische Fragestellungen aus geschichtlichen Legenden, empirischen Versuchen oder Experimenten heraus geboren wurden. Der mathematische Beweis konnte in vielen Fällen erst Jahre später erbracht werden.

Hätte Dido, die Gründerin von Karthago, 825 v. Chr. auf einen Beweis für das ihr aufgebene isoperimetrische Problem gewartet, wäre die Menschheit um eine große Geschichte beraubt worden. Der Legende nach soll sie ein Stück Land von Einheimischen erworben haben, welches von einer Kuhhaut umspannt werden konnte. Dido löste dieses Problem, indem sie die Kuhhaut in viele kleine Streifen schnitt, um so ein möglichst langes Band zu erhalten und somit ein möglichst großes Stück Land zu erwerben (Zimmermann, 2010; Dittmar, 2008; Huss, 1991). Dieses Problem ist bis heute bekannt und wird als isoperimetrische Ungleichung verstanden, bei der der Flächeninhalt A einer ebenen Fläche bei gegebenem Umfang L maximiert werden soll. Die Lösung der daraus resultierenden, zu maximierenden Gleichung $L^2 \geq 4nA$ ist der Kreis. Dies konnte erst 1875/79 von Weierstraß in einer Vorlesung über Variationsrechnung vollständig bewiesen werden (Weierstraß, 1927, S. 257 ff., 306 ff.)^{3,4}. Weniger häufig existiert der umgekehrte Fall, bei dem zunächst ein Phänomen theoretisch bewiesen und erst später in Grundzügen experimentell erforscht wurde. Ein Beispiel stellt Einsteins Relativitätstheorie dar, die gleichzeitig eine Revolution in der Physik einläutete und sie in ihren Grundfesten erschütterte (Fischer, 1996, S. 55).

Nachteil des experimentellen Weges ist der enorme Forschungsaufwand, welcher sich in großer repräsentativer Stichprobengröße, hoher Studienanzahl und enormem Datenvolumen zeigt. Weiterhin haben die Ergebnisse lediglich hypothetischen Charakter, da eine Analyse der vollständigen Grundgesamtheit (alle interessierenden Objekte) in den seltensten Fällen möglich ist. Im Falle der Verwendung repräsentativer Stichproben ist man auf die Anwendung der Wahrscheinlichkeitstheorie und induktiven Statistik angewiesen, um Aussagen zu generieren. Im ungünstigsten Fall erschließt sich aus allen einem Thema gewidmeten empirischen Studien ein Konsens, der dennoch nicht der Wirklichkeit entspricht (Eberhard, 1999, S. 34).

Auswege aus diesem Dilemma gibt es nicht viele. Entweder man wartet auf den deduktiven Beweis, man verbindet die Vorteile beider Erkenntniswege zu einem (kritischer Rationalismus von Popper)

³Die Legende der Gründung von Karthago geht auf eine Sage von Iustinus zurück. Dabei konnte bis heute nicht nachgewiesen werden, ob diese Legende der tatsächlichen Gründungsgeschichte von Karthago entspricht, noch in welches Jahr diese Gründung tatsächlich fällt. Geschichtswissenschaftler datieren diese auf ungefähr 813 oder 825 vor Christus (Huss, 1991). Deutlich wird in beiden Fällen, dass ein Warten auf den Beweis rund 2600 Jahre in Anspruch genommen hätte. Die Geschichte um Rom und Karthago wäre dann möglicherweise nie geschrieben worden.

⁴Rudolf Rothe hat die Vorlesungen von 1875 und 1879 von Karl Weierstraß im Buch „Mathematische Werke von Karl Weierstraß. Vorlesungen über Variationsrechnung“ zusammengefasst und im Jahr 1927 veröffentlicht.

oder man versucht alle möglichen Fehlerquellen auszuschließen, um die Wirklichkeit möglichst nah abzubilden und eine Fehlinterpretation zu vermeiden.

Eine Empfehlung des Wartens kann nach Kenntnis der Geschichte um Dido und die isoperimetrische Ungleichung nicht ausgesprochen werden. In einigen Fällen könnte dies insofern Ewigkeiten dauern, dass es gar keine etablierten Theorien und Beweise gibt, um das Phänomen logisch zu erklären, und dass viele unsichere Quellen existieren (u. a. der Mensch als Untersuchungsobjekt), die unmöglich alle als Randbedingungen des deduktiv zu beweisenden Satzes aufgeführt werden können.

Ein weiterer Lösungsvorschlag bedient sich der Theorien zum kritischen Rationalismus von Popper, wobei obiges Dilemma zum „Induktionsproblem“ der Erkenntnistheorie gezählt werden kann. Dieses wirft die Frage auf, ob es überhaupt möglich ist, vom Besonderen, dem Ergebnis vieler empirischer Studien, zum Allgemeinen, der wissenschaftlichen Theorie, mittels empirischer Wissenschaft zu schließen und ob dieser Schluss Berechtigung findet (Keuth, 2000, S. 27). Der kritische Rationalismus verbindet die logische deduktive Prüfung aus dem deduktiv-dogmatischen Erkenntnisweg mit der empirischen Überprüfung der Forschungsfrage aus dem induktiv-empirischen Erkenntnisweg. Dabei wird zunächst die aufgestellte Forschungsfrage einer logischen Prüfung unterzogen (u. a. Widerspruchsprüfung und logischer Vergleich bereits etablierter Methoden). Dieser folgt eine empirische Untersuchung, wobei die aufgestellten Prüfhypothesen widerlegt oder bestätigt werden (Eberhard, 1999, S. 36 ff.; Wellhöfer, 1984, S. 8). Eine empirische Widerlegung gilt dabei als empirisch-wissenschaftlich und ist insofern „eindeutiger“, da diese zum tatsächlichen Sturz der Theorie führt (Keuth, 2000, S. 54 ff.). Eine Bestätigung ist kein endgültiger Nachweis für die Gültigkeit, sondern lediglich eine vorläufige Bestätigung (Wellhöfer, 1984, S. 9).

Die mit dem Prüfprozess verbundene innere Logik der aufgestellten Theorie ist einer der Vorteile der Herangehensweise dieses Erkenntnisweges. Eine Ableitung unterschiedlicher Interpretationen des gleichen Ausgangs des experimentellen Versuchs wird hiermit erschwert (Eberhard, 1999, S. 43 ff.). Dennoch wird durch den Einsatz der empirischen Prüfung und damit der Nutzung von Wahrscheinlichkeitstheorie und schließender Statistik der Irrtum nicht vollständig ausgeschlossen (Wellhöfer, 1984, S. 10). Eine Vollerhebung ist in den meisten Fällen schlichtweg unmöglich und daher wird der Forscher die Existenz von Irrtumswahrscheinlichkeiten akzeptieren müssen. Dies kann unter Umständen zu unterschiedlichen Erklärungsansätzen derselben Theorie und somit verschiedene Forscher zu gegensätzlichen Ansichten führen. Mit dieser Herangehensweise erhöht sich zwar die Stringenz des Forschungsvorhabens und sollte daher in die eigene Forschungstätigkeit integriert werden. Es bewahrt allerdings nicht vor fehlgeleiteten Interpretationen.

Bei der Arbeit in einem bewegungswissenschaftlichen Umfeld sieht man sich mit mehreren Faktoren konfrontiert. Die Bewegungswissenschaft ist eine junge wissenschaftliche Disziplin (ganz im Gegensatz zur Mathematik), welche sich empirischer Methoden bedienen muss, um neue Erkenntnisse zu generieren. Bei der Untersuchung der menschlichen Bewegung und aller damit verbundenen Bereiche, Aufgaben und Systeme wird der Mensch in den Fokus gestellt und damit die Forschung auf deduktivem Weg erschwert oder sogar unmöglich gemacht. In diesem Zusammenhang haben alle bisher durchgeführten Studien lediglich hypothetischen Charakter und damit verbunden können die Ergebnisse ausschließlich als vorläufig gelten. Zwar werden Elemente des kritischen Rationalismus verwandt, um Fehlinterpretationen zu vermeiden, wie die stringente Aufstellung einer Theorie und die empirische Prüfung, dennoch ist sich die bewegungswissenschaftliche Welt nicht über alle Bereiche einig. Häufig fehlen auch mehrere Studien zu einem Thema, um die Theorie auf ein

festes Standbein zu stellen. Dabei wird oft unterschätzt, dass erst ein wiederholtes Durchführen von Studien mit derselben Untersuchungsbedingung die Theorien glaubhaft macht (Weihe, 2004). Vorhandene Studien sollten dabei v. a. hinsichtlich der Methodik und Seriosität der Untersuchung durch diverse Forschergruppen kritisch hinterfragt und erneut realisiert werden.

1.1 Status Quo wissenschaftlicher Experimente im Bereich der Laufschuhforschung

Die facettenreichen Studien in der Bewegungswissenschaft umfassen unterschiedlichste Bereiche. Sie reichen von der Untersuchung der sensorischen Fähigkeiten, muskulärer Aktivitäten bei bestimmten Aufgabenstellungen, motorischer Eigenschaften von Personen mit bestimmten Krankheitsbildern wie z. B. Morbus-Parkinson bis hin zur Analyse des Gangbildes, der Laufeigenschaften und der Interaktion zwischen Läufer und Laufschuh. Bei näherer Betrachtung wissenschaftlicher Experimente im Bereich der Lauf- und Laufschuhforschung werden weitere detaillierte Untersuchungsbedingungen offenbar (Tabelle 1.1.1).

Allen diesen Untersuchungen obliegt der wissenschaftliche Versuch der Klärung der Interaktion zwischen Schuhwerk und Laufbewegung und die Analyse des Laufverhaltens bei unterschiedlichen Einflussbedingungen. Dabei kommen die beteiligten Forscher zum Teil zu unterschiedlichen Ergebnissen.

Milani et al. (1995) und Brauner et al. (2009) untersuchten beispielsweise den Einfluss einer Varus-Abschrägung von 8° respektive 4° auf das Laufverhalten. Beide Autoren gelangten zum Schluss, dass eine Varus-Abschrägung, eingebaut an der Ferse des Schuhs, die Pronationsgeschwindigkeit reduzieren kann. Die tatsächlichen mittleren Ergebnisse wichen jedoch voneinander ab. Milani et al. (1995) fanden eine Abnahme für die Pronationsgeschwindigkeit von $390.6 \pm 89.6^\circ/s$ auf $292.3 \pm 101.1^\circ/s$, wobei bei Brauner et al. (2009) eine Reduktion von $614.6 \pm 129.2^\circ/s$ auf $440.9 \pm 132.6^\circ/s$ verzeichnet wurde.

Im Zuge der Untersuchung des Einflusses einer härteren Mittelsohle fanden Nigg et al. (1987) keinen Einfluss auf die vertikalen Bodenreaktionskräfte sowie die Rückfußbewegung. Wohingegen vergleichsweise Milani et al. (1997) eine Abnahme der maximalen Pronation und Zunahme der maximalen vertikalen Kraftanstiegsrate mit zunehmendem Härtegrad registrierten. Der Einfluss einer Pronationsstütze anhand einer zweiten Härte wurde von Nigg und Bahlsen (1988) und mittels eigener Vorarbeiten analysiert, wobei Nigg und Bahlsen (1988) signifikante Differenzen für die initiale Eversion angeben. Dies konnte anhand der selbst durchgeführten Studie nicht nachgewiesen werden. Der Positionierung einer zweiten Härte wurde kein Einfluss auf biomechanische Variablen zugeschrieben und stellt damit die Effektivität einer zweiten Härte in Frage (Oriwol et al., 2011).

Bezüglich der Einlagenforschung wurden viele Studien veröffentlicht, welche zum Teil diverse Diskrepanzen aufweisen. Während Stacoff et al. (2000) keine Effekte durch Einlagen nachweisen konnten, fanden Eslami et al. (2009) eine Reduzierung der Eversionsbewegung um 4.1° und des aktiven Peaks der vertikalen Bodenreaktionskraft um 6%. Nach Nigg et al. (2003) und Mündermann et al. (2003) sind die gefundenen Effekte inkonsistent und nicht auf die Schuheinlagen zurückzuführen.

Diverse Gründe können derartige Diskrepanzen erklären, wobei ein genaues Eruiere der Fehler

Untersuchungsbedingungen	Literatur (Auszüge)
Analyse des Einflusses ...	
einer Varus- und/oder Valgus-Modifikation	Brauner et al. (2009), Grau und Horstmann (2007), Milani et al. (1995)
der Veränderung der Sohlenkonstruktion an der Ferse	Stacoff et al. (2001), Nigg und Bahlsen (1988), Nigg und Morlock (1987)
von Schuheinlagen	Eslami et al. (2009), Grau und Horstmann (2007), Nigg et al. (2003), Mündermann et al. (2003), Stacoff et al. (2000), Chen et al. (1994)
der Fersenkappe	van Gheluwe et al. (1995), Ferrandis et al. (1994)
des Mittelsohlenmaterials	Hardin et al. (2004), Milani et al. (1997), Hennig et al. (1996), Hamill et al. (1992), Nigg et al. (1987)
von Crashpads	Heidenfelder et al. (2010)
einer zweiten Härte	Nigg und Bahlsen (1988), Oriwol et al. (2011)
der Fußaufsatzcharakteristik	Pohl und Buckley (2008)
des Fußgewölbes	Butler et al. (2007), Butler et al. (2006)
der Tibiarotation	Digby et al. (2005)
von Inversion und Eversion	Perry und Lafortune (1995), Lafortune et al. (1994)
von Pronationseigenschaften	McClay und Manal (1998)
des Adaptationsverhaltens	Hardin et al. (2004)

Tabelle 1.1.1:

Zielsetzungen und Einflussfaktoren ausgewählter biomechanischer Studien im Bereich der Laufschuhforschung

oftmals nicht mehr möglich ist und lediglich Vermutungen angestellt werden können. Beim Vergleich zweier Studien kommen z. T. unterschiedliche Methoden zum Einsatz, z. B. kinematografische Aufnahmen und Elektrogoniometer zur Messung der Rückfußbewegung, welche differierende Ergebnisse erklären können. Verschiedene Versuchsbedingungen (z. B. 8° versus 4° Varus-Abschrägung) lassen eine Vergleichbarkeit ebenso wenig zu wie abweichende statistische Methoden. Vorstudien zur Einhaltung der Hauptgütekriterien (Objektivität, Reliabilität, Validität) wissenschaftlicher Untersuchungen werden nur zögerlich verwendet, missachtet oder nicht untersucht. In diesem Zusammenhang wird häufig auf eine Stichprobenplanung zur Bestimmung der optimalen Stichprobengröße verzichtet. Dies beinhaltet die Festlegung des α - und β -Fehlers und die Kenntnis über noch zu akzeptierende Unterschiede zweier vorgenommener Messungen. Oftmals haben Forscher kein genaues Wissen über derartige Abweichungen. Abhilfe kann eine vorgelagerte Reliabilitätsanalyse schaffen, welche allerdings selten angewendet wird. Wenige Autoren beschäftigen sich mit der Wiederholbarkeit von Messungen, wobei korrelative Zusammenhänge absolute Reliabilitätsmaße als Güte für die Wiederholbarkeit vorgezogen werden. Dabei sei zu beachten, dass erst absolute Relia-

bilitätsmaße Informationen über tatsächliche Abweichungen mehrerer Messungen liefern (Maiwald, Axmann & Grau, 2011).

Eine weitere Ursache liegt in der Tatsache begründet, dass ein bisher nicht untersuchtes Novum vorausgesetzt wird, um eine Veröffentlichung von Studien zu erreichen. Dieser Umstand erschwert die Vergleichbarkeit zweier Studien, welche sich stetig durch unterschiedlichste Versuchsbedingungen auszeichnen. Zwei vollkommen „identische“ Studien wird man unter den bewegungswissenschaftlichen Publikationen nicht finden. Im Sinne optimaler Erkenntnisfindung ist dies jedoch notwendig, um empirische Erkenntnis generieren zu können. So sind die meisten Ergebnisse lediglich Momentaufnahmen und keine durch mehrere Forscher fundierten Erkenntnisse (Weihe, 2004). Zudem ist es weit verbreitet und durch Reviewer forciert, lediglich „positive“, also signifikante Ergebnisse zu veröffentlichen, was den „publication bias“ erhöht und somit zusätzlich in die Irre führen kann (Weihe, 2004).

Bei der Betrachtung einzelner Studien aus methodischer und erkenntnistheoretischer Sicht spielen weitere Probleme eine Rolle. Unter Umständen ist die Theorie in sich nicht logisch aufgebaut oder hat falsche Annahmen, wobei einer falsch aufgestellten Theorie vehement durch logische Prüfung der Theorie sowie der aufgestellten Hypothesen und durch angemessene Literaturrecherche von vornherein entgegengewirkt werden sollte.

In den Achtziger und Neunziger Jahren wurde bei der Untersuchung der Laufbewegung die Pronationsbewegung als Ursache für Überlastungsschäden und damit einhergehender Verletzungen angesehen (Achillessehnenbeschwerden, Iliotibiales Bandsyndrom, Shin Splints, Ermüdungsbrüche (Stressfrakturen), Tendinitis) (Nigg & Bahlsen, 1988; Nigg & Morlock, 1987; Hennig et al., 1996; Hintermann & Nigg, 1998). In den letzten zehn Jahren wurde diese Annahme und die direkte Verbindung von Pronation und Verletzung von einigen Autoren kritisch hinterfragt (Nigg, 2001; Hreljac, Marshall & Hume, 2000) und zudem als nicht evident angesehen (Richards, Magin & Callister, 2009; Knapik et al., 2010). Studien, deren Schlussfolgerungen bestimmten Schuhmodifikationen weniger Pronationsverhalten und damit geringere Verletzungsmöglichkeiten bescheinigen, müssen daher streng genommen in Frage gestellt werden.

Des Weiteren ist das Aufbringen des Bewusstseins für Versuchsplanung, Stichprobenplanung und statistische Auswertung eine der Grundvoraussetzungen für ein wissenschaftliches Forschungsvorhaben und wissenschaftliche Tätigkeit. Oftmals werden Versuchsaufbau, Versuchsplan und das damit verbundene Forschungsdesign jedoch nicht adäquat gewählt. Damit einhergehend erfolgt die Auswahl der Merkmalsträger nicht zufällig oder stellt eine sehr eingeschränkte Stichprobe dar, welche nicht die gewünschte Gesamtpopulation widerspiegelt und daher als nicht repräsentativ gelten muss. Mittels ausreichender Lektüre forschungsmethodischer und statistischer Lehrbücher oder unter Zuhilfenahme einer statistischen Methodenberatung kann in diesem Zusammenhang diversen entstehenden Mängeln entgegengewirkt werden. Nichtwissen befreit nicht vom grundlegenden Verständnis und sollte im Sinne empirischer Forschung nicht befreien, sondern beseitigt werden. Dies gilt auch für eine exakte Dokumentation und die digitale Aufbereitung der Daten, da im Zuge der Anwendung diverser Messsysteme Fehler in der Datenaufbereitung auftreten können: u. a. bei der Verwendung digitaler Filter, bei der Auswertung interpolierter Messreihen oder bei der Analyse kinematografisch aufgenommener Daten (Beachtung von Reihenfolge und Richtung der durchgeführten Rotationen (Winter, 2005; Croce, Leardini, Chiari & Cappozzo, 2005)).

Allen bisher erwähnten methodischen, statistischen, analytischen und dokumentatorischen Problemen und Fehlern können die Forscher mit einer akkuraten Versuchsplanung entgegenwirken. Schwieriger gestaltet sich die Frage nach der Auswahl charakteristischer Parameter einer Zeitreihe, welche zur weiteren Auswertung herangezogen werden. Meist wird der Informationsgehalt, welchen die gesamte Zeitreihe enthalten kann, auf z. T. einen einzigen Wert beschränkt. Dies kann zu einer Diskrepanz zwischen einer punktuellen Analyse und einer Zeitreihenanalyse führen und eröffnet die Diskussion, ob eine Informationsreduktion von Daten auf einen einzigen Wert angemessen ist. Der entstehende Informationsverlust kann möglicherweise für Fehlinterpretation der Ergebnisse oder Falschdarstellung der Wirklichkeit verantwortlich sein.

Ein weiterer Aspekt bei der Betrachtung bewegungswissenschaftlicher Studien ist die Einbeziehung der Schwankung, die ein Proband selbst produzieren kann. Nur wenige Autoren vermerken eine derartige Schwankung oder ziehen sie in Betracht. In oben skizzierten Studien erwähnen lediglich Stacoff et al. (2000) und Oriwol et al. (2011), dass die Effekte, verursacht durch den Proband, größer sind als die Effekte, die der Schuh produziert.

Fraglich ist, welche Größen die Schwankung des Probanden und daher den Messfehler der aufgenommenen Signale beeinflussen. Wiederholte Diskrepanzen zwischen den Ergebnissen unterschiedlicher Studien zur gleichen Theorie werfen daher die Frage auf, inwiefern latente unberücksichtigte Variablen die Ergebnisse beeinflussen oder verzerren und inwiefern Variabilität als beeinflussender Faktor berücksichtigt werden kann.

1.2 Variabilität - Fluch und Segen

Variabilität muss als komplexes System verstanden werden, wobei Schwankungen sowohl innerhalb eines Probanden als auch zwischen mehreren Probanden präsent sind (James, 2004; Newell & Corcos, 1993)⁵. Zum Verständnis der intersubjektiven Schwankung ist es unumgänglich den Mensch als komplexes, dynamisches System zu verstehen, welches über sein sensorisches und motorisches System auf Umwelteinflüsse reagieren kann. Betrachtet man Bewegungen im Allgemeinen und die Laufbewegung im Speziellen, lassen sich verschiedene Randbedingungen und Einschränkungen charakterisieren. Nach Bates et al. (2004) existieren biomechanische, morphologische und Umgebungseinflüsse, welche eine Bewegung beeinflussen.

Biomechanische Einschränkungen determinieren sich durch physikalische Gesetzmäßigkeiten, wie Reibung oder Gravitation. Morphologische oder anatomische Restriktionen werden durch das Individuum selbst bestimmt (seine physische Statur und psychische Verfassung). Bezogen auf die Laufbewegung können das tatsächliche Bewegungsausmaß verschiedener Gelenke, Tagesform und eventuelle Ermüdungserscheinungen oder der Trainingszustand und damit die muskuläre Beschaffenheit eine Rolle spielen. Äußere Gegebenheiten wie Temperatur, Licht oder Bodenbeschaffenheit beeinflussen zusätzlich die ausgeführte Bewegung, wobei Reize, z. B. ein veränderter Bodenbelag, über Rezeptoren wahrgenommen werden. Diese Information wird mittels elektrischer Signale zum zentralen Nervensystem weitergeleitet und verarbeitet, was letztlich zur Aktivierung der Muskulatur führt und somit eine Reaktion auf den Reiz herbeiführt, z. B. durch ein verändertes Fußauf-

⁵ Es sei hier angemerkt, dass sich das Kapitel 1.2 zur Variabilität an die Überlegungen von Bates, James und Dufek (2004) und James (2004) anlehnt.

satzmuster (Speckmann, Heschler & Köhling, 2008). Infolge dieser Komplexität des menschlichen Systems und der gestellten Bewegungsaufgabe entsteht Variabilität, da das neurophysiologische System ständig auf Bewegungsaufgaben und Umwelteinflüsse reagiert und somit immer eine veränderte Bewegung hervorruft. Diese Variabilität gehört unwiderruflich zum System Mensch. Es ist daher unmöglich, denselben Bewegungsablauf zweimal durchzuführen, auch wenn alle Bedingungen gleich gehalten werden.

In diesem Zusammenhang kann Variabilität als positiver Faktor angesehen werden, da der Organismus in der Lage ist, auf verschiedene Umwelteinflüsse zu reagieren, ohne das System zu verletzen oder gar zu zerstören. Somit ist Variabilität zur Aufrechterhaltung des dynamischen menschlichen Systems notwendig und eine erhöhte Variabilität durchaus positiv zu betrachten. Zudem belegen Forschungen, dass bei gesunden Personen die Variabilität des menschlichen dynamischen Systems erhöht ist. Goldberger, Rigney und West (1990) untersuchten die Herzfrequenz von gesunden Menschen und von Patienten kurz vor einem Herzstillstand. Die Dynamik gesunder Herzfrequenzen gemessen mittels Spektralanalyse und Phasenraumportraits ist eher chaotisch, wohingegen die Herzfrequenz erkrankter Personen periodisch erscheint.

Bei der Betrachtung der Komplexität von EEG-Signalen von Parkinsonpatienten im Vergleich zu Gesunden (im Ruhezustand) ermittelten Stam et al. (1994) eine geringere Dimensionalität, wobei der Begriff der EEG-Komplexität aus der Theorie des deterministischen Chaos stammt. Dessen Algorithmen, angewendet in der nichtlinearen EEG-Dynamik, geben Aufschluss über die Komplexität eines Systems, wobei die Dimensionalität ein wichtiger Parameter ist. Die Zahl entspricht den Dimensionen, die zur Beschreibung der Dynamik erforderlich sind. Daraus folgt, je höher die Dimensionalität ist, desto größer ist die Komplexität des Systems (Schiepek, 2004). Nach Stam et al. (1994) scheint dieser Verlust an dynamischer Komplexität im Kortex auf einer geringeren Aktivität des cholinergen Systems zu basieren.

Nach Collins, Imhoff und Grigg (1996) kann eine Verbesserung der Oberflächensensibilität der kutanen Mechanorezeptoren herbeigeführt werden, indem dem Reizsignal eine aperiodische stochastische Resonanz beigefügt wird. Diese stochastischen Methoden könnten nach van Emmerik und van Wegen (2000) zur Ermittlung von sensorischen Grenzwerten für neurologische Funktionsstörungen hilfreich sein.

Diese Beispiele verdeutlichen die positive Funktion und Bedeutung, welche der Variabilität zugeschrieben wird, und lösten u. a. eine Diskussion über den möglichen Zusammenhang von verringerter Variabilität in Verbindung mit Erkrankungen aus (van Emmerik & van Wegen, 2000). Der positive Aspekt gilt nicht uneingeschränkt. Zu viel Variabilität kann nach van Emmerik und van Wegen (2000) schädlich sein. Bei der Untersuchung von Parkinsonpatienten konnte eine sehr hohe Variabilität in den Bewegungsmustern festgestellt werden. Dies war vor allem dann der Fall, wenn starke Asymmetrien bei der Bewegung der Gliedmaßen auftraten z. B. ein halbseitiger Tremor. In diesem Fall wurde die Variabilität derart extrem, dass kein stabiles Bewegungsmuster erreicht werden konnte.

Aus diesen Überlegungen wird deutlich, dass Variabilität als beeinflussender Faktor, bei Messungen und anschließenden Auswertungen nicht außen vor gelassen werden darf und bei gesunden und erkrankten Personen verschieden diskutiert werden muss. Aus Sicht dynamischer Systeme bestimmt Variabilität die Stabilität von Bewegungsmustern, erlaubt Flexibilität innerhalb des neuromuskulä-

ren Systems, gestattet die Auswahl oder den Wechsel zu neu erlernten Fähigkeiten und stellt einen stochastischen Störeinfluss dar (Bates et al., 2004). Nach Newell und Corcos (1993) und van Emmerik und van Wegen (2000) ist die direkte Verlinkung von Variabilität mit Stabilität in Bezug auf die körperliche Haltung einer Person schwierig, wenn nicht sogar unmöglich, so dass sich Stabilität in diesem Kontext lediglich auf die Bewegungsmuster bezieht.

Die bei einer Messung aufgenommenen Signale S_T enthalten sowohl das tatsächliche biologische Signal S_B und einen variablen Faktor V_T (Bates et al., 2004; Newell & Slifkin, 1998):

$$S_T = S_B + V_T. \quad (1.2.1)$$

S_B kann bezüglich der Messtheorie als wahrer Wert eines Probanden bezeichnet werden (siehe Kapitel 1.4) (Steyer & Eid, 2001). Die Gesamtvariabilität in einem stochastischen System setzt sich zudem aus zwei Quellen zusammen, einer zufälligen Schwankung V_E und einer deterministischen chaotischen Schwankung V_n , auch als nichtlinearer dynamischer Prozess bezeichnet, welcher mathematisch vorhergesagt werden kann:

$$V_T = V_n + V_E.$$

Der zufällige Anteil V_E ist eine Kombination aus Noise V_{EB} , methodischem Fehler V_{EM} und sonstigem Fehler V_{EE} . Somit kann das gemessene Signal wie folgt dargestellt werden

$$S_T = S_B + V_n + V_{EB} + V_{EM} + V_{EE}. \quad (1.2.2)$$

Das stochastische Rauschen V_{EB} kann in verschiedenen Formen innerhalb eines biologischen Systems auftreten: als weißes Rauschen $\frac{1}{f^0}$ (zufällige Schwankung), als pinkfarbenedes Rauschen $\frac{1}{f^1}$, als braunes Rauschen $\frac{1}{f^2}$ und als schwarzes Rauschen $\frac{1}{f^b}$, $b > 2$, wobei f die mittlere Frequenz der Schwankung definiert. Je größer der Exponent ist, um so strukturierter ist das Signal (Newell & Corcos, 1993; Newell & Slifkin, 1998). Für die Betrachtung der Bewegungsvariabilität wird oftmals ein weißes Rauschen angenommen, obwohl nach Newell und Slifkin (1998) zunächst eine Überprüfung erfolgen sollte, inwiefern eine zufällige Schwankung angenommen werden kann.

Die Behandlung und Einbeziehung des variablen Terms ist je nach Quelle unterschiedlich. Bei Schwankungen, die zwischen mehreren Probanden entstehen, wird man auf die Hilfe parametrischer Testverfahren zurückgreifen z. B. das allgemeine lineare Modell, welches die Variabilität über Varianzanteile einbezieht. Effektstärken dieses Verfahrens beurteilen den gesamten Varianzanteil gegenüber dem Fehlerterm der Varianzanalyse und lassen somit Aussagen über die Güte des vorliegenden mittleren Effektes der untersuchten Bedingung zu.

Die Einbeziehung der Variabilität, die ein Proband selbst produziert, gestaltet sich schwieriger. Aus methodischer Sicht hat man mehrere Möglichkeiten den gesamten variablen Term einzubeziehen. Man könnte den Fehlerterm ignorieren. Dabei ist es bei numerischer Integration oder Ableitung möglich, dass sich der Fehler zusätzlich erhöht. Aus Sicht obiger Diskussion bezüglich des Einflusses von Variabilität wäre dies jedoch kontraproduktiv, da zwar Variabilität einbezogen werden soll, jedoch nur in ihrer tatsächlichen Ausprägung.

Des Weiteren könnte man versuchen den variablen Anteil durch Filtern zu beseitigen oder zu verringern. Bei der Verwendung von low-pass-Filtern werden die höheren Frequenzen abgeschnitten

und die niedrigeren Frequenzen werden im Signal beibehalten. Die Probleme bleiben bestehen. Bei der Anwendung zu hoher Frequenzen können die Fehler erhöht werden, indem sich diese bei Integration und Ableitung aufsummieren. Ist die ausgewählte cut-off-Frequenz zu niedrig, wird zuviel des eigentlichen Signals gelöscht und nicht mehr korrekt dargestellt. Wichtige Informationen würden damit verloren gehen.

Eine weitere Möglichkeit, Variabilität zu betrachten, ist die Untersuchung der Anzahl an Versuchen. Je mehr an einzelnen Signalen aufgenommen wird, umso größer ist die Wahrscheinlichkeit, dass das gesamte Variabilitätsspektrum des Probanden aufgenommen und abgedeckt wird. Dabei ist die Menge an Versuchen entscheidend und die Frage nach der minimalen Anzahl an Versuchen ist interessant.

1.3 Berücksichtigung der Versuchswiederholung zur Charakterisierung der Variabilität

Bei der Betrachtung der in Kapitel 1.1 vorgestellten bewegungswissenschaftlichen Studien hinsichtlich vorgenommener Standardisierungen ist die Vielfalt der eingesetzten Messsysteme augenscheinlich. Entsprechend den verwendeten Messgeräten und der Aufgabenstellung erfolgen unterschiedliche Standardisierungen hinsichtlich der Stichprobengröße, der Anzahl an Versuchen, der vorgegebenen Geschwindigkeit, angewendeter Filter und Aufnahmefrequenzen, Labor- oder Felduntersuchungen sowie Untergründen und Laufart.

Tabelle 1.3.1 zeigt einen Überblick über methodische Besonderheiten obiger ausgewählter Studien. Detaillierte Informationen können Tabelle A.1 auf Seite 112 entnommen werden.

Dabei ist auffällig, dass fast alle Studien unter Laborbedingungen durchgeführt wurden, wobei 17 Studien einen festen Untergrund und acht Studien ein Laufband verwendeten. Mittels Laufband soll eine länger anhaltende Laufbewegung, wie sie beim tatsächlichen Einsatz eines Laufschuhs entsteht, simuliert werden. Letztlich wurde in obigen Studien dennoch eine endliche Anzahl an Versuchen verarbeitet und nicht der gesamte gemessene Zeitraum. Möglichkeiten zur Auswertung von Zeitreihen wurden in den letzten Jahren in die Bewegungswissenschaft eingeführt und verwendet, z. B. der Lyapunov Exponent (Bruijn, Dieën, Meijer & Beek, 2009a, 2009b) oder der scaling Exponent α einer trendbereinigenden Fluktuationsanalyse (Nakayama, Kudo & Ohtsuki, 2010). Insofern wäre eine Auswertung des gesamten gemessenen Zeitraumes möglich. Es bleibt dennoch die Frage bei Laufbandstudien bestehen, ob die gemessene Laufbewegung auf einem Laufband die tatsächliche Laufbewegung auf festem Untergrund widerspiegelt. Letztlich bewegt sich der Untergrund beim Laufband und der Mensch muss andere neuronale Verbindungen herstellen als beim Laufen auf festem Untergrund, da er sich dort aktiv nach vorn bewegen muss (Wank, Frick & Schmidtbleicher, 1998).

Des Weiteren wurden Standardisierungen bezüglich der vorgegebenen Geschwindigkeit vorgenommen. Vor allem Kraftdaten sind u. a. hinsichtlich der Geschwindigkeit, mit der ein Lauf absolviert wird, anfällig. Je höher die Geschwindigkeit ist, um so stärker äußert sich dies in erhöhten zweiten Kraftspitzen und größeren Kraftanstiegsraten (Cavanagh & Lafortune, 1980). Somit scheint eine Standardisierung notwendig. Die hier vorgegebenen Geschwindigkeitsintervalle variieren sehr stark

Zusammenfassung	
Artikel	24 Artikel gesamt
Stichprobengröße	15 ± 8 (minimal 5 bis maximal 40 Probanden)
Anzahl an Versuchen	6 ± 3 (minimal 1 bis maximal 12 Versuche)
Laufart und Untergründe	23 × laufen, 3 × gehen 17 × Boden, 8 × Laufband
Geschwindigkeit	laufen: minimal 2.4 m/s bis maximal 4.3 m/s, 3 mal individuell gehen: minimal 1.4 m/s bis maximal 1.6 m/s, 1 mal individuell
Aufnahme- und Filterfrequenz (weitere Messgeräte)	Kraftmessplatte: 960 Hz bis 2400 Hz (10 Hz bis 100 Hz) Kinematographie: 60 Hz bis 1000 Hz (8 Hz bis 40 Hz) (Elektrogoniometer, Fragebögen, Druckverteilung, Beschleunigungsaufnehmer, Metabolische Daten)
Labor- und Felduntersuchung	24 Labormessungen, 1 integrierte Feldmessung

Tabelle 1.3.1:

Zusammenfassung methodischer Besonderheiten der Studien aus Tabelle 1.1.1

und hinterlassen einen beunruhigenden Beigeschmack hinsichtlich der Vergleichbarkeit von unterschiedlichen Studien. Nach Queen, Gross und Liu (2006) hingegen scheint der Einfluss einer vorgegebenen Geschwindigkeit gegenüber einer selbst gewählten Geschwindigkeit geringe Auswirkungen auf die Variabilität eines Probanden zu haben. Insofern kann man hinsichtlich der Vergleichbarkeit von Studien mit Hinblick auf die vorgegebene Geschwindigkeit Nachsicht üben.

Lediglich drei der 24 Autoren berechneten die Stichprobengröße vor der Messung (Butler et al., 2007, 2006; Hardin et al., 2004). Der Vorteil einer Stichprobenplanung besteht zunächst in der Festlegung von zu untersuchenden Unterschieden zwischen verschiedenen Bedingungen. Des Weiteren müssen α -Fehler und Power festgelegt werden. Mit diesen Vorgaben wird die Stichprobengröße derart ermittelt, dass der vorgegebene Effekt tatsächlich aufgedeckt werden kann. Letztlich muss man sich fragen, inwiefern der gefundene Effekt zu bewerten ist, wenn es an einer Stichprobenplanung fehlt. Oftmals mangelt es auch an der Angabe von Effektstärken, welche eine erweiterte Bewertung des Effektes zulassen, wobei für die Biomechanik die genaue Einordnung der Größe des Effektes anhand von Varianzanteilen geklärt werden muss. Eine Aufstellung möglicher Effektstärken wie bei Cohen (1992), welcher diese für psychologische Problemstellungen darstellt, fehlt in der Literatur der Forschungsmethodik der Biomechanik. Eine Orientierung an den Effektstärken von Cohen (1992) ist dennoch aufschlussreicher als die Vernachlässigung einer Stichprobenplanung.

Im Hinblick auf die Variabilität eines Probanden hat man nach Kapitel 1.2 zwei Möglichkeiten diese zu berücksichtigen: die Wahl des Filters inklusive Filterfrequenz und die Anzahl an gemessenen Versuchen pro Proband. Als Filter wurde stets ein Butterworth-Filter verwendet. Die Aufnahmefrequenzen hingegen sind vielfältig und unterscheiden sich je Messgerät. Dabei werden die aufgenom-

menen Signale der Bodenreaktionskräfte (gemessen mit Kraftmessplatten) mit höheren Frequenzen aufgenommen und gefiltert. Mittels kinematografischer Aufnahmegerate gemessene Daten wurden mit niedrigen Frequenzen gefiltert. Dabei kommt die Problematik zum Tragen, dass man im ersten Fall möglicherweise zu viele Fehler beibehält und im zweiten Fall zu viel Information verloren geht (Digby et al., 2005). Bei der gleichzeitigen Aufnahme von Bodenreaktionskräften und kinematografischen Daten werden für die Aufnahmefrequenzen der Kraftmessplatten häufig Vielfache der Frequenzen der verwendeten Kamerasysteme eingestellt. Dies ist im Sinne der Datenverarbeitung hilfreich, da bei der Interpolation der Kraftdaten auf die gleiche Anzahl an Messzeitpunkten zwar Daten verloren gehen, aber keine neuen Datenpunkte generiert werden müssen. Somit kann man sicher sein, dass der Datenverlauf homogen bleibt.

Durchschnittlich wurden in diesen Studien 15 ± 8 Probanden untersucht und 6 ± 3 Einzelversuche aufgezeichnet. Angesichts der geringen Anzahl an Versuchen pro Proband und obiger Variabilitätsbetrachtung stellt sich die Frage, inwiefern 6 ± 3 Versuche ausreichend sind, um die Variabilität des Probanden vollständig zu charakterisieren.

Ein Blick in die Literatur offenbart diverse Vorschläge, wie viele Läufe pro Proband aufgezeichnet werden sollten. Bates, Osternig, Sawhill und James (1983) schlugen für die Messung der Bodenreaktionskräfte acht Versuche vor, wobei zehn Versuche gemessen wurden. Dabei wurde die minimale Anzahl i an Versuchen anhand diskreter Werte derart bestimmt, dass der Mittelwert dieser Einzelwerte MW_{i+1} innerhalb des vorgegebenen Intervalls $[MW_{10} \pm \frac{1}{4}SD_{10}]$ noch enthalten sein musste und MW_i sich außerhalb des Intervalls befinden sollte. In einer späteren Studie von DeVita und Bates (1988) wird die Verwendung von 25 Versuchen vorgeschlagen, um die Variabilität der Bodenreaktionskräfte eines Läufers ausreichend zu charakterisieren. Die Aussage basiert auf der Verwendung von ursprünglich 50 Versuchen. In dieser Studie wurde ein ähnliches Kriterium zur Bestimmung der minimalen Anzahl wie bei Bates et al. (1983) verwendet, wobei der Mittelwert der gesamten Kurve der vertikalen Bodenreaktionskraft bestimmt wurde und verschiedene Anteile der Standardabweichung Verwendung fanden. Des Weiteren ist zu erwähnen, dass die Bestimmung der minimal notwendigen Versuche von den experimentell aufgezeichneten abhängig ist. Demzufolge wurde von DeVita und Bates (1988) zusätzlich ein theoretisches Modell aufgestellt, das die experimentell ermittelte minimale Anzahl von 25 Läufen untermauern sollte.

Nach den Untersuchungen von DeVita und Bates (1988) fand ein scheinbarer Wandel in der Auffassung über die Anzahl zu messender Versuche pro Proband statt:

- Von Kadaba et al. (1989) wurden drei Läufe zur Messung der Bodenreaktionskräfte als ausreichend angesehen. Dies wurde mittels Schrumpfungsalgorithmen und dem angepassten Bestimmtheitsmaß einer damit verbundenen multiplen Regression bewertet (Venter & Maxwell, 2000).
- Hamill und McNiven (1990) schlugen zehn Läufe mit einer Aufwärmphase von zehn Läufen vor, wobei die Beurteilung über obiges Mittelwertkriterium für 20 Läufe und ein Viertel der Standardabweichung erfolgte.
- Von fünf aufgezeichneten Läufen der Bodenreaktionskraft wurden von White et al. (1999) drei als akzeptabel angesehen. Zur Bewertung wurde die Signifikanzschwelle einer Varianzanalyse herangezogen. Hierbei muss Erwähnung finden, dass die Wiederholbarkeit von Versuchen mittels eines Äquivalenztests oder adäquaten Methoden untersucht werden sollte. Eine Varianz-

analyse ist hier völlig ungeeignet, da deren primäres Ziel in der Aufdeckung von Unterschieden zu finden ist.

- Diss (2001) ermittelte, dass sieben von zehn Signalen der Bodenreaktionskraft aufgezeichnet werden sollten, um Reproduzierbarkeit zu erreichen. Als Entscheidungskriterium wurde die Spearman-Brown-Formel mit einem Reliabilitätslevel von 90 % verwendet (Moosbrugger, 2007).
- Für kinematische Messungen gaben Laroche et al. (2011) fünf bis zehn Versuche für gute Reliabilität an (gemessen mittels 30 Versuchen anhand Intraclasskorrelationen und Bland and Altman Verfahren) und von Webster et al. (2010) wurden 6 Wiederholungen als ausreichend angesehen.
- Bezüglich der Messung auf einem Laufband gaben Belli et al. (1995) 32 bis 64 Schritte als akzeptabel an. Dies ergibt mindestens 15 bis 20 s Laufumfang auf einem Laufband.

Neben diesen Studien, welche sich der Ermittlung einer minimalen Anzahl an Versuchen widmen, untersuchten einige Autoren, welche Bedingungen die Variabilität eines Probanden beeinflussen können. Dabei wurde u. a. der Einfluss der Geschwindigkeit und Schrittfrequenz untersucht (Karamanidis et al., 2004, anhand 3 Läufen), der Einfluss einer vorgegebenen Geschwindigkeit gegenüber einer selbstgewählten Geschwindigkeit analysiert (Queen et al., 2006, anhand 6 Läufen), der Einfluss unterschiedlicher Messmethoden zur Ermittlung marker-basierter 3D-kinematographischer Parameter ermittelt (Pohl et al., 2010, anhand 5 Wiederholungen), die Beeinflussung durch verschiedene Untersuchungsmethoden kinematischer Messungen beurteilt (Long et al., 2010, 1 Messung, 4 Messtermine), der Einfluss von Schuhwerk erforscht (Kurz & Stergiou, 2003, 10 Wiederholungen) und der Einfluss des Alters bestimmt (Korhonen et al., 2010, 8 Wiederholungen). Bis auf das Alter wurde keiner obigen Bedingung ein Einfluss auf die Variabilität des Probanden zugeschrieben.

Betrachtet man die vorgeschlagene Mindestanzahl an Versuchen für die Messung der Bodenreaktionskräfte oder kinematischer Messungen, scheinen sich die Autoren uneinig zu sein. Schlugen DeVita und Bates (1988) noch 25 Versuche vor, verringerte sich diese Anzahl im Laufe der Jahre auf drei bis zehn Messungen pro Proband. Die Frage bezüglich obiger Studien, inwiefern 6 ± 3 Versuche ausreichend sind, um die Variabilität des Probanden vollständig zu charakterisieren, steht dabei weiterhin im Raum. Obige Autoren orientierten sich demnach eher an der Mindestanzahl von drei bis zehn. Dabei ist jedoch zu beachten, dass die Bestimmung der minimalen Anzahl an Läufen, die bei jeder Messung aufgezeichnet werden sollte, abhängig von der Anzahl zunächst experimentell gemessener Versuche ist. Diss (2001) zum Beispiel zeichnete lediglich zehn Läufe auf. Wenn man davon ausgeht, dass diese zehn Läufe bereits die ganze Bandbreite an Variabilität des Probanden abbilden, dann kann eine Anzahl von sieben Läufen ausreichen, um das entsprechende Reliabilitätslevel abzudecken. Dies ändert nichts an der Tatsache, dass die getroffene Annahme der vollständigen Variabilitätscharakterisierung möglicherweise inkorrekt ist.

Bei Entscheidungskriterien, welche von der Gesamtzahl gemessener Versuche abhängen wie Standardabweichungen und Reliabilitätslevel, ist die letztliche Bestimmung einer minimalen Anzahl wiederum abhängig von dieser Anzahl zunächst aufgenommener Versuche. Wenn diese Gesamtanzahl nicht die gesamte Variabilität charakterisiert, sind Aussagen über die Mindestanzahl wertlos, da sie sich lediglich auf einen Teil der Variabilität des Probanden beziehen. Daher sind auch die Studien hinsichtlich der Beeinflussung der Variabilität, welche ausschließlich auf kleinen Wiederho-

lungszahlen beruhen, mit Vorsicht zu betrachten. Somit lässt die Berechnung der Mindestanzahl, die auf einer geringen Gesamtanzahl beruht, zunächst keine theoretische Verallgemeinerung auf minimal zu messende Versuche zu.

Ein weiteres Problem einer geringen Anzahl an Versuchen liegt in der möglichen Beeinflussung mehrerer Versuchsbedingungen untereinander, so dass die geforderte statistische Unabhängigkeit von Ereignissen verletzt werden könnte. Man nimmt zwar an, dass die Messungen durch den Proband beeinflusst werden und daher als abhängige Ereignisse gelten. Die untersuchte Bedingung, z. B. der Einfluss von Schuhwerk oder der Einfluss des Alters, sollte jedoch unabhängige Ereignisse generieren. Im Falle des Alters wird dies absolut eingehalten, da diese Bedingung durch unterschiedliche Probandengruppen geschaffen werden muss. Der Einfluss des Schuhwerks wird dagegen innerhalb eines Probanden untersucht und muss dennoch unabhängig werden. Beim Wechsel der zu untersuchenden Schuhbedingung könnten sich diese gegenseitig beeinflussen und die ersten Läufe im neuen Schuh durch den vorhergehenden Schuh geprägt sein. Der Vorschlag von Hamill und McNiven (1990) zehn Probeversuche vor die eigentliche Messung sowie die Empfehlung von Stergiou und Scott (2005) sogenannte Basismessungen zwischen die Versuchsbedingungen zu schieben, könnte dieses Problem lösen. Der Vorteil einer größeren Anzahl an Messversuchen liegt insbesondere darin, die ersten und letzten Messungen einer Versuchsbedingung nicht in die eigentliche Auswertung aufzunehmen, um die Unabhängigkeit der Messungen zu gewährleisten und dennoch die vollständige Variabilität des Probanden zu charakterisieren. Schlussfolgernd ist eine Untersuchung und Bestimmung der Mindestanzahl an Versuchen anhand einer großen Datenmenge durch adäquate Methoden sinnvoll und es ergeben sich folgende Fragestellungen:

- *Kann die Variabilität eines Probanden mittels einer endlichen Anzahl an Wiederholungen charakterisiert werden?*
- *Welche Algorithmen stehen zur Ermittlung dieser endlichen Anzahl an Versuchen zur Verfügung?*

Die Auswertung aller Studien aus Tabelle 1.1.1 auf Seite 5 (↗ A.1, S. 112) erfolgte fast ausschließlich anhand des Mittelwertes diskreter Parameter der gemessenen Kurven der Bodenreaktionskräfte, der tibialen Beschleunigung oder von Winkelverläufen. Bei diesem Vorgehen wird die aufgenommene Information über den Probanden erheblich reduziert. Damit einhergehend muss die Frage geklärt werden, ob die Variabilität eines Probanden anhand diskreter Werte oder anhand der gesamten Kurve Beachtung finden muss und inwiefern der Mittelwert als die den Probanden repräsentierende Kennzahl verwendet werden kann.

1.4 Der Mittelwert als charakteristische Kennzahl eines Probanden

Alle bisherigen Überlegungen bezogen sich ausschließlich auf die Variabilität, deren Auswirkungen, Messung und Beeinflussung. Ausgehend vom gemessenen Signal eines Probanden $S_T = S_B + V_T$ (↗ Glg. 1.2.1) wurde somit lediglich der variable Term V_T betrachtet und dem biologischen Signal S_B bisher keinerlei Beachtung geschenkt. Legt man messtheoretische Konzepte zugrunde, wird das biologische Signal des Probanden auch als „wahrer“ Wert bezeichnet. Nach Steyer und Eid (2001)

bezeichnet der wahre Wert den Erwartungswert $E(S_T|p_a)$ der intraindividuellen Verteilung des Probanden p_a und setzt sich mit einer Fehlervariable (hier V_T) additiv zur Testvariable (hier S_T) zusammen. Mit $S_B = E(S_T|p_a)$ ergibt sich somit

$$S_T = E(S_T|p_a) + V_T. \quad (1.4.1)$$

Welche statistische Kennzahl oder Funktion eignet sich am besten zur Beschreibung des Erwartungswertes eines Probanden und welche Eigenschaften können aus diesem Modell abgeleitet werden? Bei der Beantwortung dieser Fragen helfen zwei bekannte Sätze aus der mathematischen Statistik - das Gesetz der großen Zahlen und der zentrale Grenzwertsatz.

Das Gesetz der großen Zahlen besagt, dass für eine sehr große Anzahl an Wiederholungen sich der Erwartungswert der Verteilung einstellen wird und sich das arithmetische Mittel diesem derart annähert, dass die Differenz aus Erwartungswert und arithmetischem Mittel für $n \rightarrow \infty$ gegen Null konvergiert (Fahrmeir, Künstler, Pigeot & Tutz, 2007). Es sei hier angemerkt, dass der Erwartungswert einer Verteilung sich zunächst vom arithmetischen Mittel (Stichprobenmittelwert) unterscheidet und als dessen Realisierung gilt. Das arithmetische Mittel charakterisiert die Lage von Daten innerhalb einer Stichprobe, wobei der Erwartungswert die Lage einer Verteilung im Allgemeinen definiert, ohne dass Daten vorliegen müssen (Fahrmeir et al., 2007). Übertragen auf obige Situation einer bewegungswissenschaftlichen Messung ist es somit möglich, anhand einer sehr großen Wiederholungszahl den Erwartungswert des Probanden $E(S_T|p_a)$ sehr genau mit Hilfe des arithmetischen Mittels zu schätzen.

Es liegt somit nahe, bei genügend großer Anzahl das arithmetische Mittel als die den Probanden repräsentierende Kennzahl zu verwenden. Dies entspricht dem üblichen Vorgehen, wobei die wiederholten Messungen innerhalb bewegungswissenschaftlicher Studien im Mittel zusammengefasst und weiterverarbeitet werden. Durch Verwendung des Mittelwertes wird der variable Term herausgerechnet und geht für alle weiteren Auswertungen verloren. Das arithmetische Mittel ist ohne die Kenntnis der Streuung jedoch nicht aussagekräftig, um genaue Angaben über eine Verteilung treffen zu können und den Probanden vollständig zu charakterisieren (Krämer, 1991).

Eine weitere Eigenschaft folgt aus dem zentralen Grenzwertsatz. Dieser besagt, dass „beim unabhängigen, additiven Zusammenwirken vieler kleiner Einflussgrößen“ eine Normalverteilung zu beobachten ist (Hilgers, Bauer & Scheiber, 2007, S. 60). Eine visuelle experimentelle Überprüfung dieses Satzes liefert das Galton Brett. Bei diesem Experiment werden Kugeln in ein mit gitterförmig versetzten Hindernissen versehenes Brett geworfen und überprüft, in welcher Röhre sie am unteren Ende landen (siehe Abbildung 1.4.1). Jede einzelne Kugel folgt dabei einer Binomialverteilung, da an jedem Hindernis zwei Möglichkeiten bestehen, in welche Richtung sich die Kugel orientieren kann. Die Wahrscheinlichkeit, dass eine Kugel ganz rechts oder links landet, ist dabei viel geringer als dass sie sich in Richtung Mitte bewegt. In Summe ergibt sich daraus die gaußsche Glockenkurve, auch Normalverteilungskurve genannt. Aus dem zentralen Grenzwertsatz folgt, dass sich alle unabhängigen, identisch verteilten Zufallsgrößen in Summe derart verhalten, dass sie bei genügend großer Anzahl eine Normalverteilungskurve bilden und die Verteilungsart für dieses Verhalten unerheblich ist.

Betrachtet man beispielhaft die maximalen Kraftanstiegsraten der vertikalen Bodenreaktionskraft eines Probanden aus 100 Versuchen gemessen mittels einer Kraftmessplatte, dann wird das Verhalten, welches der zentrale Grenzwertsatz beschreibt, deutlich (siehe Abbildung 1.4.2). Obwohl die

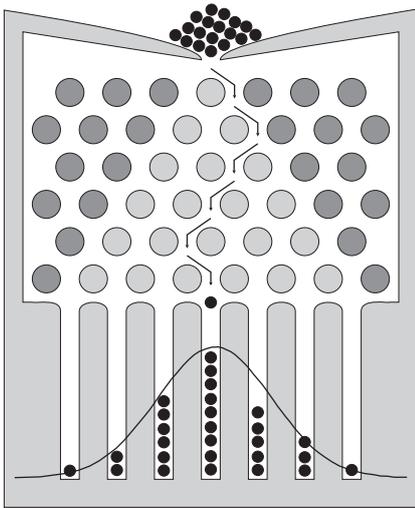


Abbildung 1.4.1:
Beispiel eines Galton Bretts zur Veranschaulichung des zentralen Grenzwertsatzes

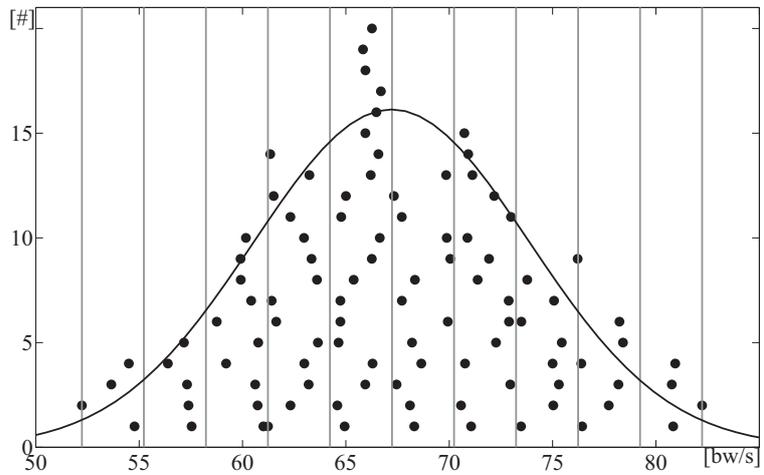


Abbildung 1.4.2:
Verteilung der Werte der maximalen Kraftanstiegsrate der vertikalen Bodenreaktionskraft eines Probanden bei 100 Messungen auf einer Kraftmessplatte

Verteilung jedes einzelnen Wertes unbekannt ist, aber die Werte der maximalen Kraftanstiegsrate durch die Betrachtung eines Probanden identisch verteilt sind, folgen sie in Summe einer Normalverteilung.

Für das Residuum, den Fehlerterm, lassen sich zudem noch weitere Eigenschaften schlussfolgern. Versteht man das gemessene Signal $S_T = E(S_T|p_a) + V_T$ als Regression auf den wahren Wert (geschrieben als bedingter Erwartungswert und Residuum), dann folgt aus der Definition des Fehlerterms einer Regression und den Eigenschaften eines bedingten Erwartungswertes, dass das Residuum zunächst den Erwartungswert Null hat und der wahre Wert und das Residuum unkorreliert sind. Für eine Beweisführung sei auf Steyer und Eid (2001) Seite 364 verwiesen. Die Annahme, dass sich die Fehler eines biologischen Systems normalverteilen und im Mittel Null sind, ergibt sich somit aus der Definition des Residuums selbst und aus den Aussagen des zentralen Grenzwertsatzes.

Vereinfacht soll dies an einem bewegungswissenschaftlichen Beispiel verdeutlicht werden - dem Lauf über einen Schotterweg, wobei bei jedem Schritt der Fuß den unebenen Untergrund ausgleichen muss. Der wahre Wert in diesem Beispiel wäre der optimale Fußaufsatz auf einem ebenen Untergrund, z. B. asphaltierter Boden. Der Fehlerterm bezeichnet dann alle Abweichungen von diesem optimalen Fußaufsatz. Als Vereinfachung sei angenommen, dass diese Abweichung als einzige Fehlerquelle existiert. Wenn die Fehler dieses Systems im Erwartungswert nicht Null ergeben, würde der Läufer auf halber Strecke umfallen oder auch etwas später, da sich der Fehler mit jedem Schritt aufsummiert. Der Läufer wäre irgendwann nicht mehr in der Lage, die extreme Fußstellung auszugleichen und würde in eine arge Schiefelage geraten oder sich schlimmstenfalls Verletzungen zuziehen. Es sei an dieser Stelle angemerkt, dass in einer tatsächlichen Messsituation wesentlich mehr Fehlerquellen auftreten können. Bei Existenz mehrerer Fehler wird oftmals deren Unkorreliertheit vorausgesetzt, dies ist jedoch keine aus der Definition des Residuums ableitbare Eigenschaft (Steyer

& Eid, 2001).

Aus diesen Überlegungen heraus kann geschlussfolgert werden, dass sich die Daten eines Probanden normalverteilen sollten und die Messfehler im Erwartungswert Null ergeben. Des Weiteren folgt, dass das arithmetische Mittel als Repräsentant des Probanden für eine biomechanische Messung verwendet werden kann. Über eine genügend große Anzahl an Wiederholungen müsste sich dieser Wert nach einer gewissen Zeit einstellen. Nach Definition des arithmetischen Mittels haben alle anderen Zahlen in Summe den kleinsten quadratischen Abstand zu dieser Zahl. „Diese minimierende Eigenschaft arithmetischer Mittelwerte spricht [zudem] für ihre Repräsentativität“ (Axmman, Groten, Rupp & Scheideler, 2006, S. 215) und es ergeben sich folgende Fragestellungen:

- *Kann in einer biomechanischen Messsituation das arithmetische Mittel über eine große Anzahl an Versuchen repräsentativ für den Erwartungswert verwendet werden?*
- *Ist eine Reduzierung der Information einer Zeitreihe auf einen einzigen Parameter angemessen?*

1.5 Fragestellungen

Die sich aus Kapitel 1.2, 1.3 und 1.4 ergebenden Problematiken können in vier Bereiche aufgeteilt werden:

- die Verwendung von diskreten Parametern einer Zeitreihe gegenüber der gesamten Information einer Zeitreihe,
- die Überprüfung der Verwendung des Mittelwertes als standardisierte Größe für die Repräsentation des Probanden,
- die vollständige Charakterisierung der Variabilität eines Probanden anhand einer minimal zu messenden Anzahl an Läufen und
- die Entwicklung und Validierung geeigneter Algorithmen zur Berechnung dieser minimalen Anzahl.

Ausgehend von diesen Überlegungen werden folgende Hypothesen generiert.

Hypothese I: *Der Mittelwert von*

- (a) *diskreten Parametern einer Zeitreihe und/oder*
- (b) *einer gesamten Zeitreihe*

kann als ausreichende und stabile Charakterisierung des Probanden angesehen werden.

Hypothese II: *Die Variabilität eines Probanden kann mit einer endlichen Anzahl an Wiederholungen vollständig und stabil mittels*

- (a) *diskreten Parametern einer Zeitreihe und/oder*
- (b) *einer gesamten Zeitreihe*

charakterisiert werden.

Unter der Annahme, dass Hypothese I und/oder Hypothese II bestätigt werden können, ist die Entwicklung eines einheitlichen Algorithmus zur Bestimmung einer minimalen Anzahl zu messender Versuche sinnvoll:

Hypothese III: *Unter der Annahme der Bestätigung von Hypothese I und/oder Hypothese II kann die endliche Anzahl an Wiederholungen mit einem einheitlichen Algorithmus anhand*

- (a) diskreter Parameter einer Zeitreihe und/oder*
- (b) einer gesamten Zeitreihe*

bestimmt werden.

Kapitel 2

Methodische Aspekte

Für die Untersuchungen zur Charakteristik der Variabilität eines Probanden wurden zunächst zwei Studien durchgeführt, wobei die für diese Arbeit interessierenden Messungen ausschließlich im Labor statt fanden. Jeder Proband absolvierte zwei Messungen im Abstand einer Woche, wobei lediglich vollständige Datensätze zur Auswertung kamen. Innerhalb von Studie II erfolgte eine Feldmessung zwischen den beiden Labormessungen. Diese ist jedoch nicht Gegenstand der Auswertungen dieser Arbeit. Die Ethikkommission der Fakultät für Human- und Sozialwissenschaften der Technischen Universität Chemnitz begutachtete beide Studien und befand diese für ethisch unbedenklich.

2.1 Probandengut

Tabelle 2.1.1 liefert eine Zusammenfassung des Probandengutes beider Studien. Außer für die Anzahl der Probanden sind jeweils das arithmetische Mittel und die Standardabweichung für entsprechende Intervalle angegeben.

	Studie I	Studie II
Anzahl Probanden	14 (4 Frauen, 10 Männer)	20 (11 Frauen, 9 Männer)
Alter [Jahre]	29 ± 4	34 ± 11
Gewicht [kg]	74.4 ± 11.5	68.8 ± 9.0
Größe [cm]	178.9 ± 10.2	171.7 ± 8.8
Schuhgröße [EUR]	42 ± 2	41 ± 2
<i>Training pro Woche</i>		
Häufigkeit:	2 ± 1	3 ± 1
Kilometer pro Einheit:	10.1 ± 2.4	8.0 ± 1.9
Kilometer pro Woche:	19.3 ± 10.0	15.4 ± 16.6

Tabelle 2.1.1:

Probandengut von Studie I und II

Von allen Probanden haben lediglich zwei an beiden Studien teilgenommen. Das Probandengut beider Studien ist in etwa homogen, wobei die Personen aus Studie I (25 bis 38 Jahre) durchschnittlich jünger waren als von Studie II (21 bis 59 Jahre). Des Weiteren sind die durchschnittlich gelaufenen

Kilometer pro Woche bei Studie I größer als bei Untersuchung II. Dabei sei zu beachten, dass innerhalb der Studie II sechs untrainierte Probanden teilgenommen haben sowie sieben Probanden über Laborerfahrung verfügten. Als Ausschlusskriterium wurden lediglich momentane Verletzungen der unteren Extremität verzeichnet, wobei bei länger zurückliegenden Verletzungen Beschwerdefreiheit vorliegen musste.

2.2 Datenaufnahme

In Studie I trug jeder Proband seinen eigenen Laufschuh. Dies sollte den Tragekomfort des Schuhs erhöhen und eine mögliche Behinderung durch den Einsatz von Kabeln bei der Messung verhindern. Eine Verwendung von Messgeräten, deren Daten ohne an den Probanden anzubringende Kabel aufgezeichnet werden, wurde dabei erforderlich. Bei der Aufzeichnung von Bodenreaktionskräften ist diese Voraussetzung erfüllt, da die Kraftmessplatte (KISTLER 9287BA; 60x90 cm) im Labor-Fußboden integriert und extern mit einem Computer verbunden ist (siehe Abbildung 2.2.1). Es können die Bodenreaktionskräfte in allen drei Richtungen gemessen werden (F_z - vertikale Richtung, F_y - anterior-posterior Richtung, F_x - mediolaterale Richtung).

In Studie II wurden die Probanden mit einem extra präparierten Laufschuh in ihrer Schuhgröße ausgestattet. Ein im Schuh integrierter Beschleunigungsaufnehmer (ADXL78) zur Ermittlung des Fußaufsatzes und ein zugehöriges Gyrometer (Murata ENC-03R) zur Messung der Winkelgeschwindigkeit des Schuhs erlaubten die zusätzliche Messung von kinematischen Größen (siehe Abbildung 2.2.2). In dieser Ausfertigung standen mehrere Schuhgrößen zur Verfügung, so dass ein dem von Studie I ähnliches Spektrum an Schuhgrößen erreicht werden konnte. Alle drei verwendeten Messgeräte (Kraftmessplatte, Beschleunigungsaufnehmer und Gyrometer) zeichneten die Daten mit jeweils 1000 Hz auf.

Nach einer kurzen Einweisung in den Versuchsablauf und die verwendete Messtechnik unterzeichneten die Probanden eine Einverständniserklärung und ihre anthropometrischen Daten wurden erfasst. Danach erfolgte die eigentliche Messung.

Studie I Nach der Ermittlung des Körpergewichts, welche eine Normalisierung der Bodenreaktionskräfte ermöglichte, absolvierte jeder Proband zunächst Probeläufe, um den richtigen Abstand zur Kraftmessplatte zu finden und die vorgegebene Geschwindigkeit von 3.5 ± 0.1 m/s einzuhalten. Zwei Lichtschranken überprüften diese Bedingung und gaben im Fall der Einhaltung den gültigen Datensatz zur Speicherung frei. Ein Versuch galt dann als gültig, wenn die Kraftmessplatte mit dem rechten Fuß vollständig getroffen und mit der Ferse aufgesetzt wurde. Nach 100 gültigen Versuchen des rechten Fußes war die Messung beendet, wobei jeder Proband durchschnittlich 23 ± 13 Fehlmessungen durchlief. Den Testpersonen war freigestellt, ob sie die gesamte Zeit laufen oder nach jedem Versuch zurück gehen wollten. Die getroffene Entscheidung musste über die gesamte Studie beibehalten werden, dabei wählten sechs Probanden Variante eins (durchgehend laufen) und acht Probanden Variante zwei (zurück gehen).

Studie II Nach der Auswahl des Schuhs in der korrekten Schuhgröße wurde der Bauchgurt mit integriertem Aufzeichnungsgerät und Stromversorgung angelegt und Beschleunigungsaufnehmer und

**Abbildung 2.2.1:**

Linkes Bild: Labor mit Kraftmessplatte Studie I und II: (a) Startpunkt (Anlauf ≈ 5 m); (b) Kraftmessplatte

Abbildung 2.2.2:

Rechtes Bild: Messgeräte Studie II: (a) Gyrometer und Beschleunigungsaufnehmer; (b) Kabelverbindungen zwischen Aufnahmegerät und Messgeräten zur Stromversorgung; (c) Bauchgurt mit Aufnahmegerät

Gyrometer an den Schuh angebracht. Der Eingabe der bevorzugten Laufgeschwindigkeit und Einstellung in das zugehörige Datenerfassungsprogramm der Kraftmessplatte folgte die Ermittlung des Körpergewichts anhand der Kraftmessplatte, wobei alle Messgeräte bereits angelegt waren. Dies diente der Normalisierung der Bodenreaktionskräfte. Zudem erhielt der Proband die Möglichkeit, sich mit der Messtechnik vertraut zu machen. Mittels Drucktaster musste vor jeder Messung die Aufzeichnung gestartet und der Schritt notiert werden, welcher auf der Kraftmessplatte auftraf. Nach dem Lauf beendete die Testperson selbstständig die Datenaufzeichnung per Hand wiederum durch Betätigung des Tasters. Dies unterband die Option des durchgehenden Laufens (Variante eins aus Studie I), da vor allem beim Start eine präzise Ermittlung des Bodenkontaktes durch den Beschleunigungsaufnehmer ermöglicht werden musste. Die individuelle Laufgeschwindigkeit ($\pm 2\%$) wurde mittels Lichtschranken gemessen und durch das System überprüft. Im Mittel ist dabei eine geringere Laufgeschwindigkeit zu verzeichnen als in Studie I (3.0 ± 0.3 m/s). Der Aufzeichnung von 100 gültigen Messungen (Einhaltung des Abstandes zur Kraftmessplatte, Fersenlauf und Einhaltung der individuellen Geschwindigkeit) gingen diverse Probeläufe voraus. Die durchschnittliche Anzahl an Fehlmessungen 14 ± 8 fiel dabei geringer aus als in der ersten Studie.

Mögliche Gruppierungen Aus den Angaben zum Probandengut und der Datenaufnahme kristallisieren sich vier mögliche Gruppierungen heraus. Eine Möglichkeit ist eine Gruppierung hinsichtlich des Geschlechtes, da aufgrund unterschiedlicher Gewebestrukturen Frauen und Männer verschiedene Laufeigenschaften und damit verbundene unterschiedliche Bodenreaktionskraftkurven und Winkelgeschwindigkeiten aufweisen können. Eine Auswirkung auf den Mittelwert und die mi-

nimale Anzahl wäre daher denkbar. Eine weitere Gruppierungsmöglichkeit bieten die sechs Läufer aus Studie I, welche durchweg gelaufen sind. Diese verfügen möglicherweise über einen Vorteil, da sie nicht beständig für jede Messung Anlaufen und Stoppen mussten. Diese Form der Bewegung spiegelt eher die tatsächliche Laufbewegung im freien Feld wider und ein Einfluss auf den Mittelwert und die zu bestimmende minimale Anzahl wäre vorstellbar. Die dritte Gruppierung bezieht sich auf die sechs Probanden aus Studie II ohne sportliche Aktivität, welche gegenüber den Probanden mit einem wöchentlichen Laufpensum schnellere Ermüdungserscheinungen aufzeigen könnten. Dies müsste in den Mittelwerten und der minimalen Anzahl sichtbar werden. Die vierte Unterscheidung wird zwischen den Probanden aus Studie II, welche laborerfahren waren gegenüber den Probanden, die vorher noch nicht innerhalb eines Labores gemessen wurden, getroffen. Das neue Umfeld und eine gewisse Eingewöhnungszeit an das Equipment und die Messsituation ergeben unter Umständen unterschiedliche Kurven der Bodenreaktionskraft und der Winkelgeschwindigkeit und resultieren gegebenenfalls in einer höheren minimal zu messenden Anzahl.

2.3 Datenverarbeitung

2.3.1 Bodenreaktionskräfte

Die vertikalen Bodenreaktionskräfte F_z wurden bereits während der Datenaufnahme auf das Körpergewicht normiert (F_{znorm}). Zusätzlich mussten die Bodenreaktionskräfte in anterior-posterior Richtung (F_y) und mediolateraler Richtung (F_x) auf die Gewichtskraft $\vec{G} = m \cdot g$ des Probanden angepasst werden, wobei m das Körpergewicht und $g = 9.80665$ die Fallbeschleunigung bezeichnet:

$$F_{ynorm} = \frac{F_y}{\vec{G}} = \frac{F_y}{m \cdot g} = \frac{F_y}{m \cdot 9.80665}$$

und

$$F_{xnorm} = \frac{F_x}{\vec{G}} = \frac{F_x}{m \cdot g} = \frac{F_x}{m \cdot 9.80665}.$$

Im Folgenden werden stets die normierten Bodenreaktionskräfte verwendet und es wird vereinfachend F_z , F_y und F_x geschrieben. Nach O'Connor et al. (2007) erfolgte die Notation des Bodenkontaktes anhand der vertikalen Bodenreaktionskräfte ab drei zusammenhängenden Werten über 10 N bis zu Werten kleiner als 5 N. Zudem wurden die Daten mit einem Tiefpass Butterworth-Filter dritter Ordnung bei 100 Hz gefiltert, um hohe Frequenzen, welche auf ein zufälliges Rauschen V_{EB} (↗ Glg. 1.2.2) zurückzuführen sind, aus dem Signal zu löschen. Die Filterart und die Frequenz wurden in Übereinstimmung mit den in der Literatur angegebenen Filtern verwendet (↗ Tab. A.1), wobei die Untersuchung der Auswirkung des Filters auf die Schwankung der Daten nicht Gegenstand dieser Arbeit ist.

Die Interpolation der normierten Bodenreaktionskräfte auf 101 Datenpunkte ermöglichte zudem eine bessere Veranschaulichung und Weiterverwendung für statistische Analysen. Eine grafische Überprüfung der originären und interpolierten Kurven war dabei unumgänglich, um Fehler, wie in Kapitel 1.2 (↗ S. 10) beschrieben, zu vermeiden. Die bei einer Interpolation entstehenden marginalen Unterschiede werden im Folgenden als sehr gering mit Erwartungswert Null betrachtet.

Im Anschluss werden die berechneten diskreten Parameter der Bodenreaktionskräfte vorgestellt. Die Auswahl orientiert sich an den in der Literatur häufig vorkommenden Parametern.

Diskrete Parameter der vertikalen Bodenreaktionskraft F_z :

- F_{z_I} - bezeichnet das erste Maximum der vertikalen Bodenreaktionskraft nach dem Fußaufsatz (in der Literatur oftmals als Impact-Peak bezeichnet). Im Fall, dass kein erster Peak existiert, wird auf den minimalen Anstieg zwischen Fußaufsatz und globalem Maximum zurückgegriffen.
- t_{z_I} - stellt den Zeitpunkt des Auftretens von F_{z_I} dar.
- F_{z_A} - beschreibt das Maximum der vertikalen Bodenreaktionskraft nach dem ersten Peak (in der Literatur auch als Active-Peak benannt).
- t_{z_A} - bezeichnet den Zeitpunkt des Auftretens von F_{z_A} .
- FR_{z_I} - stellt die maximale Kraftanstiegsrate zwischen Bodenkontakt und erster Kraftspitze dar

$$FR_{z_I} = \max_{t_i} \frac{F_{z_{i+1}} - F_{z_i}}{t_{i+1} - t_i} \quad \text{mit } i = 1, \dots, z_I.$$

F_{z_I} und FR_{z_I} werden unter anderem zur Charakterisierung des Fußaufsatzes verwendet. In fast allen betrachteten Studien aus Kapitel 1.1 werden als beschreibende Parameter der vertikalen Bodenreaktionskraft F_{z_I} sowie FR_{z_I} zur Bewertung von unterschiedlichen Schuhbedingungen herangezogen. Neben diesen beiden Variablen wird zudem F_{z_A} verwendet, wobei dieser zweite Peak von der Geschwindigkeit und der Körpermasse des Läufers abhängt und weder vom verwendeten Schuhwerk noch vom Abrollvorgang beeinflusst werden sollte (Nigg, 1986). Die Verwendung dieses Parameters ermöglicht die Bewertung ohne Beeinflussung des in diesen Studien eingesetzten Schuhwerks.

In Abbildung 2.3.1 sind die verwendeten diskreten Parameter der vertikalen Bodenreaktionskraft an einem Beispiel verdeutlicht und in Abbildung 2.3.2 sind alle 100 aufgenommenen Kurven eines Probanden von F_z dargestellt.

Diskrete Parameter der Bodenreaktionskraft in anterior-posterior Richtung F_y und in mediolateraler Richtung F_x :

- F_{y_B} - bezeichnet die minimale Kraft während der Abbremsphase des Bodenkontaktes.
- t_{y_B} - stellt den Zeitpunkt des Auftretens von F_{y_B} dar.
- F_{y_A} - beschreibt die maximale Kraft während der Beschleunigungsphase des Bodenkontaktes.
- t_{y_A} - bezeichnet den Zeitpunkt des Auftretens von F_{y_A} .

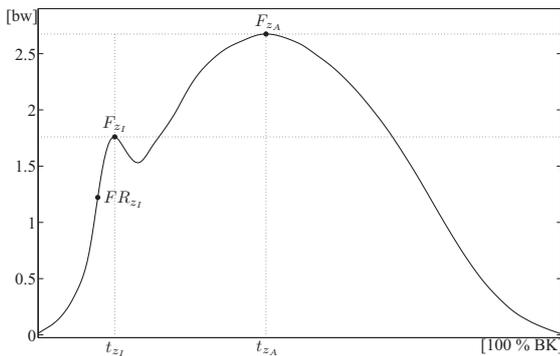


Abbildung 2.3.1:
Beispiel einer Kurve der vertikalen Bodenreaktionskraft F_z mit diskreten Parametern

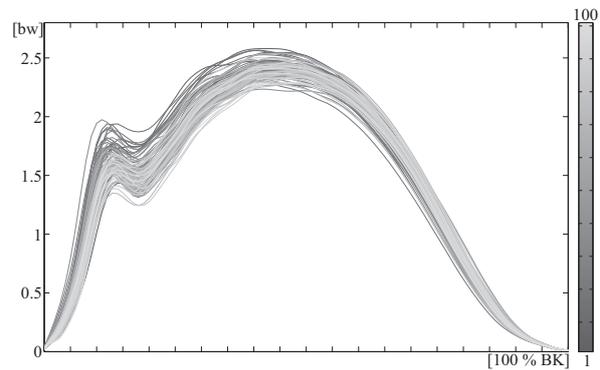


Abbildung 2.3.2:
100 aufgezeichnete Signale der vertikalen Bodenreaktionskraft F_z eines Probanden

FR_{y_B} - stellt die maximale Kraftanstiegsrate zwischen Bodenkontakt und t_{y_B} dar

$$FR_{y_B} = \max_{t_i} \frac{F_{y_{i+1}} - F_{y_i}}{t_{i+1} - t_i} \quad \text{mit } i = 1, \dots, y_B.$$

FR_{y_A} - beschreibt die maximale Kraftanstiegsrate zwischen t_{y_B} und t_{y_A}

$$FR_{y_A} = \max_{t_i} \frac{F_{y_{i+1}} - F_{y_i}}{t_{i+1} - t_i} \quad \text{mit } i = y_B, \dots, y_A.$$

Alle Parameter der Bodenreaktionskraft in anterior-posterior Richtung sind abhängig von der Geschwindigkeit des Läufers. Die Vereinheitlichung dieser sollte den Einfluss der Geschwindigkeit ausschließen (Cavanagh, 1990).

Abbildung 2.3.3 gibt einen Überblick über die diskreten Parameter der Bodenreaktionskraft in anterior-posterior Richtung und Abbildung 2.3.4 zeigt alle 100 Kurven von F_y eines Probanden.

Für die Bodenreaktionskraft in mediolateraler Richtung wird lediglich die gesamte Kurve betrachtet (Abbildung 2.3.5 und 2.3.6).

2.3.2 Winkelgeschwindigkeit

Die mit einem Gyrometer aufgenommene Winkelgeschwindigkeit ω spiegelt die Geschwindigkeit der Eversions- und Inversionsbewegung des Fußes, streng genommen des Schuhs, entlang der Sagittalachse wider. Durch Integration dieser Kurve erhält man das Bewegungsausmaß. Vor der Berechnung einzelner diskreter Parameter werden die Daten sowohl mit einem Hochpass Butterworth-Filter (erster Ordnung, 0.3 Hz) als auch mit einem Tiefpass Butterworth-Filter (dritter Ordnung, 100 Hz) bereinigt. Der Tiefpass-Filter dient der Eliminierung eines durch das Gyrometer verursachten Offsets von ungefähr 1.35 V und der Hochpass-Filter der Beseitigung der zufälligen Schwankung

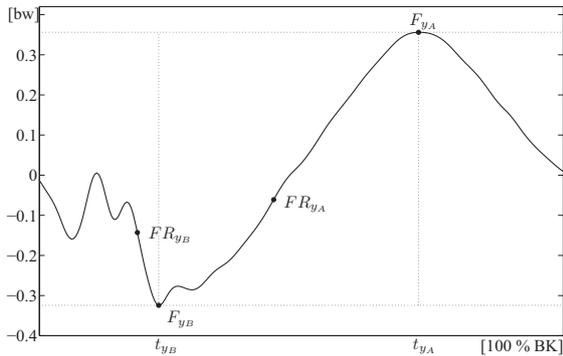


Abbildung 2.3.3: Beispiel einer Kurve der Bodenreaktionskraft in anterior-posterior Richtung F_y mit diskreten Parametern

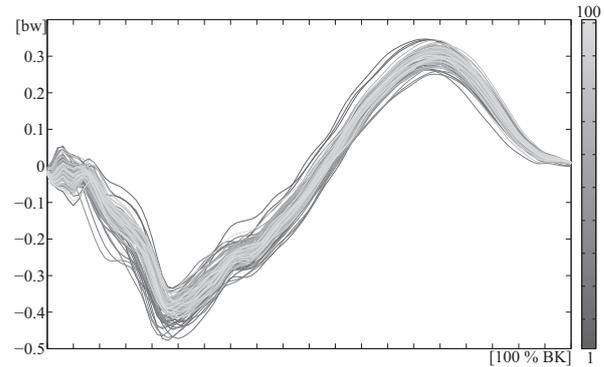


Abbildung 2.3.4: 100 aufgezeichnete Signale der Bodenreaktionskraft in anterior-posterior Richtung F_y eines Probanden

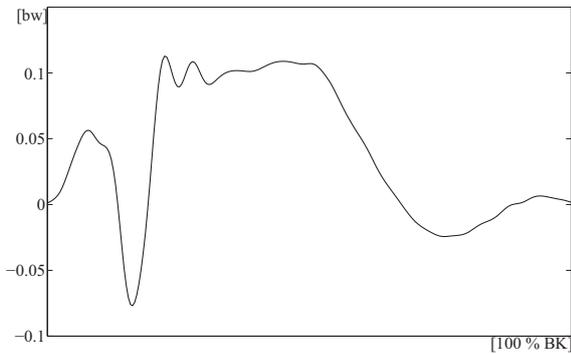


Abbildung 2.3.5: Beispiel einer Kurve der Bodenreaktionskraft in mediolateraler Richtung F_x

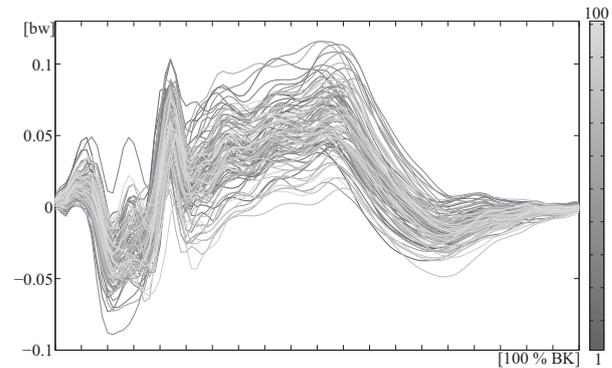


Abbildung 2.3.6: 100 aufgezeichnete Signale der Bodenreaktionskraft in mediolateraler Richtung F_x eines Probanden

innerhalb des aufgenommenen Signals.

Bei der Messung wurden fünf bis sieben Schritte eines Probanden gespeichert, wobei letztlich für die Auswertungen der Labor-Daten lediglich der Schritt von Bedeutung ist, welcher auf der Kraftmessplatte absolviert wurde. Die Schrittseparation erfolgte entsprechend der Schrittnotation und Bestimmung des Bodenkontaktes anhand des Peaks des Beschleunigungsaufnehmers. Dieser datiert gerade den Aufsatz des Fußes bei jedem Schritt (siehe Abbildung 2.3.7).

Zunächst wurde der entsprechende Schritt anhand der maximalen Beschleunigung ausgewählt, wobei in Abbildung 2.3.7 Beginn und Ende der Schrittphase durch das dritte und vierte Maximum separiert werden. Die Phase des Bodenkontaktes konnte schließlich mit Hilfe der Kraftmessplatte und der ermittelten Bodenkontaktzeit ausgewählt werden, wobei zum zeitlichen Auftreten des Peaks der Beschleunigung die Bodenkontaktzeit notiert wurde. Aufgrund der gleichen Aufnahme Frequenz konnte dies ohne Umstände realisiert werden.

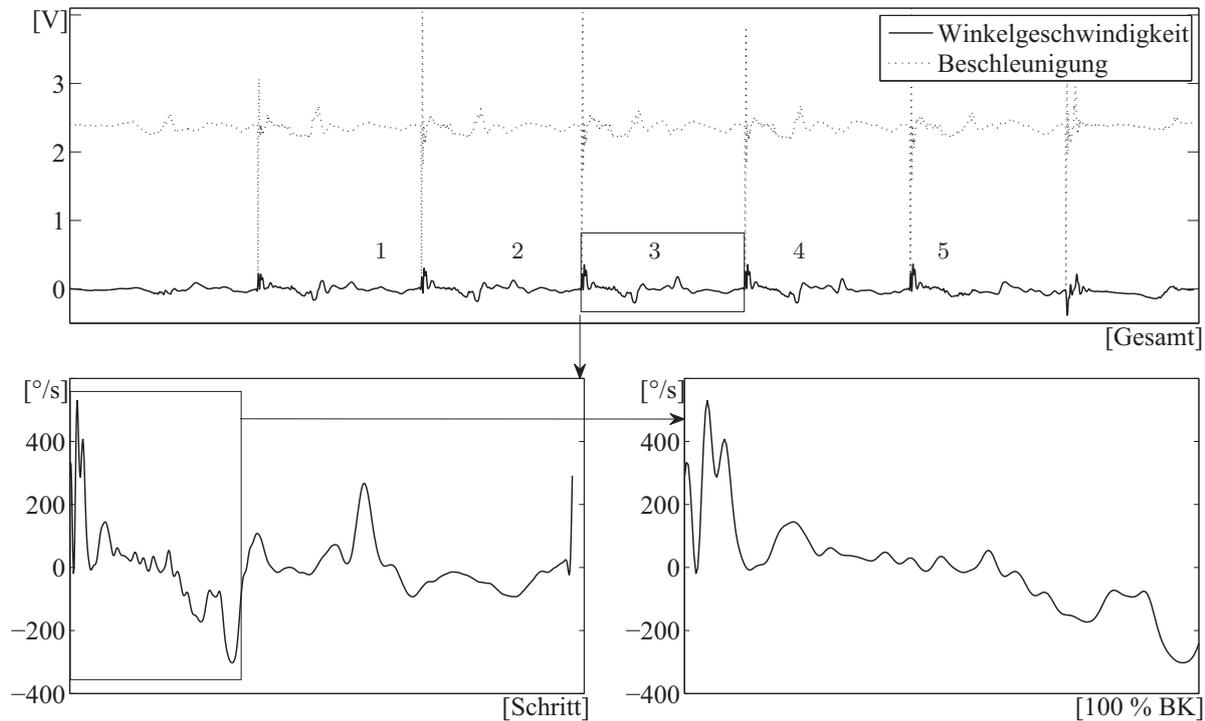


Abbildung 2.3.7:

Schrittseparation des Gyrometers bei Laboraufzeichnungen (Bild oben: Winkelgeschwindigkeit (über gesamten Zeitraum); Bild unten links: interessierende Schrittphase (Schritt drei); Bild unten rechts: Winkelgeschwindigkeit während des Bodenkontaktes)

Das in Volt gespeicherte Signal musste mit Hilfe der Umrechnung $0.67 \text{ mV} \cong 1^\circ/\text{s}$ angepasst und zur besseren Veranschaulichung auf 101 Datenpunkte interpoliert werden.

Diskrete Parameter der Winkelgeschwindigkeit ω und der integrierten Winkelgeschwindigkeit φ :

- ω_{max} - bezeichnet das Maximum der Kurve der Winkelgeschwindigkeit und
- $t_{\omega_{max}}$ - stellt den Zeitpunkt des Auftretens von ω_{max} dar.
- φ_{max} - beschreibt das Maximum der integrierten Winkelgeschwindigkeit und stellt das Bewegungsausmaß dar.
- $t_{\varphi_{max}}$ - bezeichnet den Zeitpunkt des Auftretens von φ_{max} .

In Abbildung 2.3.8 und 2.3.10 sind die entsprechenden Parameter für einen Probanden gekennzeichnet.

Darstellung 2.3.9 und 2.3.11 stellen jeweils die 100 aufgenommenen Kurven der Winkelgeschwindigkeit und der integrierten Winkelgeschwindigkeit dar. Dabei ist zu erkennen, dass die Kurven der integrierten Winkelgeschwindigkeit einen gemeinsamen Anfangspunkt aufweisen. Dies wird durch

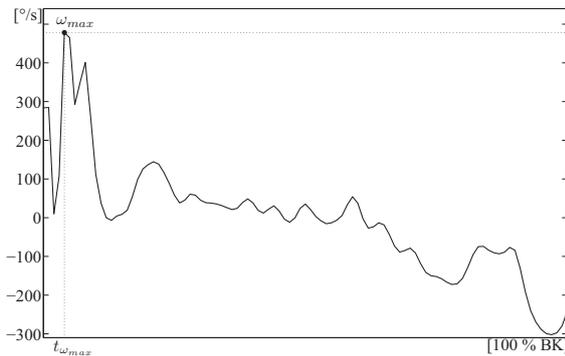


Abbildung 2.3.8:
Beispiel einer Kurve der Winkelgeschwindigkeit ω mit diskreten Parametern

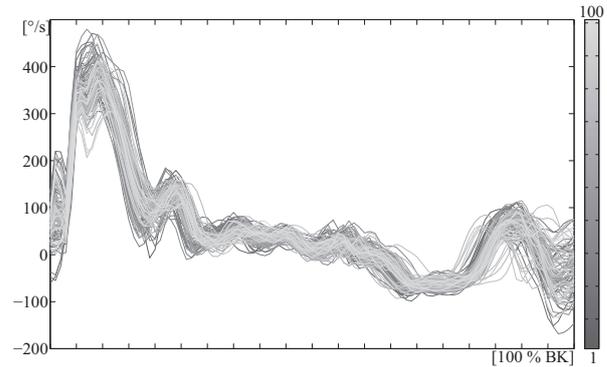


Abbildung 2.3.9:
100 aufgezeichnete Signale der Winkelgeschwindigkeit ω für einen Probanden

die Integration verursacht, da bei dieser das Absolutglied einer Funktion nicht wiederhergestellt werden kann. Daher wird für die Ergebnisse lediglich das Bewegungsausmaß φ_{max} dieser Kurven betrachtet.

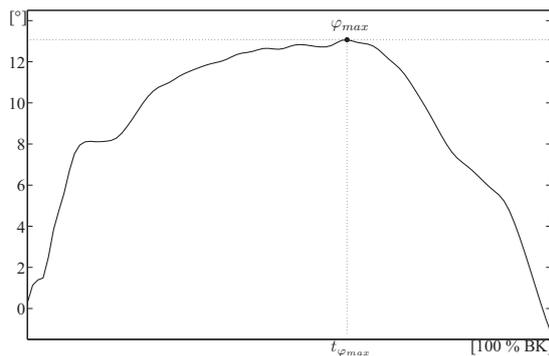


Abbildung 2.3.10:
Beispiel einer Kurve der integrierten Winkelgeschwindigkeit φ

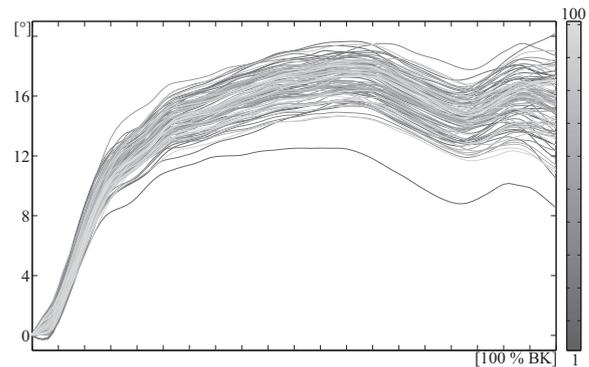


Abbildung 2.3.11:
100 aufgezeichnete Signale der integrierten Winkelgeschwindigkeit φ für einen Probanden

2.4 Statistische Auswertungen und Berechnungen (Hypothese I & II)

Ausgehend von der Definition des gemessenen Signals eines Probanden aus Gleichung 1.2.2

$$S_T = S_B + V_n + V_{EB} + V_{EM} + V_{EE}$$

bezieht sich Hypothese I auf das biologische Signal und somit den wahren Wert des Probanden S_B . Nach den Ausführungen aus Kapitel 1.4 kann dieser anhand des Mittelwertes geschätzt werden.

Die Auswertungen zu Hypothese I beziehen sich daher auf die Betrachtung von Mittelwerten für diskrete Parameter und Mittelwertkurven im Fall der Analyse gesamter Kurven. Die Überprüfung der Endlichkeit der Variabilität eines Probanden (Hypothese II) bezieht sich auf den variablen Faktor $V_T = V_n + V_{EB} + V_{EM} + V_{EE}$, wobei das zufällige Rauschen V_{EB} anhand des bei der Datenaufbereitung eingesetzten Filters beseitigt sein sollte. Der methodische und der sonstige Fehler $V_{EM} + V_{EE}$ werden als konstant angenommen, so dass sich die Überlegungen zu Hypothese II größtenteils auf die deterministische Schwankung V_n beziehen. Die Messung dieser Schwankung erfolgt innerhalb eines Probanden hauptsächlich über Standardabweichungen für diskrete Parameter, Standardabweichungskurven und Prädiktionsbänder für die gesamte Kurve. Die Variabilität zwischen den Probanden wird anhand des Root Mean Square Error (RMSE) betrachtet.

Die Messung von 100 Versuchen ergab 100 der entsprechenden Signale S_T . Die Betrachtung von diskreten Parametern reduziert die Information von S_T auf den jeweils verwendeten Wert, so dass sich die Auswertung von Hypothese I und II lediglich auf die Variabilität der Menge dieser diskreten Parameter bezieht. Die gesamten Kurven beinhalten zwei Variabilitätskomponenten - die Variabilität der einzelnen Signale und die Variabilität, welche punktwise durch die 100 gemessenen Signale entsteht. Ausschlaggebend für die Auswertungen zu Hypothese I und II ist die Variabilität, welche durch die 100 gemessenen Läufe verursacht wird.

Zunächst umfassen die Auswertungen die Betrachtung der Konvergenz von Folgen kumulierter Mittelwerte, Standardabweichungen und Root Mean Square Errors für diskrete Parameter $v_p = (x_1^{pr}, \dots, x_{100}^{pr})$ für jeden Proband r mit

$$p \in (F_{z_I}, F_{z_A}, FR_{z_I}, F_{y_B}, F_{y_A}, FR_{y_B}, FR_{y_A}, \omega_{max}, \varphi_{max})$$

und der gesamten aufgezeichneten Signale der Bodenreaktionskräfte und Winkelgeschwindigkeit $w_k = (x_{mi}^{kr}, \dots, x_{m100}^{kr})$ für alle Probanden r mit $m = 1, \dots, 101$ für

$$k \in (F_z, F_y, F_x, \omega).$$

Im Folgenden wird für die diskreten Parameter vereinfachend p und für die Gesamtkurven vereinfachend k geschrieben. Dabei bezeichnet p immer einen der Parameter $F_{z_I}, F_{z_A}, FR_{z_I}, F_{y_B}, F_{y_A}, FR_{y_B}, FR_{y_A}, \omega_{max}, \varphi_{max}$ und k eines der aufgezeichneten Signale F_z, F_y, F_x, ω .

2.4.1 Kumulation

Mittelwerte und Standardabweichungen Für die Bewertung der Konvergenz des Mittelwertes und der Standardabweichung von diskreten Parametern und Gesamtkurven werden diese zunächst kumuliert betrachtet, wobei eine Mindestanzahl von fünf Versuchen angenommen wird. Diese ergibt sich aus der Literaturrecherche aus Kapitel 1.3 (↗ Tab. 1.3.1 und A.1). Eine Versuchsanzahl von fünf stellte innerhalb dieser Studien die häufigste Anzahl an Wiederholungen dar. Daraus ergibt sich für jeden Proband r jeweils ein Vektor mit 96 Mittelwerten und 96 Standardabweichungen für jeden diskreten Parameter p

$$MW_n^{pr} = \frac{1}{n} \sum_{i=1}^n x_i^{pr} \quad \text{und} \quad SD_n^{pr} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i^{pr} - MW_n^{pr})^2}$$

für $n = 5, \dots, 100$ und jeweils eine Matrix mit 96 punktweise berechneten Mittelwertkurven und 96 punktweise berechneten Standardabweichungskurven für jedes aufgezeichnete Signal der Bodenreaktionskräfte und der Winkelgeschwindigkeit (k)

$$MW_{mn}^{kr} = \frac{1}{n} \sum_{i=1}^n x_{mi}^{kr} \quad \text{und} \quad SD_{mn}^{kr} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{mi}^{kr} - MW_{mn}^{kr})^2}$$

für $n = 5, \dots, 100$ und $m = 1, \dots, 101$.

RMSE Der Root Mean Square Error wird als Maß für die absolute Abweichung der entsprechenden Parameter p zwischen den Probanden herangezogen. Es wird die Folge kumulierter Abweichungen

$$RMSE_n^p = 1.96 \sqrt{2 \cdot MSE_n^p} \tag{2.4.1}$$

für $n = 5, \dots, 100$ betrachtet. MSE_n^p entspricht dem mittleren Fehlerterm einer Varianzanalyse mit n Messwiederholungen (Atkinson & Nevill, 1998). Zur Berechnung von MSE_n^p sei auf Sheskin (2007, S. 1024 ff.) verwiesen. Der Vorteil des RMSE besteht in einer Aussage über die Stabilität der Variabilität über alle Probanden und gibt weiterhin eine absolute Größe an, wie viel Schwankung das Probandenkollektiv produziert.

2.4.2 Konvergenz

Für die Beurteilung der Stabilisierung der Folge der kumulierten Mittelwerte, Standardabweichungen und RMSEs werden Aussagen über deren Konvergenz getroffen, wobei eine Folge reeller Zahlen (a_n) mit $n \in \mathbb{N}$ als konvergent bezeichnet wird, wenn eine Zahl $a \in \mathbb{R}$ existiert, für die gilt: Für alle $\varepsilon > 0$ gibt es eine natürliche Zahl N , so dass für $N \geq n$ gilt

$$|a_n - a| < \varepsilon. \tag{2.4.2}$$

Man sagt, dass die Folge a_n für $n \rightarrow \infty$ gegen den Grenzwert a konvergiert und schreibt $a = \lim_{n \rightarrow \infty} a_n$ (Hering, Martin & Stohrer, 2009). Falls ein solcher reeller Grenzwert nicht existiert, heißt eine Folge divergent. Die Ermittlung des Grenzwertes einer Folge ist dann möglich, wenn eine genaue Bildungsvorschrift für a_n vorliegt oder die Folge mittels Regression an eine Funktion angepasst werden kann. Im ersten Fall ist der Grenzwert exakt bestimmbar. Im Fall einer Regression wäre der Grenzwert der Funktion lediglich eine näherungsweise Lösung des Grenzwertes der vorliegenden Folge.

Nach der Theorie der großen Zahlen sollte sich der Mittelwert nach einer gewissen Anzahl an Versuchen einstellen. Die Berechnung des Grenzwertes der entstehenden Folge ist an dieser Stelle ungeeignet, da keine direkte Bildungsvorschrift vorliegt und das funktionale Modell der kumulierten Mittelwerte unbekannt und nicht einheitlich ist. Die Annahme der Konvergenz des Mittelwertes impliziert nicht, dass die Folge kumulierter Mittelwerte einem einheitlichen Verlauf und damit dem gleichen Modell folgt. Eine Vereinheitlichung gestaltet sich somit schwierig und es wird auf das Mittelwertkriterium zurückgegriffen, welches bei Hamill und McNiven (1990) und DeVita und

Bates (1988) vorgestellt wurde, um eine Aussage über die Anzahl an Versuchen zu erhalten, ab wann die Folge konvergiert. Dieses bezieht sich lediglich auf die Formulierung eines Intervalls, in welchem der Grenzwert der Folge der kumulierten Mittelwerte enthalten sein muss und setzt das Vorliegen von Konvergenz voraus.

Dies bedarf einer visuellen Überprüfung, inwiefern ein tatsächlicher Grenzwert erreicht wurde. Zur besseren Veranschaulichung und Vergleichbarkeit der Folgen kumulierter Mittelwerte für diskrete Parameter wird eine Z-Transformation

$$\frac{MW_n^{pr} - MW_{100}^{pr}}{SD_{100}^{pr}} \quad (2.4.3)$$

vorgenommen. Durch die Differenzenbildung orientieren sich die letzten Folgenglieder auf Null. Im Falle der Konvergenz sollten sich ab einer gewissen Anzahl die kumulierten Mittelwerte Null annähern. Für die gesamte Kurve k wird dabei eine punktweise Überprüfung für $m = 1, \dots, 101$ vorgenommen, wobei ebenfalls für die Mittelwertfolgen zur Vergleichbarkeit eine Normierung mittels z-Transformation

$$\frac{MW_{m,n}^{kr} - MW_{m,100}^{kr}}{SD_{m,100}^{kr}} \quad (2.4.4)$$

für jeden Punkt $m = 1, \dots, 101$ erfolgt. Diese entstehenden 100 punktwisen Folgen sollten um Null konvergieren unter der Annahme, dass der Mittelwert $MW_{m,100}^{kr}$ und damit die Mittelwertkurve aller 100 Läufe das Optimum für jede untersuchte Kurve k pro Proband r darstellt. Eine Normierung ist notwendig, da zu jedem Zeitpunkt die Kurven eine andere Ausprägung haben. In Abbildung 2.4.1 und 2.4.2 ist diese Prozedur beispielhaft für drei frei ausgewählte $m = 21, 46, 71$ dargestellt.

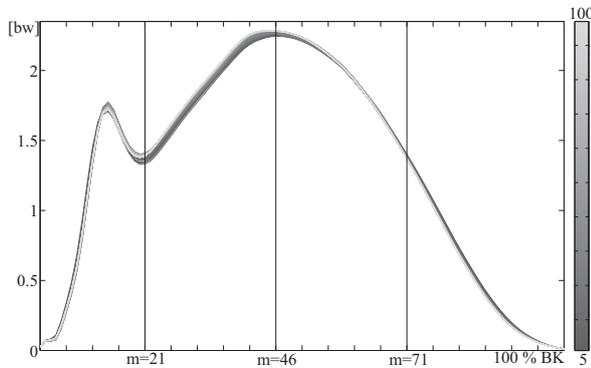


Abbildung 2.4.1:

Beispiel der kumulierten Mittelwertkurven der vertikalen Bodenreaktionskraft F_z mit $m = 21, 46, 71$

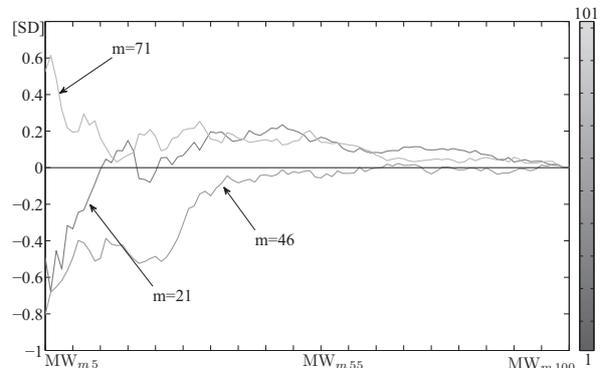


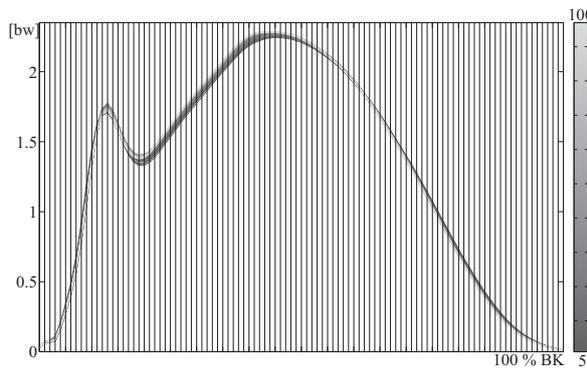
Abbildung 2.4.2:

Entsprechende punktweise Mittelwertkurven für $m = 21, 46, 71$ (z-transformiert)

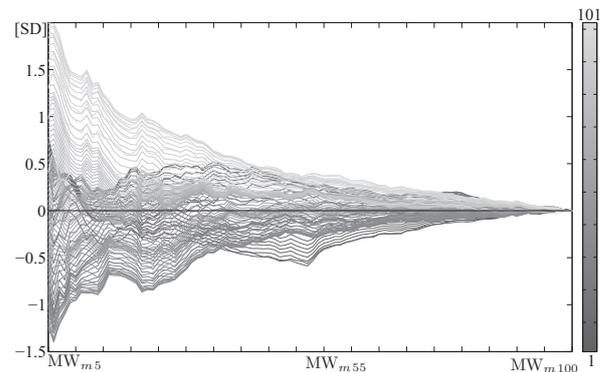
Angewendet auf alle 100 Datenpunkte m der entsprechenden Kurven ergibt sich das Bild aus den Darstellungen 2.4.3 und 2.4.4.

Für das Mittelwertkriterium wird zunächst das interessierende Intervall für die entsprechenden diskreten Parameter p und die gesamten Kurven k über alle Probanden r bestimmt

$$MW_{100}^{pr} \pm q \cdot SD_{100}^{pr} \quad \text{und} \quad MW_{m,100}^{kr} \pm q \cdot SD_{m,100}^{kr} \quad (2.4.5)$$


Abbildung 2.4.3:

Beispiel der kumulierten Mittelwertkurven der vertikalen Bodenreaktionskraft F_z mit $m = 1, \dots, 101$


Abbildung 2.4.4:

Entsprechende punktwise Mittelwertkurven für $m = 1, \dots, 101$ (z-transformiert)

für $q = 0.1, 0.25, 0.5$. Es werden somit 20 % bis 100 % der Standardabweichung (SD) der gesamten 100 Versuche als Intervalllänge akzeptiert. Die Stabilisierung des Mittelwertes ist dann erreicht, wenn ab einer bestimmten Anzahl alle kumulierten Mittelwerte oder Mittelwertkurven in dieses Intervall fallen. Für die Betrachtung der kumulierten Standardabweichungen und kumulierten RMSEs wird lediglich eine visuelle Überprüfung der Konvergenz durchgeführt, da zum einen für diese entstehenden Kurven kein einheitliches konvergentes funktionales Modell gefunden werden kann und somit keine Anpassung mittels nichtlinearer Regression möglich ist. Zum anderen ist das Mittelwertkriterium ungeeignet, da die Standardabweichung der kumulierten SD-Werte bestimmt werden müsste, welche jedoch keine unabhängigen Werte darstellen. Die Grundvoraussetzung zur Bildung eines erwartungstreuen Schätzers für die Standardabweichung wäre damit verletzt (Czado & Schmidt, 2011).

Für die visuelle Überprüfung der Standardabweichungskurven der gesamten Kurven k wird ähnlich wie für die Mittelwertkurven verfahren. Diese werden für jedes $m = 1, \dots, 101$ punktwise betrachtet und auf die Standardabweichungskurve SD_{m100}^{kr} normiert. Hierbei wird keine z-Transformation verwendet, sondern eine Art Variationskoeffizient berechnet, indem von den einzelnen Kurven die letzte Standardabweichung SD_{m100}^{kr} abgezogen und durch diese geteilt wird. Im Falle der Konvergenz sollten sich alle Kurven auf Null einpendeln.

2.4.3 Bootstrap-Verfahren zur Bildung von Prädiktionsbändern

Das Bootstrap-Verfahren basiert auf dem wiederholten Ziehen aus einer vorhandenen Stichprobe mit unbekannter Verteilung. Dieses Resampling ermöglicht die Ermittlung von Standardfehlern und Konfidenzintervallen/-bändern und Prädiktionsintervallen/-bändern ohne Kenntnis der Verteilung (Albers, Klapper & Konradt, 2007). Zur Schätzung von Prädiktionsbändern gesamter Kurven sei im Detail auf Lenhoff et al. (1999) verwiesen. Dabei werden die N vorliegenden Kurven $Y_i^{kr}(t_1), \dots, Y_i^{kr}(t_m)$ ($i = 1, \dots, n(=N)$) eines der Signale k für jeden Probanden r zunächst als

Fourier-Summe für $0 \leq t_1 \leq \dots \leq t_m \leq T$

$$f_i^{kr}(t) = \mu_i^{kr} + \sum_{l=1}^L \left(a_{il}^{kr} \cos\left(\frac{2\pi lt}{T}\right) + b_{il}^{kr} \sin\left(\frac{2\pi lt}{T}\right) \right)$$

geschrieben, wobei $\mu_i^{kr}, a_{i1}^{kr}, b_{i1}^{kr}, \dots, a_{iL}^{kr}, b_{iL}^{kr}$ unbekannte Koeffizienten der Fourier-Summe darstellen. Diese werden mit Hilfe der Methode der kleinsten Quadrate anhand $Y_i^{kr}(t) = f_i^{kr} + \varepsilon_i^{kr}$ ermittelt⁶ und es kann eine Schätzung für $f_i^{kr}(t)$ mit $\hat{f}_i^{kr}(t)$ für jede Kurve und die mittlere Kurve $\hat{f}_{\cdot N}^{kr}(t)$ angegeben werden. Ein Prädiktionsband gibt die Streuung der Signale derart an, dass bei wiederholter Messung des Probanden diese neue Kurve $f_{i+1}^{kr}(t)$ mit einer bestimmten Wahrscheinlichkeit in dieses Intervall fallen würde. Dies ergibt ein breiteres Band als bei Konfidenzintervallen, da diese lediglich mit einer bestimmten Wahrscheinlichkeit die Abweichung der vorhandenen Kurven vom wahren Wert des Probanden angeben. Zur Bestimmung des Prädiktionsbandes wird somit die Wahrscheinlichkeit bestimmt, dass

$$P \left\{ \max_t \left(\frac{|\hat{f}_{i+1}^{kr}(t) - \hat{f}_{\cdot N}^{kr}(t)|}{\hat{\sigma}_{\hat{f}_{\cdot N}^{kr}(t)}} \right) \leq C_N^{kr} \right\} = 1 - \alpha, \quad (2.4.6)$$

wobei α die vorgegebene Irrtumswahrscheinlichkeit darstellt und $\hat{\sigma}_{\hat{f}_{\cdot N}^{kr}(t)}$ die Variabilität von $\hat{f}_{\cdot N}^{kr}(t)$ (zur genauen Berechnung von $\hat{f}_{\cdot N}^{kr}(t)$ und $\hat{\sigma}_{\hat{f}_{\cdot N}^{kr}(t)}$ siehe Lenhoff et al. (1999)). Mit

$$\hat{f}_{\cdot N}^{kr}(t) \pm C_N^{kr} \cdot \hat{\sigma}_{\hat{f}_{\cdot N}^{kr}(t)} \quad (2.4.7)$$

ergibt sich das $(1 - \alpha)$ -Prädiktionsband für jeden Probanden. Das Intervall wird somit aus der mittleren Kurve, deren Variabilität und einem zu bestimmenden Quantil C_N^{kr} berechnet. In dieser Arbeit wird α auf 0.05 gesetzt. Mit $PB_{No}^{kr}(t)$ wird die obere Grenze und mit $PB_{Nu}^{kr}(t)$ die untere Grenze des Prädiktionsbandes für N Kurven bezeichnet.

An dieser Stelle wird das Bootstrap-Verfahren eingesetzt, um das unbekannte Quantil C_N^{kr} zu bestimmen. Die Verteilung der vorliegenden Kurven ist nicht ersichtlich und das Quantil C_N^{kr} folgt keiner gewöhnlichen Verteilung, so dass durch den Einsatz des Bootstrap-Verfahrens dieses Quantil geschätzt wird. Durch wiederholtes Ziehen mit Zurücklegen werden $1 \leq b \leq B$ Resampling-Stichproben der gleichen Stichprobengröße wie die Originalstichprobe gezogen und jeweils die mittlere Fourier-Kurve $\hat{f}_b^{kr}(t)$ und deren Variabilität $\hat{\sigma}_{\hat{f}_b^{kr}(t)}$ bestimmt. Dabei sei angemerkt, dass B möglichst groß sein sollte. Hier wurden $B = 500$ Wiederholungen gewählt. Anhand dieser neuen Stichproben wird die Wahrscheinlichkeit in Gleichung 2.4.6 näherungsweise durch

$$\frac{1}{B} \sum_{b=1}^B \left[\frac{1}{n} \sum_{i=1}^n \mathbf{I} \left\{ \max_t \left(\frac{|\hat{f}_i^{kr}(t) - \hat{f}_b^{kr}(t)|}{\hat{\sigma}_{\hat{f}_b^{kr}(t)}} \right) \leq C_N^{kr} \right\} \right], \quad (2.4.8)$$

bestimmt, wobei $\mathbf{I}()$ eine logische Matrix darstellt, deren Einträge 1 oder 0 ergeben, falls die Ungleichung zutrifft oder nicht. Das Quantil C_N^{kr} wird derart ermittelt, dass die Wahrscheinlichkeit

⁶ Der Fehleranteil ε_i^{kr} sollte für alle t_j mit $j = 1, \dots, m$ unkorreliert mit Mittelwert Null und Varianz σ^2 sein.

in Gleichung 2.4.8 gleich $1 - \alpha$ ist. Prädiktionsbänder für die vertikale Bodenreaktionskraft werden bei Null abgeschnitten, da negative Werte größer eines zufälligen Rauschens bei einem vertikalen Auftreffen auf die Kraftmessplatte nicht auftreten können.

Für die Auswertung bezüglich der Endlichkeit der Variabilität eines Probanden (Hypothese II) wird zusätzlich die kumulierte Fläche, welche das Prädiktionsband aus Gleichung 2.4.7 umschließt, berechnet. Dafür wird das Bootstrap-Verfahren mit $N = 5, \dots, 100$ Kurven wiederholt durchgeführt und jeweils die Fläche

$$APB_N^{kr} = \int_0^T \left| PB_{No}^{kr}(t) - PB_{Nu}^{kr}(t) \right| dt$$

ermittelt. Bei einer empirischen Ermittlung dieser Fläche werden die Differenzen punktweise aufsummiert:

$$APB_N^{kr} = \sum_{t=1}^m \left| PB_{No}^{kr}(t) - PB_{Nu}^{kr}(t) \right|. \tag{2.4.9}$$

Die Abbildungen 2.4.5 und 2.4.6 zeigen das entsprechende Prädiktionsband mit und ohne die jeweiligen 100 Kurven der vertikalen Bodenreaktionskraft für das Beispiel aus Abbildung 2.3.2.

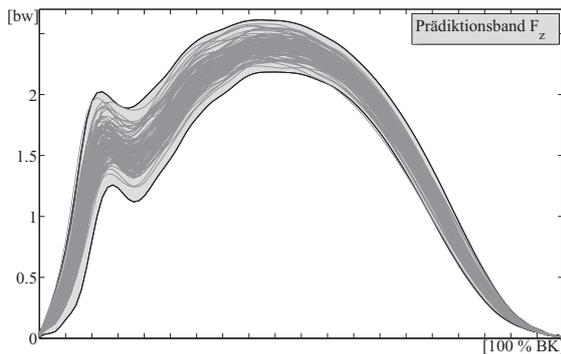


Abbildung 2.4.5:
Beispiel der 100 Versuche von F_z aus Abbildung 2.3.2 mit Prädiktionsband

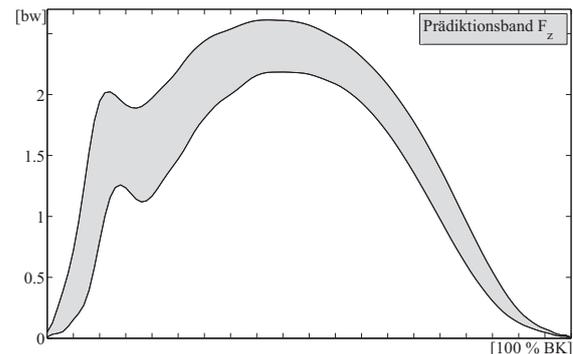


Abbildung 2.4.6:
Entsprechendes alleiniges Prädiktionsband zu Abbildung 2.4.5 ($APB_N^{kr} = 46$ (bw pro 100% Bodenkontakt))

2.4.4 Maßzahlen für die Reliabilität innerhalb und zwischen den Messtagen

Die Berechnungen zu Hypothese I und II umfassen zusätzlich prozentuale Abweichungen aller kumuliert berechneten statistischen Größen innerhalb eines Messtages für die ersten fünf Messungen und der Gesamtanzahl sowie zwischen Messtag 1 und Messtag 2. Dies dient der Analyse der Reliabilität, wobei innerhalb eines Messtages die entsprechenden Größen nach der herkömmlichen Messmethodik mit allen gemessenen Kurven verglichen werden. Die Reliabilitätsbetrachtung zwischen den Messtagen lässt Aussagen über die Genauigkeit der wiederholten Messungen zu.

Differenzen innerhalb eines Messtages Dabei fällt die Auswahl jeweils auf die ersten fünf Läufe und die gesamten 100 zur Verfügung stehenden Versuche. Die Anzahl fünf ergibt sich wiederum aus der Literaturrecherche, wobei eine Versuchsanzahl von fünf die häufigste Anzahl an Wiederholungen darstellte. Die Gesamtzahl wurde gewählt, da unter der Annahme der Konvergenz der Werte diese Anzahl das Optimum darstellt, sowohl für die Mittelwerte und Standardabweichungen als auch für die Gesamtkurven. Dies wird zusätzlich durch die Begrenzung des Datenvolumens von 100 Messversuchen oktroyiert, so dass intuitiv im Rahmen dieser Auswertung die letzte Mittelwertkurve zunächst als Optimum gelten muss.

Zunächst wird die prozentuale Abweichung D_{intra}^{pr} des Mittelwertes diskreter Parameter p der ersten fünf Läufe bezüglich aller 100 Läufe normiert auf deren Mittelwert betrachtet:

$$D_{intra}^{pr} = \frac{|MW_{100}^{pr} - MW_5^{pr}|}{(MW_{100}^{pr} + MW_5^{pr})/2} \cdot 100. \quad (2.4.10)$$

Hohe prozentuale Abweichungen deuten auf erhebliche Unterschiede innerhalb eines Tages. Die Normierung auf den Mittelwert aus beiden Differenzen wurde gewählt, um im Fall hoher Diskrepanzen in den entsprechenden Mittelwerten eine Über- oder Unterrepräsentation der einzelnen Werte zu vermeiden. Für die gesamten Signale k wird ein ähnliches Maß verwendet, wobei die prozentuale Abweichung der Mittelwertkurve der ersten fünf Läufe bezüglich aller 100 Läufe normiert auf die Mittelwertkurve aus diesen Werten berechnet wird. Dies entspricht der prozentualen Abweichung der absoluten Differenzenfläche der Mittelwertkurve der ersten fünf Läufe und allen Läufen normiert auf die Fläche der mittleren Kurve aus diesen beiden Mittelwertkurven:

$$DA_{intra}^{kr} = \frac{\sum_{m=1}^{101} |MW_{m100}^{kr} - MW_{m5}^{kr}|}{\sum_{m=1}^{101} (MW_{m100}^{kr} + MW_{m5}^{kr})/2} \cdot 100. \quad (2.4.11)$$

Für die Bewertung der Zuverlässigkeit der Schwankungen innerhalb der diskreten Parameter p über alle Probanden wird die prozentuale Abweichung des RMSE der ersten fünf Läufe bezüglich aller 100 Läufe normiert auf deren Mittelwert bestimmt.

$$DE_{intra}^p = \frac{|RMSE_{100}^p - RMSE_5^p|}{(RMSE_{100}^p + RMSE_5^p)/2} \cdot 100. \quad (2.4.12)$$

Die Beurteilung der Variabilität der gesamten betrachteten Kurven k wird anhand der Prädiktionsbänder vorgenommen. Hierfür wird die Differenzenfläche der nicht überlagerten Flächen der Prädiktionsbänder der ersten fünf Läufe und der gesamten 100 Läufe berechnet und dies auf die mittlere Fläche der beiden Prädiktionsbänder normiert:

$$DAPB_{intra}^{kr} = \frac{\sum_{m=1}^{101} (|PB_{100o}^{kr}(m) - PB_5o^{kr}(m)| + |PB_{100u}^{kr}(m) - PB_5u^{kr}(m)|)}{(APB_{100}^{kr} + APB_5^{kr})/2} \cdot 100$$

Je größer die nicht überlagerten Flächen der beiden involvierten Prädiktionsbänder sind und damit je mehr diese voneinander abweichen, um so ungenauer konnte der Proband anhand der 100 Kurven charakterisiert werden und die Messungen müssen als unreliabel gelten. Die prozentualen Angaben dienen dem Vergleich der Probanden untereinander.

Differenzen zwischen den Messtagen Für die Reliabilitätsbetrachtung zwischen den Messtagen werden jeweils die statistischen Größen für alle 100 aufgezeichneten Signale verwendet. Zunächst wird die Genauigkeit des Mittelwertes für diskrete Parameter p und der Mittelwertkurven für die betrachteten Kurven k überprüft, wobei jeweils die prozentuale Abweichung zwischen Messtag 1 und Messtag 2 berechnet wird. Diese Differenz wird normiert auf den Mittelwert aus diesen Werten:

$$D_{inter}^{pr} = \frac{\left| MW_{100T1}^{pr} - MW_{100T2}^{pr} \right|}{\left(MW_{100T1}^{pr} + MW_{100T2}^{pr} \right) / 2} \cdot 100$$

und

$$DA_{inter}^{kr} = \frac{\sum_{m=1}^{101} \left| MW_{m100T1}^{kr} - MW_{m100T2}^{kr} \right|}{\sum_{m=1}^{101} \left(MW_{m100T1}^{kr} + MW_{m100T2}^{kr} \right) / 2} \cdot 100.$$

Für die Auswertung anhand der Mittelwertkurven der gesamten Signale k entspricht DA_{inter}^{kr} der prozentualen Abweichung der absoluten Differenzenfläche der Mittelwertkurve aller 100 Läufe normiert auf die Fläche der mittleren Kurve aus diesen beiden involvierten Mittelwertkurven. Die Bewertung der Reliabilität diskreter Parameter wird zusätzlich durch die Betrachtung von Konfidenzintervallen unterstützt:

$$KI_{100}^{pr} = \left[MW_{100}^{pr} \pm t_{1-\alpha/2,99} \frac{SD_{100}^{pr}}{\sqrt{100}} \right], \quad (2.4.13)$$

wobei die Abweichung dieser Konfidenzintervalle betrachtet wird. Bei Überlappung der Konfidenzintervalle wird von einer reliablen Messung zwischen den Tagen ausgegangen. Im Fall, dass sich diese nicht überlappen, sind die Unterschiede zwischen den Messtagen mit einer Irrtumswahrscheinlichkeit von $\alpha = 0.05$ als signifikant und damit nicht reliabel anzusehen. Entsprechend wäre es nicht möglich den Mittelwert des Probanden an zwei unterschiedlichen Tagen genau zu messen. $t_{1-\alpha/2,99}$ entspricht dem $(1 - \alpha/2)$ -Quantil der t-Verteilung mit $(n - 1) = 99$ Freiheitsgraden.

Für die Bewertung der Variabilität innerhalb diskreter Parameter p eines Probanden wird die prozentuale Abweichung des RMSE aller 100 Läufe zwischen Messtag 1 und Messtag 2 normiert auf deren Mittelwert berechnet:

$$DE_{inter}^p = \frac{\left| RMSE_{100M1}^p - RMSE_{100M2}^p \right|}{\left(RMSE_{100M1}^p + RMSE_{100M2}^p \right) / 2} \cdot 100. \quad (2.4.14)$$

Für die gesamten Kurven werden entsprechend die nicht überlagerten Differenzenflächen der betrachteten Prädiktionsbänder für Messtag 1 und Messtag 2 auf die mittlere Fläche aus diesen Prädiktionsbändern betrachtet:

$$DAPB_{inter}^{kr} = \frac{\sum_{m=1}^{101} \left(\left| PB_{NoT1}^{kr}(m) - PB_{NoT2}^{kr}(m) \right| + \left| PB_{NuT1}^{kr}(m) - PB_{NuT2}^{kr}(m) \right| \right)}{(APB_{NT1}^{kr} + APB_{NT2}^{kr}) / 2} \cdot 100.$$

2.4.5 Auswertung der vier Gruppierungen

Im Fall, dass die einzelnen Analysen jeweils getrennt nach den vier Gruppierungen (Geschlecht, sportliche Aktivität, Ausdauer, Laborerfahrung) ausgeführt werden und metrisches Messniveau vorliegt, wird der Mann-Whitney Test durchgeführt, um mögliche signifikante Unterschiede zu eruieren. Die Irrtumswahrscheinlichkeit wird dabei auf $\alpha = 0.05$ gesetzt. Dieser Test wird durchgeführt, da aufgrund der geringen Stichprobengröße und der zum Teil wenigen Probanden innerhalb der Gruppen keine normalverteilten Daten zu erwarten sind.

Bei der Auswertung nominaler Daten und damit verbundener Häufigkeiten wird lediglich eine deskriptive Analyse vorgenommen und auf die Verwendung eines χ^2 -Tests verzichtet, da die kleine Anzahl innerhalb der jeweiligen Gruppierung auf geringe erwartete Häufigkeiten schließen lässt.

2.5 Statistische Auswertungen und Berechnungen (Hypothese III)

Die Analysen zu Hypothese III und die Bestimmung der minimal notwendigen Anzahl an Versuchen werden auf Basis der Auswertungen zu den vorangegangenen Hypothesen aufgestellt. Die Untersuchungen behandeln vorrangig die Flächen der Prädiktionsbänder APB_N^{kr} und die kumulierten Standardabweichungskurven SD_{mn}^{kr} der Bodenreaktionskräfte sowie der Winkelgeschwindigkeit.

2.5.1 Nichtlineare Regressionsmodelle für die Anpassung der Fläche der Prädiktionsbänder

Unter der Annahme der Endlichkeit der Variabilität (Hypothese II) und der damit verbundenen genauen Schätzung der Grenzen des Prädiktionsbandes der involvierten Kurven müsste die Fläche des Prädiktionsbandes bei kumulierter Berechnung stetig abnehmen und gegen eine bestimmte Fläche konvergieren. Für die Entwicklung der Fläche des Prädiktionsbandes APB_N^{kr} (\nearrow Glg. 2.4.9) und der normierten Folge $APBZ_N^{kr}$ wurden jeweils die aufgenommenen Versuche für $N = 5, \dots, 100$ verwendet, wobei die Normierung auf die letzte bestimmte Fläche APB_{100}^{kr} vorgenommen wurde. Im Fall einer stetigen Abnahme des Ausmaßes dieser Fläche sollte die kumulierte Kurve einen abfallenden exponentiellen Verlauf aufweisen, so dass eine Anpassung dieser entstehenden Folge an ein theoretisches Modell möglich ist. Im Folgenden werden zwei Modelle mit exponentiellem Verhalten vorgestellt, wobei eine Orientierung an den nichtlinearen Regressionsmodellen von Ratkowsky

(1989) erfolgt. Für die Ermittlung der minimal aufzuzeichnenden Anzahl an Versuchen wird der maximale Krümmungspunkt der mittels nichtlinearer Regression angepassten Funktion verwendet. Dieser wird nach Bronstein, Semendjajew, Musiol und Mühlig (2008) unter Zuhilfenahme der ersten und zweiten Ableitung der entsprechenden Funktion mit

$$K = \frac{y''(x)}{(1 + (y'(x))^2)^{\frac{3}{2}}}$$

bestimmt. Das erste Modell (**Modell 1**), ein zwei-Parameter-Modell, mit

$$f(x) = \frac{\alpha}{\beta^x} \tag{2.5.1}$$

folgt für $0 < \beta < 1$ einem exponentiell abfallenden Verlauf (Ratkowsky, 1989, S. 86 (4.2.9)). Die Funktion 2.5.1 kann mit Hilfe der Parametrisierung $\delta = \ln \alpha$ und $\gamma = \ln \beta$ zu folgendem Modell umgewandelt werden:

$$f(x) = e^{-\gamma x + \delta},$$

wobei $\gamma < 0$ (Ratkowsky, 1989, S. 86 (4.2.11)). Eine Fläche des Prädiktionsbandes von Null ist nicht zu erwarten, daher wird das parametrisierte Modell von Gleichung 2.5.1 um eine Verschiebung auf der y-Achse erweitert und man erhält Gleichung 2.5.2 mit den entsprechenden Ableitungen mit $\gamma > 0$:

$$f(x) = e^{(-\gamma x + \delta)} + \eta, \quad \text{mit} \quad f'(x) = -\gamma e^{(-\gamma x + \delta)}, \quad f''(x) = \gamma^2 e^{(-\gamma x + \delta)}. \tag{2.5.2}$$

Modell 2.5.2 hat nach Ratkowsky (1989) bessere statistische Eigenschaften als Modell 2.5.1 und ist jenem daher vorzuziehen. Diese beziehen sich vor allem auf die schnellere und fast sichere Konvergenz des nichtlinearen Regressionsverfahrens und auf die Kleinste-Quadrate-Schätzer, welche einer approximativen Normalverteilung folgen. Auf die Funktionsweise einer nichtlinearen Regression wird an dieser Stelle nicht eingegangen. Für eine genaue Beschreibung des Verfahrens sei auf Seber und Wild (2003) verwiesen.

Die Startwerte c und d für γ und δ können mittels linearer Regression mit $\ln y = -cx + d$ gewonnen werden. Im Fall des Erhaltes von $c < 0$ mittels linearer Regression wird $c = 1$ gesetzt.

Modell 2 zur Anpassung der kumulierten Fläche der Prädiktionsbänder besteht in

$$\alpha + \beta x^\gamma \quad \text{mit} \quad f'(x) = \beta \gamma x^{\gamma-1}, \quad f''(x) = \beta \gamma (\gamma - 1) x^{\gamma-2}, \tag{2.5.3}$$

wobei $\beta < 0$ und $0 < \gamma < 1$ (Ratkowsky, 1989, S.99 (4.3.12)). Die Startwerte a , b und c für α , β und γ der nichtlinearen Regression wurden auf $a = 1$, $b = -1$ und $c = 0.5$ gesetzt. Dabei sei angemerkt, dass für $\beta > 0$ und $\gamma < 0$ ein ähnliches exponentiell abnehmendes Verhalten erreicht wird. Diese Parameterkombination wurde von Ratkowsky (1989) nicht weiter verfolgt.

Die Verwendung der normierten Folge kumulierter Flächen $APBZ_N^{kr}$ impliziert, dass die Absolutglieder in den oben beschriebenen Modellen wegfallen, da die so konstruierten Folgen auf Null enden. Die Parameteranzahl verringert sich somit um 1.

Güte der Anpassung einer nichtlinearen Regression Die Verwendung des Bestimmtheitsmaßes R^2 einer linearen Regression als Maß für die Güte einer nichtlinearen Regression ist nicht zulässig, da es zwar berechnet, aber dessen Bedeutung in Bezug auf eine nichtlineare Regression nicht geklärt werden kann (Ratkowsky, 1989). Neben der Überprüfung der Unabhängigkeit, Übereinstimmung und Normalverteilung des Fehlerterms ε der Regression sollte die Fehlervarianz σ_ε^2 der Anpassung betrachtet werden. Im Vergleich der beiden obigen Modelle ist das Modell mit der kleinsten Fehlervarianz vorzuziehen. Die Verwendung der Funktion *nlinfit* in MATLAB stellt die Berechnung dieser Fehlervarianz innerhalb der Funktion zur Verfügung, wobei zunächst die Differenz $d_{N\text{fit}}^{kr}$ zwischen der vorliegenden Folge kumulierter Flächen des Prädiktionsbandes und der mittels Modell 1 und 2 ermittelten Funktion bestimmt wird. σ_ε^2 wird dann wie folgt bestimmt:

$$\sigma_\varepsilon^2 = \frac{\sum_{N=5}^{100} \left(|d_{N\text{fit}}^{kr}| \right)^2}{N - P},$$

wobei P der Anzahl an Parametern entspricht. Des Weiteren wird die Wurzel aus der Fehlervarianz gezogen, um eine Zurückführung auf die tatsächliche Ausprägung der Variablen zu erhalten.

Prozentuale Abweichung der Prädiktionsbänder Des Weiteren werden die sich nicht überlappenden Flächen zweier aufeinander folgender Prädiktionsbänder bestimmt und die prozentuale Abweichung ermittelt:

$$DAPB_N^{kr} = \frac{\sum_{m=1}^{101} \left(\left| PB_{(N+1)o}^{kr}(m) - PB_{No}^{kr}(m) \right| + \left| PB_{(N+1)u}^{kr}(m) - PB_{Nu}^{kr}(m) \right| \right)}{\left(APB_{N+1}^{kr} + APB_N^{kr} \right) / 2} \cdot 100$$

für $N = 5, \dots, 100$. Es entsteht so eine Folge mit 96 prozentualen Abweichungen, wobei sich die Prädiktionsbänder ab einer bestimmten Anzahl nicht mehr groß unterscheiden und die Differenzen gegen Null tendieren sollten. Die minimale Anzahl wird anhand verschiedener Schwellenwerte und einer visuellen Überprüfung bestimmt.

2.5.2 Standardabweichungskurven

Ein weiterer Ansatz zur Bestimmung der minimalen Anzahl an aufzuzeichnenden Versuchen liegt in der Verwendung von punktweise kumulierten Standardabweichungskurven SD_{mn}^{kr} . Unter der Annahme, dass die Variabilität eines Probanden endlich ist (Hypothese II), sollten die Standardabweichungskurven ab einer gewissen Anzahl annähernd gleich werden. Für die Analyse werden zunächst die absoluten Differenzkurven zweier aufeinander folgender Standardabweichungskurven bestimmt. Man erhält so eine Folge von 95 Kurven. Die Differenzkurven werden mit DSD_{mn}^{kr} bezeichnet und anhand

$$DSD_{mn}^{kr} = \left| SD_{m(n+1)}^{kr} - SD_{mn}^{kr} \right|$$

für $m = 1, \dots, 100$ und $n = 5, \dots, 99$ berechnet. Der folgende Ansatz beruht auf der Berechnung einer Populationsdifferenzenkurve $PDSD_m^k$, um einen Algorithmus zu generieren, der während einer Messung verwendet werden kann. Dafür muss die Annahme gelten, dass die hier untersuchte Stichprobe eine Grundgesamtheit repräsentiert und dass generell eine derartige Populationsdifferenzenkurve existiert. Für die Bestimmung dieser repräsentierenden Differenzenkurve wird zunächst für jeden Proband die maximale Ausprägung aller Differenzenkurven DSD_{mn}^{kr} ermittelt

$$MDSD_{mn}^{kr} = \max_m DSD_{mn}^{kr}$$

für $m = 1, \dots, 100$ und $n = 5, \dots, 99$. Aufgrund von zwei verschiedenen Messungen für jeden Proband werden diese in verschiedene Gruppen randomisiert verteilt, so dass jeweils nur eine Messung von jedem Proband verwendet wird. Aufgrund dieser Randomisierung entstehen zwei Stichproben, wobei die Auswertungen für beide Stichproben durchgeführt werden. Eine doppelte Kombination in der Randomisierung ist aufgrund der vielen Möglichkeiten (2^n) und 100 Wiederholungen eher unwahrscheinlich:

$$p = \binom{100}{2} \left(\frac{1}{2^n}\right)^2 \left(1 - \frac{1}{2^n}\right)^{98} \approx 0$$

für $n = 14, 20$. Für diese Stichprobe werden alle möglichen zu bildenden absoluten Differenzenkurven gebildet. Dies ergibt für 14 Probanden $91 = (13 + \dots + 2 + 1)$ Differenzen und für 20 Probanden $190 = (20 + \dots + 2 + 1)$ Differenzen. Von diesen 91/190 Differenzenkurven wird jeweils die mittlere Kurve gebildet und diese Prozedur 100-mal wiederholt. Mit dieser entstehenden Kurvenschar wird das in Kapitel 2.4.3 beschriebene Bootstrap-Verfahren mit $B = 500$ angewendet und ein Prädiktionsband für die entstehende Populationsdifferenzenkurve ermittelt.

Für die Bestimmung der minimalen Anzahl werden für jeden Probanden die Differenzenkurven der Standardabweichung der untersuchten Bodenreaktionskräfte und der Winkelgeschwindigkeit mit diesem Prädiktionsband verglichen. Als Populationsdifferenzenkurve $PDSD_m^k$ kommen alle Quantile des Prädiktionsbandes in Frage. Zunächst werden die obere und untere Grenze sowie der Mittelwert des Prädiktionsbandes ausgewählt. Die Anzahl an Kurven wird als minimale Anzahl notiert, ab der alle Differenzenkurven kleiner als die verwendete Populationsdifferenzenkurve sind.

Unter der Annahme der Existenz einer globalen Populationsdifferenzenkurve wird obige Prozedur für alle zur Verfügung stehenden Probanden aus Studie I und II ausgeführt und mit $PDSD_{mIII}^k$ bezeichnet. Die zwei Probanden, welche an beiden Studien teilgenommen haben, werden Studie I zugeordnet. Aus Validierungsgründen wird zudem die minimal zu messende Anzahl anhand der jeweils entgegengesetzten Populationsdifferenzenkurve bestimmt. Für diese Kreuzvalidierung wird für Studie I die Populationsdifferenzenkurve von Studie II $PDSD_{mII}^k$ verwendet und für Studie II entsprechend die Populationsdifferenzenkurve von Studie I $PDSD_{mI}^k$.

Kapitel 3

Resultate zu Hypothese I: Konvergenz des Mittelwertes

Nach den Überlegungen aus Kapitel 1.4 kann der Mittelwert als Repräsentant des Probanden verwendet werden. Dieser sollte sich nach einer genügend großen Anzahl einstellen. Die Daten in diesem Kapitel stellen die Ergebnisse bezüglich der aufgeworfenen Frage der tatsächlichen, ausreichenden Charakterisierung des Probanden anhand des Mittelwertes dar. Die Auswertungen unterteilen sich bezüglich der Betrachtung der diskreten Parameter p mit $p \in (F_{z_I}, F_{z_A}, FR_{z_I}, F_{y_B}, F_{y_A}, FR_{y_B}, FR_{y_A}, \omega_{max}, \varphi_{max})$ und der Betrachtung der gesamten Kurven der Bodenreaktionskräfte F_z, F_y, F_x , und der Winkelgeschwindigkeit ω .

3.1 Betrachtung diskreter Parameter

Zunächst erfolgt die Betrachtung der Differenzen D_{intra}^{pr} innerhalb eines Messtages und D_{inter}^{pr} zwischen den zwei Messterminen. Dies dient der Gewinnung eines Überblicks, wie groß die tatsächliche Abweichung innerhalb eines Probanden ausfällt und um wie viel die Ausprägung des Parameters mit fünf Werten den angenommenen optimalen Mittelwert aller 100 Werte über- oder unterschätzt.

Für diese absolute Abweichung innerhalb eines Tages werden die Mittelwerte aus den ersten fünf Läufen mit allen 100 Läufen verglichen. Letztlich ist die Differenz abhängig von der tatsächlichen Ausprägung der Werte und wird deshalb ins Verhältnis zum Mittelwert gesetzt (\nearrow Glg. 2.4.10). In Tabelle 3.1.1 ist die maximale prozentuale Differenz sowie die mittlere prozentuale Abweichung angegeben. Unter der Annahme, dass der Mittelwert aus 100 Versuchen das Optimum darstellt, gibt diese prozentuale Abweichung eine Vorstellung wie reliabel der entsprechende Parameter innerhalb eines Probanden ist.

Es ist erkennbar, dass zum Teil erhebliche Diskrepanzen innerhalb eines Tages auftreten können. Die maximalen prozentualen Abweichungen machen dies deutlich. Bis zu 33 % Unterschied ist für die maximale Kraftanstiegsrate zu verzeichnen. Die zweite Kraftspitze F_{z_A} hingegen ist relativ stabil innerhalb eines Probanden messbar. Innerhalb der drei Gruppierungen ergibt sich ein ähnliches Bild wie in Tabelle 3.1.1. Es sind keine erheblichen Diskrepanzen oder anderweitige Systematiken innerhalb des Geschlechtes, der Laborerfahrung, der Gruppe ohne sportliche Aktivität und den Ausdauerläufern zu verzeichnen. Lediglich vereinzelt treten Signifikanzen vor allem für die Parameter der Bodenreaktionskraft in anterior-posterior Richtung auf, wobei die prozentualen signifikanten Unterschiede mit rund 5 % als relativ unbedeutend erscheinen. Zudem treten diese Unterschiede

p	Studie I		Studie II	
	Messtag 1 (%)	Messtag 2 (%)	Messtag 1 (%)	Messtag 2 (%)
F_{zI}	14 (5 ± 4)	10 (5 ± 4)	8 (4 ± 2)	12 (4 ± 3)
F_{zA}	5 (2 ± 2)	5 (2 ± 1)	10 (3 ± 2)	3 (1 ± 1)
FR_{zI}	33 (7 ± 8)	29 (9 ± 9)	15 (7 ± 5)	25 (6 ± 6)
F_{yB}	-17 (- 7 ± 5)	-12 (- 5 ± 4)	-12 (-6 ± 4)	-21 (- 5 ± 5)
F_{yA}	19 (6 ± 5)	14 (5 ± 4)	16 (5 ± 4)	11 (4 ± 3)
FR_{yB}	64 (13 ± 16)	32 (12 ± 10)	37 (8 ± 8)	17 (8 ± 4)
FR_{yA}	33 (8 ± 9)	25 (8 ± 7)	34 (8 ± 9)	16 (5 ± 5)
ω_{max}			18 (7 ± 6)	26 (10 ± 8)
φ_{max}			23 (8 ± 6)	15 (9 ± 4)

Tabelle 3.1.1:

Differenz D_{intra}^{pr} diskreter Parameter: maximale prozentuale Abweichung innerhalb aller Probanden (mittlere Differenz ± mittlere Standardabweichung der Differenz über alle Probanden)

nicht konsistent über alle Parameter, beide Studien und alle Messtage auf.⁷

Die Differenzen an unterschiedlichen Messtagen zeigen ein ähnliches Bild (Tab. 3.1.2 und Tab. B.5, S. 116). Dabei sind die Kraftanstiegsraten und die Parameter der Winkelgeschwindigkeit am unreliaabelsten. Die maximale Abweichung der maximalen vertikalen Kraftanstiegsrate beträgt bis zu 38 %. Dies ermöglicht es nicht, an zwei Messtagen den Probanden stabil zu messen. Die maximale Winkelgeschwindigkeit weicht bis zu 36 % ab.

p	Studie I (%)	Studie II (%)
F_{zI}	12 (5 ± 4)	10 (3 ± 3)
F_{zA}	5 (2 ± 2)	6 (2 ± 2)
FR_{zI}	38 (11 ± 13)	34 (7 ± 8)
F_{yB}	-11 (- 5 ± 3)	-22 (-5 ± 6)
F_{yA}	19 (5 ± 5)	16 (6 ± 4)
FR_{yB}	21 (9 ± 6)	20 (9 ± 6)
FR_{yA}	20 (6 ± 6)	37 (8 ± 9)
ω_{max}		36 (9 ± 8)
φ_{max}		21 (9 ± 5)

Tabelle 3.1.2:

Differenz D_{inter}^{pr} diskreter Parameter: maximale prozentuale Abweichung (mittlere prozentuale Abweichung ± Standardabweichung)

Zwischen den Messtagen lassen sich keine Systematiken für die vier Gruppierungen innerhalb der diskreten Parameter feststellen. Frauen haben zwischen den Messtagen eine leicht höhere prozentua-

⁷ Studie I (M2): F_{zI} 4% höhere prozentuale Abweichung für Frauen; Studie II (M1): F_{yB} 1.5% höhere prozentuale Abweichung für Frauen, FR_{yA} 7% höhere prozentuale Abweichung für laborunerfahrene Läufer.

le Abweichung der maximalen Kraftanstiegsrate für beide Studien zu verzeichnen, jedoch ist dieser Unterschied nicht signifikant. Innerhalb eines Messtages lässt sich keine derartige Tendenz feststellen. Für die anderen Parameter waren für keine der Gruppierungen durchweg kleinere oder höhere prozentuale Abweichungen zu verzeichnen. Diese gleichen in etwa denen der gesamten Gruppe.

Des Weiteren sind die Unterschiede der Differenzen innerhalb eines Messtages und zwischen zwei Messtagen weitgehend ausgeglichen, wobei die maximalen prozentualen Abweichungen zwischen den Messtagen etwas größer sind als innerhalb eines Tages (vor allem in Studie II).

Neben der tatsächlichen absoluten Abweichung sowohl innerhalb eines Messtages als auch zwischen den Messtagen ist zudem interessant, ob und ab welcher Anzahl die kumulierten Mittelwertfolgen konvergieren. Diesbezüglich erfolgt eine visuelle Überprüfung der Konvergenz der kumulierten Mittelwertkurve MW_n^{pr} der diskreten Parameter p für jeden einzelnen Probanden r . In Tabelle B.1 (↗ S. 113) ist für jeden dieser Parameter die Häufigkeit an Probanden angegeben, für die eine konvergente Mittelwertkurve festgestellt werden konnte. Für diese Fälle wurde zudem die Anzahl der Messungen notiert, die notwendig waren, bis eine Stabilisierung erkennbar war. Die kumulierten Mittelwertkurven konvergierten in ungefähr 50 % der Fälle und es waren für diese Fälle im Mittel rund 60 bis 70 Messungen pro Proband nötig, um diese Stabilisierung des Mittelwertes zu erreichen (Tabelle 3.1.3). MW_n^{pr} des diskreten Parameters für die maximale Kraft in der Beschleunigungsphase F_{yA} konvergierte in nur 4 von 20 Fällen innerhalb Studie II.

	Studie I		Studie II	
	Messtag 1 (#)	Messtag 2 (#)	Messtag 1 (#)	Messtag 2 (#)
Konvergenz	6 (45 %)	8 (57 %)	9 (46 %)	11 (55 %)
	65 ± 16	62 ± 17	69 ± 13	61 ± 19

Tabelle 3.1.3:

Zusammenfassung der Konvergenz des Mittelwertes: mittlere Anzahl konvergierender Mittelwerte (absolut und prozentualer Anteil) / mittlere Anzahl ± Standardabweichung

Die hohen Standardabweichungen verdeutlichen die erhebliche Streuung innerhalb der stabilisierenden Anzahl an Messungen. Einige Probanden benötigten lediglich 30 Versuche, andere wiederum mindestens 90 Versuche, um Konvergenz zu erreichen. Dabei konnten keine Systematiken festgestellt werden. Weder eine einheitliche Divergenz innerhalb eines Probanden unter allen Parametern noch eine einheitliche stabilisierende Anzahl an Messungen konnte erzielt werden. Lediglich ein Proband aus Studie II wies durchweg konvergierende Mittelwerte auf (nur Messtag 2), wobei für diese Testperson am ersten Messtag sechs von neun Parametern divergierten. An Studie II nahmen sechs Probanden teil, welche keiner sportlichen Aktivität nachgingen. Deren nicht vorhandenes Laufpensum und damit verbundene eventuell auftretende Ermüdungszustände ließen die Vermutung zu, dass deren Mittelwertkurven divergieren könnten. Ein derartiger Trend konnte nicht festgestellt werden. Die Einzelwerte bestätigen dies, wobei innerhalb der einzelnen diskreten Werte kein durchweg einheitlicher Trend ersichtlich war. Die Probanden weisen keine signifikanten Unterschiede für jeden diskreten Parameter auf. Jedoch benötigen diese Probanden im Schnitt fünf bis 10 Versuche zusätzlich, um Konvergenz zu erreichen.

Ein Vergleich der stabilisierenden Anzahl zwischen den Geschlechtern zeigt, dass lediglich ein Pa-

parameter FR_{y_B} (Studie II, Messtag 1) signifikant unterschiedliche Werte mit einer Differenz von 20 aufweist. Für alle anderen diskreten Parameter für beide Studien wurden nicht signifikante Unterschiede erzielt. Für die Anzahl divergenter Kurven konnte keine Abhängigkeit vom Geschlecht festgestellt werden. Auch die Gruppe der Probanden von Studie I, welche durchgängig ohne Unterbrechung im Labor gelaufen sind, zeigen keine Systematiken.

Die Analyse der Gruppe mit Laborerfahrung zeigt für einige Parameter (F_{z_I} (M2), FR_{z_I} (M1/M2), FR_{y_B} (M2)) signifikante Unterschiede, wobei jeweils die Probanden, welche über Laborerfahrung verfügten, rund 20 Läufe mehr benötigten, um Konvergenz zu erreichen. Dabei sei angemerkt, dass aufgrund der relativ geringen Anzahl konvergenter Mittelwertfolgen zum Teil sehr wenig Probanden in diese Berechnung eingehen.

Die Abbildungen 3.1.1 und 3.1.2 zeigen zwei Beispiele für konvergierende kumulierte Mittelwertkurven der maximalen Kraftanstiegsrate der vertikalen Bodenreaktionskraft. Nach visueller Einschätzung konvergieren die Kurven ab 70 bzw. ab 55 Werten.

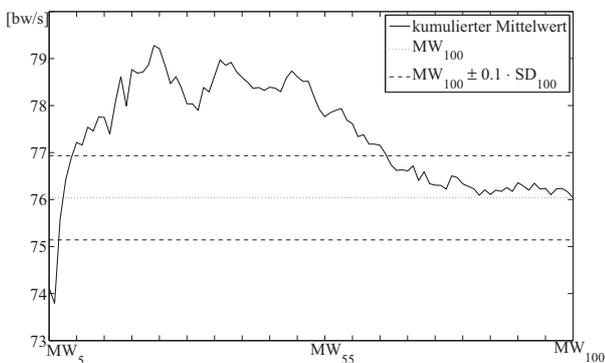


Abbildung 3.1.1:
Beispiel eines konvergierenden Mittelwertes (Stabilisierung von FR_{z_I} ab 70 Werten)

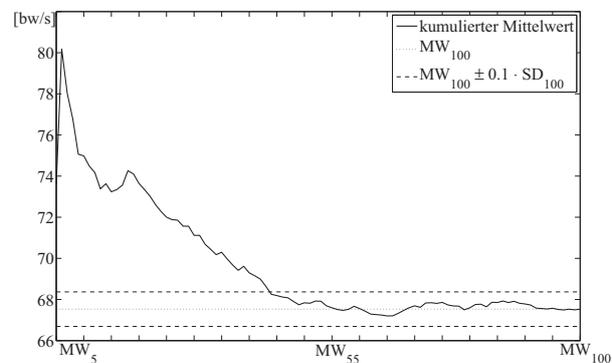


Abbildung 3.1.2:
Beispiel eines konvergierenden Mittelwertes (Stabilisierung von FR_{z_I} ab 55 Werten)

In den Abbildungen 3.1.3 und 3.1.4 sind divergierende kumulierte Mittelwertkurven des Impact-Peaks der vertikalen Bodenreaktionskraft dargestellt.

Bei Betrachtung der Differenzen innerhalb eines Messtages lediglich für die Probanden, welche eine visuelle Konvergenz aufwiesen, ist im Mittel keine wesentliche Änderung der Differenzen gegenüber denen für alle Probanden zu verzeichnen (↗ Tab. 3.1.1). Die maximalen Abweichungen innerhalb eines Messtages fallen für konvergierende Mittelwertkurven geringer aus (↗ Tab. B.4, S. 115). Auf eine zusätzliche Betrachtung der Differenzen zwischen den Messtagen bezüglich der Probanden mit visueller Konvergenz wird verzichtet, da lediglich 28 % (Studie I) und 23 % (Studie II) der Probanden an beiden Messtagen konvergente Mittelwerte aufwiesen.

Neben der visuellen Überprüfung der Konvergenz der kumulierten Mittelwertkurve existiert die Möglichkeit, die Anzahl, ab wann der Mittelwert kumuliert, über das Mittelwertkriterium zu bestimmen (↗ Glg. 2.4.5). In den Abbildungen 3.1.3 und 3.1.4 ist deutlich zu erkennen, dass ein alleiniges Verlassen auf das Mittelwertkriterium für diskrete Parameter in die Irre führen kann und dass eine visuelle Überprüfung unumgänglich ist. Abbildung 3.1.5 zeigt die kumulierten Mittelwert-

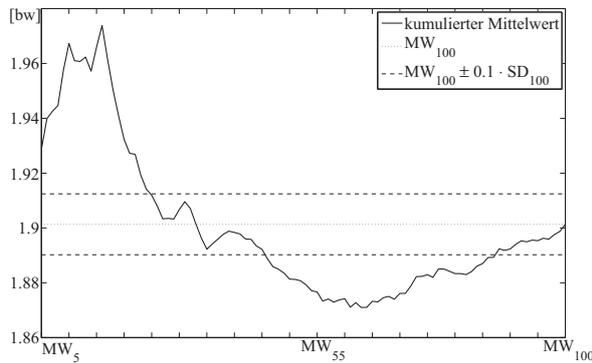


Abbildung 3.1.3:
Beispiel eines divergierenden Mittelwertes (F_{z_I})

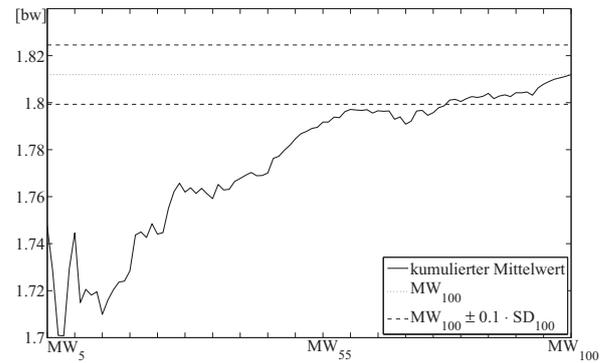


Abbildung 3.1.4:
Beispiel eines divergierenden Mittelwertes (F_{z_I})

folgen für die zweite vertikale Kraftspitze F_{z_A} . Zur besseren Veranschaulichung und Vergleichbarkeit wurden diese Folgen z-transformiert, wobei der Mittelwert aus allen 100 aufgenommenen Versuchen von den Folgengliedern abgezogen wurde und diese Differenz auf die Standardabweichung aus allen 100 Werten normiert wurde (↗ Glg. 2.4.3).

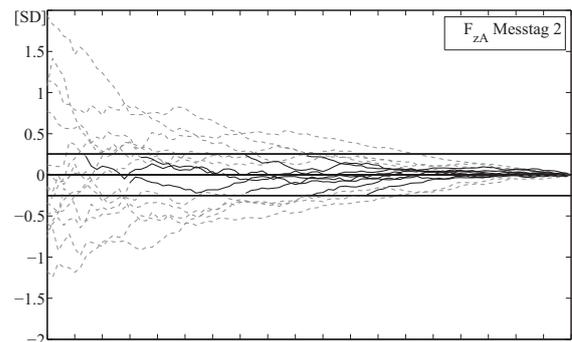
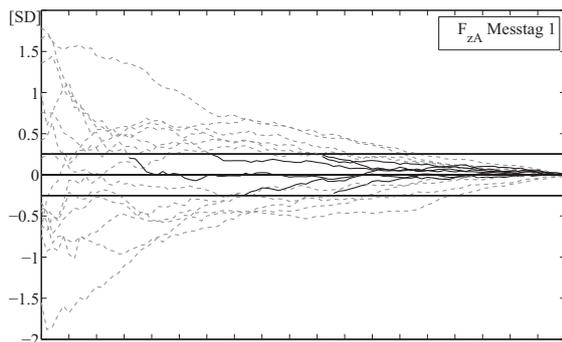


Abbildung 3.1.5:
z-transformierte Folgen kumulierter Mittelwerte der zweiten Kraftspitze (F_{z_A}) für beide Messtage mit Angabe des Mittelwertkriteriums für $q = 0.25$; durchgängig grau gestrichelt: divergente kumulierte Mittelwertfolgen, Wechsel von grau zu schwarz: minimale Anzahl für konvergente Mittelwertfolgen ermittelt mit Mittelwertkriterium

Es ist deutlich zu erkennen, dass für einige Probanden keine Annäherung an Null erfolgt und diese als divergent gelten. Weiterhin ist erkennbar, dass die minimale Anzahl, intuitiv bestimmt, größer ausfällt als jene durch das Mittelwertkriterium ermittelte Anzahl.

In Tabelle B.2 (↗ S. 114) ist die mittlere Anzahl der gemessenen Werte angegeben, ab welcher die Werte in das vorgegebene Intervall fallen. Für die Berechnung dieser mittleren Anzahl werden lediglich die Probanden verwendet, welche nach visueller Überprüfung eine konvergente Mittelwertkurve aufwiesen. In Tabelle 3.1.4 ist eine Zusammenfassung anhand der gemittelten Anzahl über alle Parameter dargestellt.

	Studie I			Studie II		
	M 1 (#)	M 2 (#)	Differenz	M 1 (#)	M 2 (#)	Differenz
$q = 0.1$	74 ± 13	69 ± 23	16 ± 10	58 ± 13	51 ± 16	14 ± 12
$q = 0.25$	53 ± 19	38 ± 23	20 ± 16	30 ± 16	26 ± 17	18 ± 14
$q = 0.5$	18 ± 13	17 ± 16	10 ± 11	9 ± 11	8 ± 9	11 ± 9

Tabelle 3.1.4:

Zusammenfassung Mittelwertkriterium: mittlere Anzahl über alle Parameter der Probanden mit visueller Konvergenz \pm Standardabweichung, mittlere Differenz an Messtag 1 (M1) und Messtag 2 (M2) \pm Standardabweichung

Dabei ist auffällig, dass mit zunehmender Akzeptanz der Standardabweichung aller 100 gemessenen Werte ($q = 0.1, 0.25, 0.5$) die mittlere Anzahl, welche in das Intervall fällt, abnimmt. Werden für 20 % Abweichung ($q = 0.1$) 60 bis 70 Werte benötigt, ergibt eine 100 %-ige Abweichung ($q = 0.5$) lediglich 10 bis 20 zu absolvierende Läufe. Zudem zeigen die hohen Standardabweichungen, dass die minimale Anzahl zwischen den Probanden große Unterschiede aufweist. Die individuelle Betrachtung zeigt, dass die Anzahl ermittelt mit dem Mittelwertkriterium nicht stabil über zwei Messtage ist. Zum Teil treten erhebliche Differenzen innerhalb eines Probanden auf (siehe Tabelle 3.1.4). Dabei sei zu beachten, dass lediglich 28 % (Studie I) und 23 % (Studie II) der möglichen Probanden-Parameter-Kombinationen an beiden Messtagen konvergierende Mittelwertkurven aufwiesen.

Das Mittelwertkriterium für diskrete Parameter ist neben der Abhängigkeit von der Akzeptanz der Standardabweichung zusätzlich von der verwendeten Gesamtanzahl abhängig. Dies zeigen die Tabellen B.3 (\nearrow S. 115) und 3.1.5. Dabei wurde das Mittelwertkriterium jeweils für $n = 20, \dots, 100$ berechnet und die gemittelte Anzahl auf n normiert, so dass eine prozentuale Anzahl bestimmt wurde. Im Mittel entspricht die minimale Anzahl analysiert anhand Mittelwertkriterium für $q = 0.1$ 65 % der eingesetzten Gesamtanzahl. Für $q = 0.25$ ergibt sich eine mittlere prozentuale Anzahl von 35 ± 4 [%] und für $q = 0.5$ ein Intervall von 13 ± 3 [%].

	Studie I		Studie II	
	Messtag 1 (#)	Messtag 2 (#)	Messtag 1 (#)	Messtag 2 (#)
$q = 0.1$	66 ± 4	64 ± 4	65 ± 4	62 ± 3

Tabelle 3.1.5:

Zusammenfassung Abhängigkeit Mittelwertkriterium: mittlere prozentuale Anzahl über alle Parameter \pm Standardabweichung

Bezüglich der Gruppierungen ergibt sich für die mittels Mittelwertkriterium bestimmte Anzahl ein ähnliches Bild wie für die visuell bestimmte Anzahl. Innerhalb der Gruppierungen für Studie I (Geschlecht, Ausdauerlauf) konnten keine signifikanten Unterschiede ermittelt werden. Die Differenzierungen bezüglich Studie II (Geschlecht, sportliche Aktivität, Laborerfahrung) weisen lediglich für die Gruppe mit sportlicher Aktivität für die maximale Kraft in der Beschleunigungsphase (FR_{yA} (M1)) einen signifikanten Unterschied auf. Dieser ist auf eine geringe Anzahl (2) an eingehenden sportlich aktiven Läufern zurückzuführen und muss daher als nicht relevant gelten. Die Läufer,

welche bereits über Laborerfahrung verfügen, zeigen für einige Parameter (FR_{z_I} (M2), F_{y_B} (M2), FR_{y_A} (M2)) einen signifikanten Unterschied von ungefähr 20 zusätzlichen Läufen.

Eine individuelle Betrachtung der Folgen kumulierter Mittelwerte für jede Parameter-Proband-Kombination zwischen zwei Messtagen offenbart sieben verschiedene, wiederkehrende Muster:

- Typ 1: Beide Kurven folgen einem identischen Verlauf und konvergieren nach der gleichen endlichen Anzahl. Dieser Typ stellt den optimalen Verlauf dar. Es ist möglich den diskreten Parameter mittels Mittelwert mit einer gewissen Anzahl reliabel zu schätzen (Abb. 3.1.6 Reihe 1, links).
- Typ 2: Beide Kurven folgen einem identischen Verlauf und divergieren. Eine Stabilisierung wird über den gesamten Zeitraum nicht erreicht (Abb. 3.1.6 Reihe 1, mittig).
- Typ 3: Die Kurven verlaufen parallel. Hier wurde nicht zwischen konvergierend oder divergierend unterschieden, da in beiden Fällen keine reliable Messung erreicht wird. Selbst wenn sich der Mittelwert stabilisiert, möglicherweise mit der gleichen Anzahl, wird dennoch nicht der gleiche Wert zwischen zwei Messtagen erreicht (Abb. 3.1.6 Reihe 1, rechts).
- Typ 4: Die Kurven starten mit unterschiedlichen Werten und enden auf dem gleichen Niveau. Dies ist dann optimal, wenn die Stabilisierung bei einer endlichen, möglichst gleichen Anzahl einsetzt (Abb. 3.1.6 Reihe 2, links).
- Typ 5: Die Kurven starten auf dem gleichen Niveau und enden mit unterschiedlichen Werten. An dieser Stelle wurde wie bei Typ 3 nicht zwischen Konvergenz und Divergenz unterschieden, da unter beiden Bedingungen keine reliable Messung erzielt wurde (Abb. 3.1.6 Reihe 2, mittig).
- Typ 6: Die Folgen starten und enden auf unterschiedlichem Niveau, verlaufen jedoch nicht parallel. Die Kurven kreuzen sich (Abb. 3.1.6 Reihe 2, rechts).
- Typ 7: Die Folgen starten und enden auf dem gleichen Niveau, verlaufen jedoch konvex (Abb. 3.1.6 Reihe 3, links).
- Typ 8: Diese Kategorie umfasst alle sonstigen Kombinationen.

Diese unterschiedlichen Typen sind in allen Parametern vorhanden. Eine Abhängigkeit zwischen Typisierung und Parameter ist nicht gegeben. Eine für die Verwendung des Gyrometer eventuell mögliche Häufung von Typ 3 (Auftreten eines Offsets) konnte nicht beobachtet werden. Dabei ergeben sich für Studie I 98 Kombinationen (7 mal 14) und für Studie II 180 Kombinationen (9 mal 20). In Tabelle 3.1.6 sind die relativen Häufigkeiten in Prozent für das Auftreten der unterschiedlichen Typen von kumulierten Mittelwertfolgen zwischen zwei Messtagen notiert. Der optimale Verlauf (Typ 1) tritt dabei lediglich in 4% bzw. 6% der Fälle auf. Die ungünstigen Verläufe von Typ 2, 3, 5 und 6 sind dabei am häufigsten vertreten (72% für Studie I und 71% für Studie II). In ungefähr 7 von 10 Vergleichen verlaufen die beiden Kurven entweder divergierend oder enden auf unterschiedlichem Niveau.

Einige Probanden wiesen eine Häufung jener Typen mit starken Abweichungen nach Ablauf der 100 Messungen für alle Parameter auf (Typ 3, 5, 6), wobei ein Proband durchgängig parallele Mittelwertkurven zeigte (Typ 3). Für Studie I wurde eine Häufung ab 6 von 7 Parametern angenommen, bei Studie II ab 8 von 9 Parametern. Unter dieser Annahme wurden in Studie I 3 von 14 Probanden

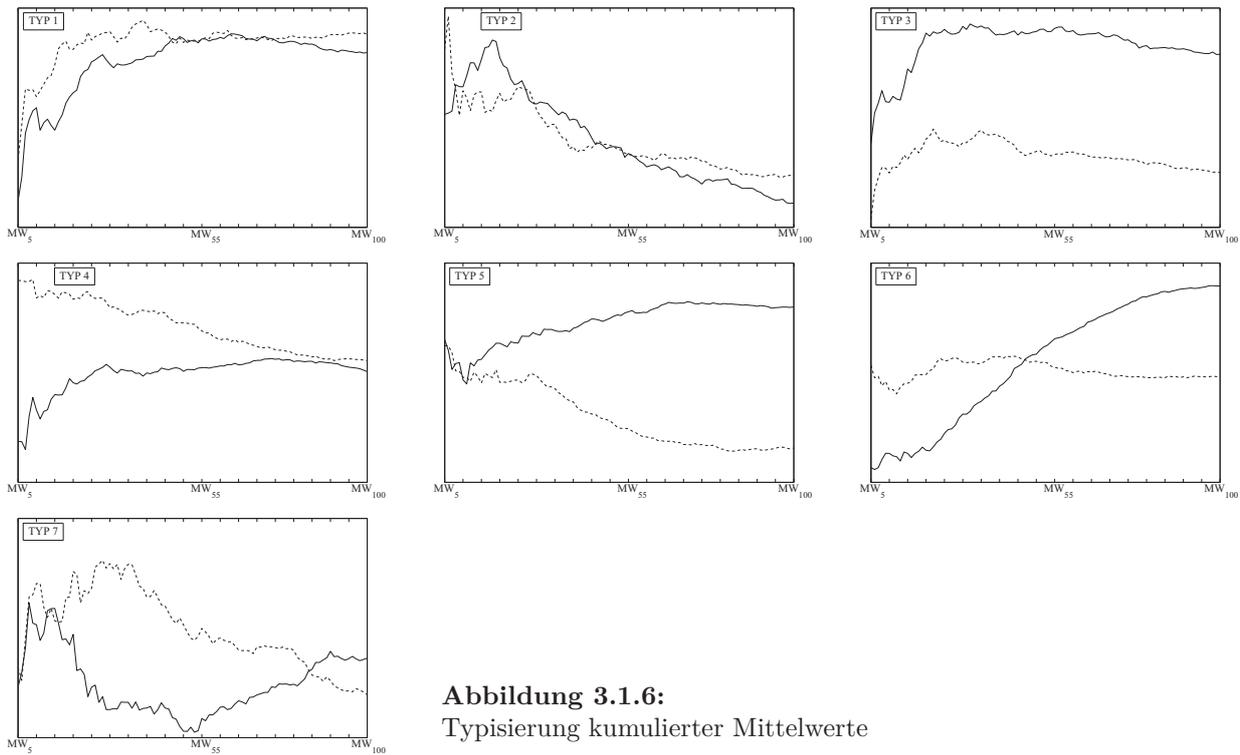


Abbildung 3.1.6:
Typisierung kumulierter Mittelwerte

Typen	1	2	3	4	5	6	7	8
Studie I	4 %	7 %	34 %	14 %	19 %	12 %	5 %	4 %
Studie II	6 %	3 %	42 %	18 %	18 %	8 %	2 %	3 %

Tabelle 3.1.6:
Typisierung kumulierter Mittelwertkurven für zwei Messtage (prozentualer Anteil)

mit einer derartigen Häufung registriert und für Studie II 4 von 20 Probanden. Eine Systematik zwischen den Geschlechtern, der Laborerfahrung, der sportlich aktiven Gruppe oder den Ausdauerläufern konnte nicht festgestellt werden. Die verschiedenen Typisierungen treten nicht gehäuft für diese verschiedenen Gruppierungen auf.

Ein Vergleich der Konfidenzintervalle (\nearrow Glg. 2.4.13) für alle 100 Versuche zwischen Messtag 1 und 2 zur Überprüfung der Zuverlässigkeit der Messung des Mittelwertes zeigt, dass sich für rund 50 % der Probanden die Mittelwerte zwischen den Messtagen signifikant unterscheiden (Tabelle 3.1.7). In 50 % der Fälle kann der Mittelwert somit nicht zuverlässig an unterschiedlichen Tagen gemessen werden.

Die Betrachtung der Gruppierungen zeigte keine Systematiken. Im Vergleich der Geschlechter zwischen den Bodenreaktionskräften von Studie I und II wird dies besonders deutlich. Weisen in Studie I die Männer mehr signifikant unterschiedliche Konfidenzintervalle auf, dominieren in Studie II die Frauen die signifikant unterschiedlichen Konfidenzintervalle.

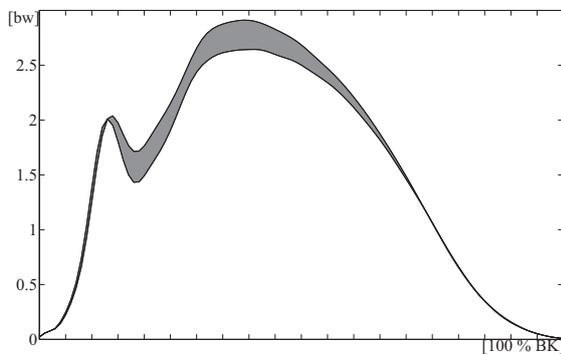
	F_{z_I}	F_{z_A}	FR_{z_I}	F_{y_B}	F_{y_A}	FR_{y_B}	FR_{y_A}	ω_{max}	φ_{max}
Studie I (14)	6	8	4	5	7	11	7		
Studie II (20)	11	8	13	11	9	9	10	9	12

Tabelle 3.1.7:

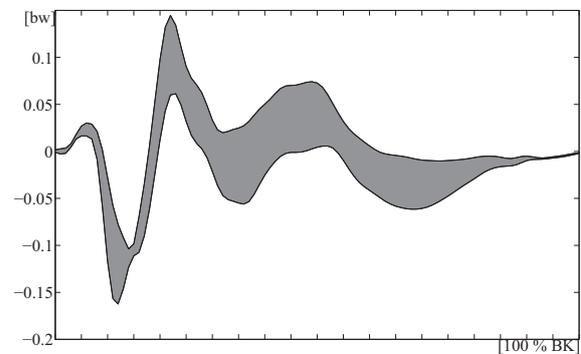
Anzahl signifikant unterschiedlicher Konfidenzintervalle diskreter Parameter für MW_{100}^{pr} an zwei unterschiedlichen Messtagen

3.2 Betrachtung der Gesamtkurve

Vorab werden die Differenzen DA_{intra}^{kr} innerhalb eines Messtages betrachtet, wobei die Differenzenfläche gegenüber der mittleren Fläche der gesamten ersten fünf Kurven und aller 100 Kurven berechnet wurde (↗ Glg. 2.4.11). Unter der Annahme, dass die Kurve aus allen 100 Kurven das Optimum darstellt, gibt die prozentuale Differenz an, wie reliabel die entsprechenden Kurven innerhalb eines Tages gemessen werden können. Abbildung 3.2.1 zeigt ein Beispiel, wie stark sich die vertikale Bodenreaktionskraft innerhalb eines Tages unterscheiden kann und Abbildung 3.2.2 zeigt dies für die Bodenreaktionskraft in mediolateraler Richtung. Es ist zudem deutlich, dass die absolute Abweichung für kleiner ausgeprägte Kurven wie die F_x -Kurve wesentlich mehr ins Gewicht fällt und eine höhere prozentuale Abweichung ermittelt wird, da die mittlere Fläche, auf welche normiert wird, relativ klein ist. Nach grafischer Betrachtung würde man eine kleinere prozentuale Abweichung für F_x als die angegebenen 131 % vermuten.


Abbildung 3.2.1:

Beispiel Differenzenfläche DA_{intra}^{kr} für $k = F_z$ (prozentuale Abweichung 7 %)


Abbildung 3.2.2:

Beispiel Differenzenfläche DA_{intra}^{kr} für $k = F_x$ (prozentuale Abweichung 131 %)

Tabelle 3.2.1 zeigt die für jede Kurve ermittelte maximale prozentuale Differenz und die über alle Probanden gemittelte prozentuale Differenz. Es wird deutlich, dass die vertikale Bodenreaktionskraft, die geringsten Schwankungen zu verzeichnen hat. Die Bodenreaktionskraft in mediolateraler Richtung und die Winkelgeschwindigkeit sind weniger reliabel innerhalb eines Tages zu messen. Es treten durchaus erhebliche Differenzen innerhalb eines Tages auf. Jedoch konnten keine signifikanten Unterschiede für diese Abweichungen innerhalb der drei Gruppierungen Geschlecht, sportliche

Aktivität und Ausdauerläufer gefunden werden. Lediglich für die Bodenreaktionskraft in mediolateraler Richtung konnte innerhalb der Gruppe ohne wöchentliches Laufpensum für Studie II und Messtag 1 eine signifikant niedrigere mittlere Differenz von 20 % festgestellt werden. Die prozentualen Differenzen bezüglich der Winkelgeschwindigkeit sind am ersten Messtag für Probanden, welche über keine Laborererfahrung verfügen, um 10 % höher. Am zweiten Messtag nahmen die prozentualen Unterschiede der laborerfahrenen Gruppe zu, wohingegen die prozentuale Abweichung der laborunerfahrenen Gruppe auf dem Niveau von Messtag 1 blieb.

k	Studie I		Studie II	
	Messung 1 (%)	Messung 2 (%)	Messung 1 (%)	Messung 2 (%)
F_z	8 (4 ± 2)	7 (4 ± 2)	7 (4 ± 2)	4 (3 ± 1)
F_y	24 (11 ± 6)	23 (10 ± 5)	19 (10 ± 5)	25 (9 ± 5)
F_x	103 (37 ± 26)	63 (27 ± 14)	131 (37 ± 27)	85 (31 ± 17)
ω			34 (21 ± 7)	34 (21 ± 8)

Tabelle 3.2.1:

Differenz DA_{intra}^{kr} : maximale prozentuale Abweichung (mittlere Differenz ± mittlere Standardabweichung der Differenz)

Die Differenzen DA_{inter}^{kr} , welche innerhalb zwei Messtagen auftreten können, fallen für die Bodenreaktionskräfte im Mittel geringer aus als die Differenzen, welche an einem Tag auftreten können, wobei dies für rund 75 % der Probanden von Studie I und 66 % von Studie II zutrifft. Für die Winkelgeschwindigkeit wurde im Mittel eine ähnliche Differenz ermittelt, welche allerdings variabler innerhalb der Probanden ist. Die maximale prozentuale Abweichung fiel höher aus als die maximale Differenz, welche an einem Tag erreicht wurde (siehe Tabelle 3.2.2). Abhängigkeiten innerhalb der verschiedenen Gruppierungen konnten für die Probanden mit wöchentlichem Laufpensum nachgewiesen werden, wobei für die Bodenreaktionskraft in anterior-posterior Richtung eine signifikante Differenz von plus 3 % und die Bodenreaktionskraft in mediolateraler Richtung von mehr als 10 % belegt werden konnte. Die laborerfahrenen Probanden wiesen für Studie II eine höhere prozentuale Abweichung um 10 % für die Bodenreaktionskraft in anterior-posterior Richtung auf. Dies konnte nicht für die drei weiteren Kurven festgestellt werden.

k	Studie I (%)	Studie II (%)
F_z	5 (2 ± 1)	5 (3 ± 1)
F_y	12 (7 ± 3)	18 (8 ± 3)
F_x	54 (23 ± 13)	45 (22 ± 10)
ω		64 (23 ± 15)

Tabelle 3.2.2:

Differenz DA_{inter}^{kr} zwischen zwei Messtagen: maximale prozentuale Abweichung (mittlere prozentuale Abweichung ± Standardabweichung)

Abbildung 3.2.3 zeigt das Beispiel für die Differenzenkurve mit der maximalen Abweichung (64 %) für die Winkelgeschwindigkeit zwischen den Messtagen und in Abbildung 3.2.4 die dazugehörigen

entsprechenden Differenzenkurven DA_{intra}^{kr} innerhalb eines Messtages. Für diesen Probanden war es somit nicht möglich, die Winkelgeschwindigkeit zwischen zwei unterschiedlichen Messtagen reliabel zu messen.

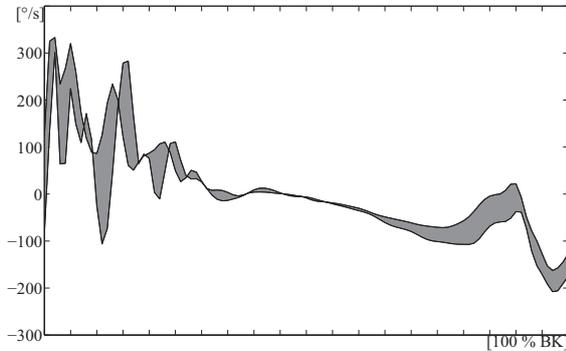


Abbildung 3.2.3:

Beispiel Differenzenfläche DA_{inter}^{kr} für ω zwischen Messtag 1 und 2 (prozentuale Abweichung 64%)

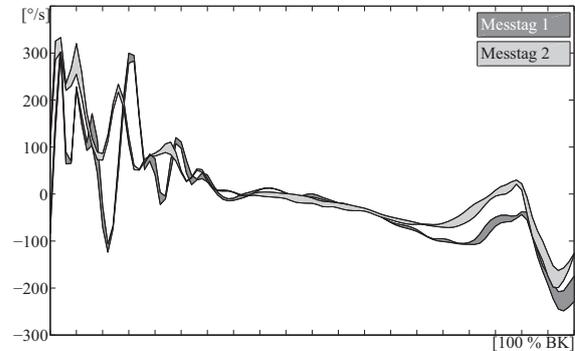


Abbildung 3.2.4:

Differenzenkurven innerhalb eines Tages für ω aus Abbildung 3.2.3 (prozentuale Abweichung Messtag 1: 17% / Messtag 2: 21%)

Neben der absoluten Abweichung innerhalb eines Tages als auch zwischen den Tagen ist zudem interessant, ob und ab welcher Anzahl die Mittelwertkurven konvergieren. Hierfür erfolgte eine visuelle Überprüfung der Stabilisierung der gesamten kumulierten Kurven k punktweise für $m = 1, \dots, 101$, wobei die Folgen auf MW_{m100}^{kr} und SD_{m100}^{kr} normiert wurden (\nearrow Glg. 2.4.4). Die entstehenden zum Teil trichterförmigen Grafiken sollten um Null konvergieren. Eine Entscheidung über die tatsächliche Konvergenz aller Kurven zu treffen, gestaltet sich dabei schwieriger als im diskreten Fall. Es müssen nicht alle 100 Kurven bei der gleichen Anzahl konvergieren und dies kann für jedes m unterschiedlich sein. Wenn eine große Anzahl der 100 kumulierten Folgen der Kurve k konvergiert, wird die Gesamtkurve als konvergent bezeichnet. Abbildung 3.2.5 zeigt ein Beispiel für konvergente kumulierte Mittelwertfolgen. Es ist erkennbar, dass ab einer bestimmten Anzahl die Kurven normiert auf deren letzten Wert um Null schwanken. In Abbildung 3.2.6 hingegen fallen die Werte erst im letzten Punkt zusammen (bedingt durch die Normierung). Diese Kurven müssen als divergent betrachtet werden.

Tabelle B.6 (\nearrow S. 117) zeigt das Ergebnis der visuellen Überprüfung für jede Kurve F_z , F_y , F_x und ω . Es ist zu erkennen, dass die mittlere Anzahl, ab wann sich die punktweisen Mittelwertfolgen stabilisieren, ungefähr gleich der für diskrete Parameter ist (rund 70). Die Zusammenfassung in Tabelle 3.2.3 zeigt, dass die Schwankung dieser mittleren Anzahl relativ stabil bleibt. Durchschnittlich sind für mehr Probanden konvergierende punktweise Mittelwertkurven als Mittelwertfolgen für diskrete Parameter zu verzeichnen. Eine Abhängigkeit bezüglich der vier Gruppierungen konnte nicht nachgewiesen werden. Lediglich für die vertikale Bodenreaktionskraft konnte für Studie I (Tag 1) für Frauen eine geringere mittlere Anzahl von 48 gegenüber 64 nachgewiesen werden. Die mittlere Anzahl ist dennoch relativ homogen. Kein Proband wies durchweg divergierende Kurven auf.

Bei Betrachtung der Differenzenflächen DA_{intra}^{kr} für Probanden mit konvergenten Mittelwertkurven zeigt sich, dass sich die mittleren und maximalen prozentualen Abweichungen nur unwesentlich

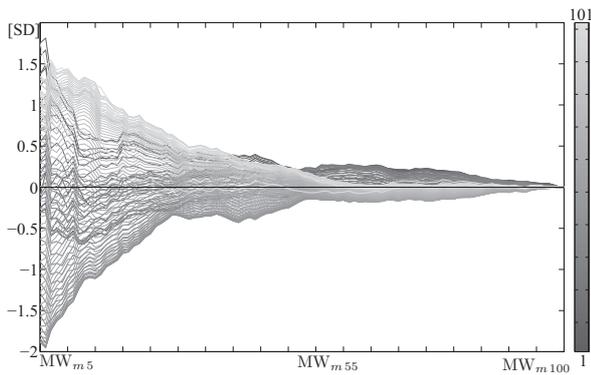


Abbildung 3.2.5:

Beispiel von konvergierenden punktwise kumulierten Mittelwertkurven MW_{mn}^{kr} eines Probanden mit $k = F_z$ (Konvergenz ab 60): mit $n = 5, \dots, 100$ (x-Achse) und $m = 1, \dots, 101$ (dunkelgraue bis hellgraue Kurven)

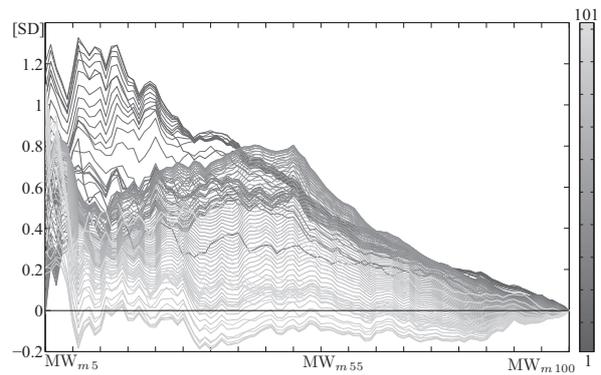


Abbildung 3.2.6:

Beispiel von divergierenden punktwise kumulierten Mittelwertkurven MW_{mn}^{kr} eines Probanden mit $k = F_z$: mit $n = 5, \dots, 100$ (x-Achse) und $m = 1, \dots, 101$ (dunkelgraue bis hellgraue Kurven)

	Studie I		Studie II	
	Messtag 1 (#)	Messtag 2 (#)	Messtag 1 (#)	Messtag 2 (#)
Konvergenz	10 (69 %)	10 (69 %)	11 (56 %)	14 (71 %)
	68 ± 11	67 ± 12	71 ± 9	70 ± 10

Tabelle 3.2.3:

Zusammenfassung der Konvergenz des Mittelwertes gesamter Kurven: mittlere Anzahl konvergierender Mittelwerte (absoluter und prozentualer Anteil) / mittlere Anzahl \pm Standardabweichung

gegenüber dem gesamten Probandenkollektiv verändern (\nearrow Tab. B.9, S. 118). Die Differenzen DA_{inter}^{kr} zwischen den Messtagen fallen etwas geringer aus als die Differenzen, welche für die gesamte Stichprobe erfasst wurden. Das Verhältnis innerhalb eines Messtages und zwischen den Messtagen wird für die Probanden mit konvergierenden Mittelwertkurven beibehalten (\nearrow Tab. B.10, S. 119).

Aufgrund der Schwierigkeit visuell die richtige Anzahl im Falle konvergierender punktwiser Mittelwertfolgen zu finden, wird zusätzlich das Mittelwertkriterium (\nearrow Glg. 2.4.5) angewendet. In Tabelle B.7 (\nearrow S. 117) ist die entsprechende Anzahl notiert, ab welcher die gesamten Mittelwertkurven in das vorgegebene Intervall für $q = 0.1, 0.25, 0.5$ fallen. Es ist erkennbar, dass für die Betrachtung gesamter Kurven die Anzahl von der Toleranz der jeweiligen Standardabweichungskurve abhängig ist. Werden für 20 % Toleranz im Mittel 84 Versuche benötigt, verringert sich die Anzahl auf rund 40 Versuche bei einer Toleranz von 100 % Standardabweichung. Die geringen Standardabweichungen verdeutlichen, dass die mittels Mittelwertkriterium errechnete Anzahl stabil über alle Probanden verteilt ist. Tabelle 3.2.4 zeigt zudem, dass die Anzahl zwischen zwei Messtagen relativ reliabel gemessen werden kann, wobei der Anteil an Probanden, welche an beiden Messtagen konvergente Mittelwertkurven aufweisen, für Studie I 52 % und für Studie II 41 % beträgt.

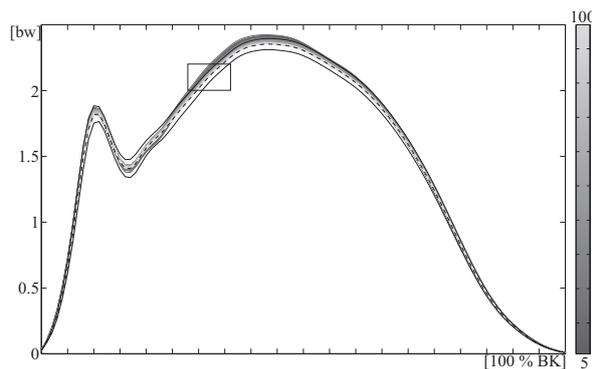
Die Abbildungen 3.2.7 und 3.2.8 zeigen die Anwendung des Mittelwertkriteriums auf die gesamten

	Studie I			Studie II		
	M 1 (#)	M 2 (#)	Differenz	M 1 (#)	M 2 (#)	Differenz
$q = 0.1$	84 ± 4	84 ± 3	3 ± 2	83 ± 3	83 ± 4	4 ± 3
$q = 0.25$	65 ± 7	66 ± 5	8 ± 5	63 ± 7	62 ± 8	7 ± 6
$q = 0.5$	44 ± 9	43 ± 9	8 ± 7	39 ± 9	36 ± 11	11 ± 9

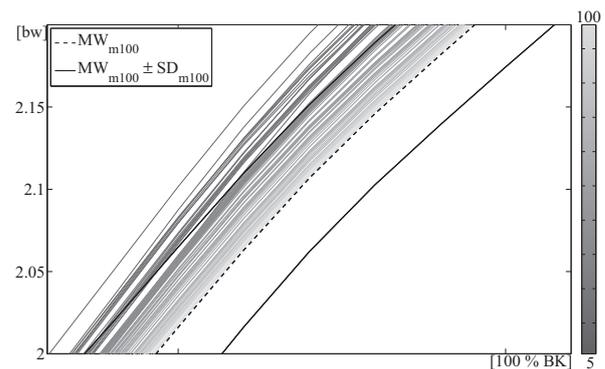
Tabelle 3.2.4:

Zusammenfassung Mittelwertkriterium: mittlere Anzahl über alle Parameter jener Probanden mit visueller Konvergenz \pm Standardabweichung / mittlere Differenz der Anzahlen über zwei Messtage

kumulierten Kurven. Es ist zu erkennen, dass die Mittelwertkurve MW_{m100}^{kr} und die Standardabweichungskurve SD_{m100}^{kr} und damit das entsprechende Intervall des Mittelwertkriteriums eine ähnliche Gestalt haben wie alle vertikalen Bodenreaktionskraftkurven. Daher ist gewährleistet, dass die kumulierten Mittelwertkurven tatsächlich innerhalb des Intervalls auftreten können. Die Analyse der entsprechenden Gruppierungen (Geschlecht, sportliche Aktivität, Ausdauer, Laborerfahrung) bezüglich der ermittelten Anzahl anhand Mittelwertkriterium ergab keinerlei signifikante Unterschiede.


Abbildung 3.2.7:

Beispiel kumulierter Mittelwertkurven der vertikalen Bodenreaktionskraft mit Mittelwertkriterium


Abbildung 3.2.8:

Ausschnitt aus Abbildung 3.2.7 (F_z) mit Mittelwertkriterium

Man kann zeigen, dass neben der Abhängigkeit von der tolerierten Standardabweichung das Mittelwertkriterium für die Betrachtung gesamter Kurven zudem abhängig von der eingesetzten Gesamtanzahl ist. Tabelle B.8 (\nearrow S. 118) zeigt die mittlere Anzahl normiert auf die eingesetzte Gesamtanzahl in Prozent an. Dabei wurden schrittweise die ersten 20 Mittelwertkurven bis hin zu allen 100 Mittelwertkurven betrachtet. Als Intervall kamen jeweils die letzte Mittelwert- und Standardabweichungskurve der verwendeten 20 bis 100 Kurven zur Anwendung. Es ist zu erkennen, dass für eine tolerierte Abweichung von 50 % von der Standardabweichungskurve etwa 60 % an Versuchen für die stabile Messung von F_z , F_x und ω benötigt werden. Für die Bodenreaktionskraft in anterior-posterior Richtung werden bis zu 84 % an Läufern benötigt. In Tabelle 3.2.5 ist zudem erkennbar, dass diese prozentuale Anzahl über die gesamten 80 Berechnungen wenig schwankt.

	Studie I		Studie II	
	Messtag 1 (#)	Messtag 2 (#)	Messtag 1 (#)	Messtag 2 (#)
$q = 0.25$	71 ± 4	70 ± 5	69 ± 4	67 ± 3

Tabelle 3.2.5:

Zusammenfassung Abhängigkeit Mittelwertkriterium: mittlere prozentuale Anzahl über alle Parameter \pm Standardabweichung

Für eine tolerierte Abweichung von 20 % von der entsprechenden Standardabweichungskurve werden 84 ± 3 [%] an Versuchen von der Gesamtanzahl benötigt und für 100 % tolerierter Abweichung müssen 45 ± 6 [%] Versuche aufgezeichnet werden. Die geringe Schwankung dieser prozentualen Anteile lässt das Mittelwertkriterium für die Verwendung innerhalb der Untersuchungen bezüglich Hypothese III ungeeignet erscheinen. Die anhand Mittelwertkriterium ermittelte minimale Anzahl wäre lediglich abhängig von der eingesetzten Anzahl an Versuchen, nicht von der tatsächlichen den Probanden charakterisierenden Variabilität.

3.3 Diskussion der Ergebnisse

Die Untersuchungen zu Hypothese I umfassten Analysen hinsichtlich der Konvergenz sowie der Differenzen innerhalb und zwischen verschiedenen Messtagen des Mittelwertes diskreter Parameter und der Mittelwertkurven von Gesamtkurven. Die erheblichen Differenzen für diskrete Parameter innerhalb eines Messtages deuten darauf hin, dass mehr als die bisher üblichen fünf Versuche notwendig sind, um den Mittelwert zu bestimmen und dessen Konvergenz zu erreichen, so denn er existiert und den Probanden tatsächlich repräsentiert. Im Fall konvergenter Mittelwertkurven fielen die Differenzen innerhalb und zwischen den Messtagen leicht geringer aus als für das gesamte Probandenkollektiv. Die maximalen Abweichungen reduzierten sich für einige Parameter wesentlich, für andere Parameter blieben sie gleich, wobei eine prozentuale Abweichung unter Konvergenz von bis zu 25 % in den häufig verwendeten Parametern die Notwendigkeit der Ermittlung einer adäquaten minimal zu messenden Anzahl an Versuchen unterstreicht. Bei der Betrachtung der unterschiedlichen Parameter fällt auf, dass es stabil messbare Parameter gibt und Parameter, die stark innerhalb eines Tages schwanken. Die Bestimmung der minimal zu verwendenden Anzahl an Versuchen sollte die starke Streuung der Werte mit einbeziehen, so dass auch die höher schwankenden Parameter adäquat gemessen werden können.

Für diskrete Parameter wurden für 50 % der Probanden divergierende Folgen kumulierter Mittelwerte festgestellt. Für diese Probanden ist es somit nicht möglich, den Mittelwert als die den Probanden charakterisierende Kennzahl zu verwenden, wobei fraglich bleibt, ob für diese Probanden generell eine Konvergenz des Mittelwertes bei mehr als 100 Messversuchen eintreten würde. Die Option, mehr als 100 Messversuche pro Untersuchungsbedingung aufzunehmen, ist hingegen ökonomisch nicht durchführbar. Im Fall von Konvergenz streuen die mindestens zu absolvierenden Läufe derart hoch, dass von einer individuell zu bestimmenden minimalen Anzahl auszugehen ist. Die visuell bestimmte Anzahl von 60 bis 70 Werten erscheint sehr hoch und deutet zudem auf eine hohe, in einigen Studien nicht mehr ökonomische Anzahl hin. Probanden, welche eine derart hohe

Anzahl an Versuchen benötigen, wären in einer Untersuchung mit mehr als zwei Untersuchungsbedingungen nicht mehr tragbar.

Die verschiedenen Gruppierungen zeigen kaum Systematiken auf, wobei zu beachten ist, dass die Stichprobengrößen innerhalb der einzelnen Gruppierungen relativ gering ausfallen. Dies erschwert die Diskussion über den generellen Einfluss dieser unterschiedlichen Gruppierungen. Für die Reduzierung der minimalen Anzahl scheint tendenziell die Untersuchung von sportlich aktiven Probanden von Vorteil, da diese 5 bis 10 Versuche weniger benötigen. Im Hinblick auf die zeitliche Planung von Studien ist eine möglichst geringe Anzahl an Versuchen von Vorteil. Dies könnte jedoch dazu animieren lediglich bestimmte Probandengruppen auszuwählen, welche tendenziell weniger Messversuche benötigen. Im Sinne einer optimalen Versuchsplanung minimiert dies hingegen die Repräsentativität der Untersuchung.

Den optimalen Probanden mit immer konvergierenden Mittelwertfolgen mit einer geringen Mindestanzahl an Versuchen scheint es für diskrete Parameter nicht zu geben. Zusätzliche Aufteilungen der Analysen hinsichtlich der Differenzen der Mittelwerte innerhalb und zwischen den Messtagen zeigen keine Beeinflussung durch das Geschlecht, nicht vorhandenes Laufpensum oder die dauerhaft ausgeübte Lauftätigkeit im Labor und stärken die Vermutung der Individualität jedes Probanden. Lediglich die Gruppierung bezüglich der Laborerfahrung erweckt den Anschein, einen relevanten Einfluss auszuüben. Bei näherer Betrachtung scheinen die signifikanten Unterschiede auf Artefakten zu beruhen, da diese Unterschiede nicht konsistent über alle Parameter auftreten und zum Teil widersprüchliche Aussagen liefern. Zum Teil treten die Unterschiede an Messtag 2 auf, wobei nach Gewöhnung am zweiten Messtag keine oder zumindest geringere Unterschiede als an Messtag 1 auftreten sollten. Beim Vergleich der Abweichung der Anzahl, ab wann die Folgen kumulierter Mittelwerte konvergieren, wurde für laborerfahrene Probanden teilweise eine höhere Anzahl ermittelt als für laborunerfahrene Testpersonen. Dies sollte theoretisch genau umgekehrt auftreten.

Das Mittelwertkriterium unterstützt die visuelle Analyse und reflektiert die dabei bestimmte Mindestanzahl an Versuchen. Die gewählten drei Toleranzbereiche entsprechen den Angaben von Bates et al. (1983) und Hamill und McNiven (1990), wobei auf eine kontinuierliche Betrachtung verzichtet wurde. Die Abhängigkeit von diesem gewählten Toleranzbereich und von der verwendeten Gesamtzahl würde durch ein stetiges q lediglich untermauert. Diese Interdependenz macht deutlich, dass das Mittelwertkriterium für die einheitliche Bestimmung einer minimalen Anzahl über alle Probanden ungeeignet ist. Je weniger Läufe insgesamt absolviert werden, um so weniger Versuche sind minimal notwendig. Die Anzahl pendelt sich auf keinem einheitlichen Niveau ein, sondern entspricht lediglich einem bestimmten Anteil an der Gesamtanzahl. Je toleranter die Abweichung angegeben wird, um so weniger Versuche werden benötigt. Somit ist dieses Kriterium nicht objektiv, sondern kann mit Hilfe der Toleranzgrenze und der Gesamtanzahl manipuliert werden. Die Untersuchungen von Hamill und McNiven (1990) und Bates et al. (1983), welche das Mittelwertkriterium zur Bestimmung der minimal notwendigen Anzahl nutzen, müssen daher kritisch betrachtet werden. Die dort angegebene minimale Anzahl von 10 und 8 kommt lediglich durch die verwendete Gesamtanzahl von 20 und 10 und die Toleranzgrenze von 50 % der Standardabweichung zu Stande. Dabei sei zu beachten, dass die maximale Anzahl angegeben wurde, um für alle verwendeten Variablen die korrekte Anzahl zu garantieren. Das Mittelwertkriterium ist somit für die Bestimmung einer minimalen Anzahl ungeeignet. Für die Beschreibung und Bewertung des vorliegenden Datenmaterials kann dieses Kriterium dennoch verwandt werden.

Interessanterweise treten lediglich für 28 % (Studie I) und 23 % (Studie II) der Probanden an beiden Messtagen konvergente Mittelwertkurven auf. Dies reduziert die Anzahl der Probanden, welche für eine Messung überhaupt zur Verfügung stehen. Unter der Annahme der Repräsentativität dieser Studie und der Verallgemeinerbarkeit auf zukünftige Testpersonen wäre das Probandenkollektiv, welches für eine Studie in Frage kommen würde, erheblich eingeschränkt. Die Repräsentativität der Untersuchung würde minimiert und die zugehörige Grundgesamtheit sich entsprechend ändern, so dass eine Verallgemeinerbarkeit nicht mehr zulässig wäre.

Zudem schwankt die minimale Anzahl zur Erreichung von Konvergenz zwischen den Messtagen erheblich, so dass keine stabile einheitliche minimale Anzahl für einen Probanden mittels Mittelwertkriterium bestimmt werden konnte. Die reliable Messung diskreter Parameter muss ferner in Frage gestellt werden aufgrund der rund 50 % signifikant unterschiedlichen Konfidenzintervalle und der damit verbundenen zum Teil erheblichen Abweichung der absoluten Mittelwerte zwischen den Messtagen. Dabei muss Erwähnung finden, dass die Länge der Konfidenzintervalle abhängig ist von der Anzahl an eingehenden Werten. Bei einer Aufnahme von zusätzlichen Messversuchen pro Proband ergeben sich automatisch nach einer gewissen Zeit sich nicht mehr überlappende Konfidenzintervalle des bestimmten Mittelwertes.

Gravierender verdeutlicht der Vergleich der kumulierten Mittelwertfolgen an zwei Messtagen pro Proband die nicht reliablen Messungen. Die Häufung von divergierenden Folgen oder Folgen, welche auf unterschiedlichem Niveau enden, unterstreichen diesen Umstand. Für beide Studien sind davon rund 70 % der Probanden-Parameter-Kombinationen betroffen. Das bei einigen Probanden festgestellte durchgängige Muster von parallelen oder auf anderem Weg auf unterschiedlichem Niveau endenden Folgen für alle untersuchten Parameter deutet ungeachtet ihrer Laborerfahrung auf einen Lerneffekt der Probanden zwischen den Messtagen hin. Die Überlegung von Hamill und McNiven (1990), den Lerneffekt im Labor mit einigen Probeversuchen innerhalb eines Tages auszuschließen und den Messfehler dadurch zu verringern, scheint daher hinfällig. Es bleibt zudem die Frage bestehen, ob mit einem vorgeschobenen Probemesstag ein derartiger Lerneffekt eliminiert und für weitere Messungen ausgeschlossen werden kann und welche Probanden davon betroffen wären, da Laborerfahrung als Begründung ausgeschlossen werden kann. Dies bedarf einer weiteren gesonderten Überprüfung mittels drei oder mehr Messterminen mit derselben Aufgabenstellung. Falls ein Lerneffekt über mehrere Tage eine Ursache für die erheblichen Unterschiede darstellt, müssten sich die kumulierten Mittelwertkurven irgendwann annähern. Diese Überlegungen stellen sich gegen eine ökonomische Bearbeitung von Studien. Ungeachtet der Bereitschaft, eine höhere Versuchsanzahl zu akzeptieren und damit die Untersuchungszeit pro Proband zu erhöhen, würden zusätzliche Probemesstage mit denselben Aufgabenstellungen den zeitlichen Rahmen wissenschaftlicher Studien sprengen, da sie zudem vor jeder Studie durchgeführt werden müssten.

Die hohe Anzahl an signifikant unterschiedlichen Mittelwerten zwischen zwei Messtagen hebt zudem die Untauglichkeit des Mittelwertes diskreter Parameter als die den Probanden repräsentierende Kennzahl hervor.

Die für diskrete Parameter aufgestellte Hypothese I der stabilen und ausreichenden Charakterisierung eines Probanden anhand des Mittelwertes muss daher abgelehnt werden.

Die Analysen bezüglich der gesamten gemittelten Kurven der Bodenreaktionskräfte und der Winkelgeschwindigkeit ergaben für die Betrachtung der Differenz innerhalb eines Tages relativ hohe Ab-

weichungen für F_y , F_x und ω im Vergleich zur vertikalen Bodenreaktionskraft. Dies macht deutlich, dass mehr als fünf Versuche notwendig sind, um eine reliable Messung für alle hier aufgenommenen Kurven pro Proband zu erhalten. Dies wird durch die in etwa gleichbleibenden Differenzen für die Probanden mit konvergenter Mittelwertkurve unterstrichen. Mittels der visuellen Überprüfung konnten für 70 % der Probanden konvergente gesamte punktweise kumulierte Mittelwertkurven festgestellt werden. Dies sind 20 % mehr als für diskrete Parameter, wobei 70 Messversuche notwendig waren. Diese Mindestanzahl an Versuchen konnte im Gegensatz zu den diskreten Parametern relativ stabil über mehrere Tage gemessen werden. Das Mittelwertkriterium angewendet auf die gesamten gemittelten Kurven konnte diese Stabilität bestätigen. Dennoch ist auch für die Gesamtkurven eine Abhängigkeit des Mittelwertkriteriums vom Toleranzbereich und der verwendeten Gesamtanzahl nachweisbar. Dies macht die Anwendung im Rahmen der Auswertungen von Hypothese III unbrauchbar.

Die absolute prozentuale Differenz zwischen verschiedenen Messtagen zeugt von reliablen Messungen für die vertikalen Bodenreaktionskräfte und die Bodenreaktionskräfte in anterior-posterior Richtung. F_x und ω wiesen erhebliche Abweichungen auf, so dass diese als nicht reliabel gelten müssen. Zwischen zwei Messtagen weisen lediglich 41 % der Probanden für Studie I und 52 % für Studie II an beiden Tagen punktweise konvergente Mittelwertkurven auf. Dies ergibt bezogen auf diskrete Parameter mehr Testpersonen. Rund 50 % sind dennoch ein äußerst geringer Prozentsatz. Eine einheitliche minimale Anzahl zu messender Versuche für alle Probanden scheint es aufgrund der Schwankungen innerhalb des Probandenkollektivs für gesamte Kurven ebenfalls nicht zu geben. Dies konstatiert die Notwendigkeit der Bestimmung einer minimalen Anzahl eines Probanden entweder vor der Messung oder anhand eines Algorithmus, welcher während der Messung eingesetzt werden kann.

Eine Betrachtung der divergierenden Kurven deutet nicht auf eine Stabilisierung mit mehr als 100 Versuchen hin. Die Z-Transformation der Mittelwertkurven suggeriert mit dem letzten Wert (=Null) zunächst eine Art Optimum, welches von vornherein durch das relativ hohe Datenvolumen mit der letzten Mittelwertkurve angenommen wurde. Die tatsächliche Existenz eines solchen Optimums über die 100 aufgenommenen Versuche hinaus wird durch diese Auswertungen jedoch in Frage gestellt. Zudem sprengt eine empirische Messung mit mehr als 100 Versuchen pro Untersuchungsbedingung und pro Proband unter Betrachtung von gesamten Kurven den Rahmen wissenschaftlicher Experimente. Bei einer Untersuchung mit lediglich zwei Untersuchungsbedingungen unter der Annahme der Messung von Bodenreaktionskräften beim Laufen und einer Messdauer von 47 ± 9 Minuten pro Bedingung (durchschnittlich ermittelt in Studie I) ergäbe dies eine Gesamtmessdauer pro Proband von zwei Stunden. Im Mittel wurden in den Studien aus der ausgewählten Literatur in Kapitel 1.1 drei bis fünf Bedingungen untersucht. Dies würde in exorbitanten Untersuchungseinheiten enden, so dass selbst unter Annahme der Konvergenz mehr als 100 Versuche nicht tragbar sind.

Die verschiedenen Gruppierungen bezüglich des Geschlechtes, der Laborerfahrung, der sportlichen Aktivität und der ausdauernd ausgeübten Lauftätigkeit im Labor wiesen keinerlei Systematiken auf. Weder innerhalb der Betrachtung der Konvergenz noch hinsichtlich der absoluten, prozentualen Abweichungen innerhalb und zwischen den Messtagen konnte eine Beeinflussung festgestellt werden. Lediglich für die vertikale Bodenreaktionskraft benötigten Frauen visuell betrachtet weniger Versuche als Männer. Die Gruppe der Probanden ohne wöchentliche Lauftätigkeit wiesen prozentual

gesehen sowohl innerhalb als auch zwischen den Messtagen weniger absolute Abweichungen für die Bodenreaktionskraft in anterior-posterior und mediolateraler Richtung auf als die sportlich aktiven Probanden. Dies könnte auf erhöhte Variabilität für aktive Probandengruppen hinweisen und sollte innerhalb der Betrachtungen zu Hypothese II, der Endlichkeit der Variabilität, näher untersucht werden. Zudem konnte die inkonsistente Beeinflussung der Laborerfahrung für Gesamtkurven nachgewiesen werden, welche sich bereits in Bezug auf diskrete Parameter äußerte.

Insgesamt sind die gesamten Kurven der Bodenreaktionskraft im Mittel stabiler messbar als einzelne diskrete Parameter, wobei 30 % divergierender Probanden dennoch bedenklich erscheinen.

Hypothese I, welche eine ausreichende und stabile Charakterisierung des Probanden postuliert, kann im Hinblick auf gesamte Kurven der Bodenreaktionskraft in vertikaler und anterior-posterior Richtung daher nur teilweise bestätigt werden. Für die Bodenreaktionskraft in mediolateraler Richtung und die Winkelgeschwindigkeit sind gemittelte Kurven keine ausreichende Charakterisierung eines Läufers.

Kapitel 4

Resultate zu Hypothese II: Charakterisierung der Probandenvariabilität

Die alleinige Betrachtung des Mittelwertes ist für eine Charakterisierung der Messung eines Probanden nicht ausreichend. Zusätzlich sollte stets die Schwankung des Mittelwertes einbezogen werden. Das getroffene Postulat von Hypothese II besagt, dass die Variabilität der Messung eines Probanden mit endlich vielen Versuchen ermittelt und charakterisiert werden kann. Die Ergebnisse in diesem Kapitel betrachten die Ausprägung der Variabilität diskreter Parameter und der gesamten Kurve der Bodenreaktionskräfte und der Winkelgeschwindigkeit. Es wird überprüft, inwiefern die Variabilität stabil mit einer endlichen Anzahl an Versuchen charakterisiert werden kann. Die diskreten Parameter p entsprechen $F_{zI}, F_{zA}, FR_{zI}, F_{yB}, F_{yA}, FR_{yB}, FR_{yA}, \omega_{max}, \varphi_{max}$ und die Gesamtkurven k werden mit $F_z, F_y, F_x,$ und ω notiert.

4.1 Betrachtung diskreter Parameter

In einem ersten Schritt werden die Standardabweichungskurven einzelner diskreter Parameter p betrachtet. Dabei sei angemerkt, dass, wie in Kapitel 2.4.2 beschrieben, in diesen Fällen die Konvergenz zunächst ausschließlich visuell betrachtet wird und daher die Anzahl, ab wann die Standardabweichungskurven konvergieren, lediglich anhand der Grafik bestimmt werden kann. In Abbildung 4.1.1 ist eine Beispielkurve für eine konvergierende und in Abbildung 4.1.2 für eine divergierende Standardabweichungskurve dargestellt. Als Orientierung für die Konvergenz ist zudem die Standardabweichung aller 100 Werte der maximalen Winkelgeschwindigkeit eingezeichnet. Im Falle der Konvergenz müssten die Werte ab einer gewissen Anzahl von dieser Standardabweichung in nur geringem Maße abweichen. In der ersten Grafik ist eindeutig ein konvergentes Verhalten ab 50 Werten erkennbar. In der zweiten Abbildung erreicht erst der letzte Wert der kumulierten Standardabweichungskurve die Standardabweichung aller 100 Werte. Die Bedingung der Konvergenz, wie sie in Kapitel 2.4.2 angegeben ist und besagt, dass ab einer gewissen Anzahl an Versuchen diese in ein bestimmtes Intervall fallen, ist für diese angegebene Standardabweichungskurve nicht gegeben.

In Tabelle C.1 (↗ S. 121) ist ein Überblick über die Konvergenz der Standardabweichungskurven dargestellt. Angegeben ist die Anzahl der Probanden, welche konvergente Standardabweichungskurven pro Parameter aufweisen sowie die mittlere visuell ermittelte Anzahl, ab wann diese Konvergenz einsetzt. Tabelle 4.1.1 zeigt eine Zusammenfassung mit der mittleren Anzahl konvergenter Standardabweichungskurven und der über alle Probanden und Parameter gemittelten Anzahl an Versuchen, ab wann die jeweiligen Kurven konvergieren.

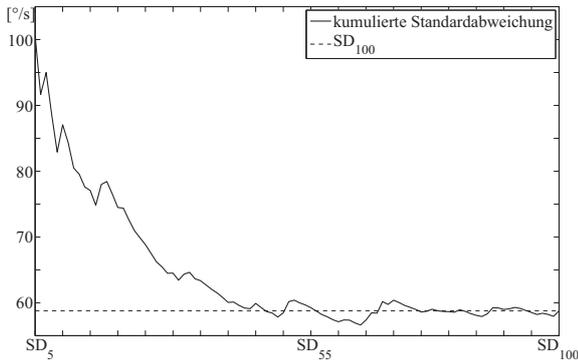


Abbildung 4.1.1:
 Beispiel konvergierender kumulierter Standardabweichungen SD_n^{pr} für $p = \omega_{max}$ (Konvergenz bei ungefähr 50 Werten)

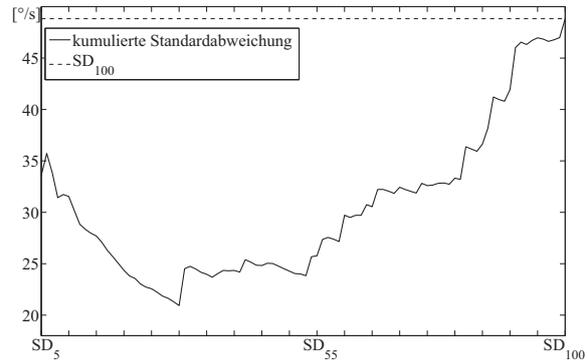


Abbildung 4.1.2:
 Beispiel divergierender kumulierter Standardabweichungen SD_n^{pr} für $p = \omega_{max}$

Es ist erkennbar, dass etwa 60 bis 70 % der Probanden konvergente Standardabweichungskurven aufweisen. Dies sind annähernd 10 bis 15 % mehr konvergierende Kurven als für die Konvergenz des Mittelwertes verzeichnet werden konnten. Des Weiteren wird deutlich, dass die mittlere Anzahl an Versuchen ungefähr 55 % beträgt. Anhand der hohen Standardabweichungen wird jedoch deutlich, dass diese Anzahl enorm schwankt und nicht einheitlich über alle Probanden und Parameter zu bestimmen ist. Für die vier Gruppierungen hinsichtlich des Geschlechtes, der Laborerfahrung, sportlicher Aktivität und des Ausdauerlaufens ergaben sich keine signifikanten Differenzen. Lediglich für den Parameter der minimalen Kraft in der Abbremsphase für die Bodenreaktionskraft in anterior-posterior Richtung benötigten die inaktiven Probanden 30 Versuche weniger als die Testpersonen, welche einer wöchentlichen Lauftätigkeit nachgingen. In Bezug auf die maximale Kraftanstiegsrate ergaben sich für Probanden mit Laborerfahrung 20 Läufe weniger als für laborunerfahrene Probanden, um eine endliche und damit konvergierende Folge kumulierter Standardabweichungen zu erreichen.

	Studie I			Studie II		
	M 1 (#)	M 2 (#)	Differenz	M 1 (#)	M 2 (#)	Differenz
Konvergenz	9 (61 %)	10 (69 %)		12 (61 %)	14 (68 %)	
	54 ± 17	56 ± 16	20 ± 17	53 ± 19	56 ± 20	22 ± 18

Tabelle 4.1.1:
 Zusammenfassung Konvergenz kumulierter Standardabweichungen SD_n^{pr} : Anzahl konvergierender Standardabweichungsfolgen (absolut und prozentualer Anteil) / mittlere Anzahl \pm Standardabweichung / Differenzanzahl zwischen Messtag 1 (M 1) und Messtag 2 (M 2)

Es gibt einige Testpersonen, die vollständig konvergierende Standardabweichungskurven aufweisen, wobei für jene Probanden die Anzahl an Versuchen mit 5 in die Berechnung der mittleren Anzahl einging. Dies trat bei insgesamt 6 Kurven für Studie I und 4 Kurven für Studie II auf. Dabei ist auffällig, dass oftmals mehrere solcher vollständig konvergierender Kurven für eine Testperson regis-

triert wurden⁸. Für keinen Probanden ergaben sich durchweg divergierende Folgen. Fünf Probanden zeigten ausschließlich konvergierende Folgen, wobei drei zum Teil vollständig konvergierende Folgen aufwiesen. Im Großen und Ganzen ließen sich jedoch keine Systematiken feststellen. Im Fall, dass die Standardabweichungskurven sowohl für Messtag 1 als auch für Messtag 2 konvergierende Folgen zeigten (46 % der Probanden-Parameter-Kombinationen für Studie I, 41 % für Studie II), weist die minimale Anzahl an Versuchen zwischen zwei Messtagen erhebliche Diskrepanzen auf (siehe Tabelle 4.1.1). Die an Studie II teilnehmenden sportlich inaktiven Probanden ließen vermuten, dass unter ihnen eine größere Anzahl an divergierenden Folgen oder zumindest hohe Schwankungen zu verzeichnen wären. Dies konnte hier nicht bestätigt werden. Diese sechs Probanden reihen sich in die Unsystematik aller Probanden ein. Für die Gruppe der laborerfahrenen Probanden zeigt sich hingegen ein anderes Bild. Drei der vollständig konvergierenden Standardabweichungsfolgen wurden für zwei der laborerfahrenen Probanden notiert. Des Weiteren konnten für 50 % der divergierenden Folgen kumulierter Standardabweichungen divergierende Mittelwertfolgen festgestellt werden (↗ Tab. C.4, S. 123).

Neben der Betrachtung der Standardabweichung als Indikator für die Variabilität gibt der Root Mean Square Error (RMSE) zudem eine Vorstellung von der absoluten Abweichung innerhalb des gesamten Probandengutes für einen diskreten Parameter p . Im Unterschied zu den bisherigen Auswertungen erhält man eine kumulierte RMSE-Folge für jeden Parameter, in welche die Informationen aller Probanden eingehen. Der RMSE lässt Aussagen bezüglich der Wiederholbarkeit von Messungen zu, wobei man erwartet, dass zwei Messungen um den RMSE (↗ Glg. 2.4.1) schwanken können. Insofern gibt dieser eine gute Vorstellung über die Gesamtschwankung des Probandenkollektivs. Tabelle C.2 (↗ S. 122) gibt einen Überblick über die Analysen bezüglich des RMSE. Angegeben sind der RMSE über alle 100 Kurven, der minimale und der maximale Wert der kumulierten RMSE-Folge sowie die Angabe der Konvergenz (+) und Divergenz (-). Es wird eine Stabilisierung auf dem gleichen Niveau erwartet. Die Angabe der Konvergenz und Divergenz der Folgen offenbart, dass lediglich rund 50 % der RMSE-Folgen konvergieren, wobei keine der Folgen vollständig konvergiert. Des Weiteren ist ersichtlich, dass einige Parameter eine relativ große Gesamtschwankung und eine hohe Spannweite aufweisen. Für die maximale Kraftanstiegsrate ist eine Gesamtschwankung um bis zu 37 bw/s und eine Schwankung des RMSE um 8 bw/s zu verzeichnen. Für die maximale Winkelgeschwindigkeit zeigt sich ein ähnliches Bild. Diese schwankt um bis zu 150°/s und kann innerhalb der Messungen um bis zu 18°/s schwanken (siehe Tabelle 4.1.2). Für die jeweiligen Gruppierungen unterscheiden sich die gefundenen Schwankungen innerhalb des gesamten Probandengutes nicht signifikant und zeigen zudem keine wesentlichen Veränderungen zur gesamten Stichprobe.

Die prozentuale Differenz sowohl innerhalb eines Messtages als auch zwischen mehreren Tagen ist in Tabelle C.3 (↗ S. 123) angegeben. Unter der Bedingung, dass der RMSE über alle 100 gemessenen Läufe die optimale Schwankung innerhalb der Probanden darstellt, gibt diese prozentuale Differenz Aufschluss über die Genauigkeit der Messung des RMSE mit den in der Literatur üblichen fünf Läufen. Es ist deutlich zu erkennen, dass für einzelne Parameter der RMSE erheblich

⁸Streng genommen kann man nicht von Konvergenz sprechen, wie sie in Kapitel 2.4.2 formuliert ist, da ein $\varepsilon > 0$, das die Bedingung in Glg. 2.4.2 erfüllt, so nicht existiert. Im Sinne der Empirie reicht eine Formulierung der Konvergenz über ein festzulegendes Intervall, wie viel Abweichung toleriert wird, jedoch aus. In den hier beschriebenen Fällen wurde die Ausprägung der Standardabweichung als ausreichend minimal angenommen.

	Messtag 1				Messtag 2			
Studie I	$RMSE_{100}^p$	Min	Max		$RMSE_{100}^p$	Min	Max	
FR_{z_I} [bw/s]	32.55	27.77	35.71	–	35.19	33.14	37.52	+
Studie II								
FR_{z_I} [bw/s]	29.03	26.5	30.84	–	29.31	28.85	32.48	+
ω_{max} [°/s]	139.21	130.83	144.58	–	132.17	130.93	149.39	–

Tabelle 4.1.2:

Betrachtung des kumulierten RMSE für FR_{z_I} und ω_{max} : $RMSE_{100}^p$ / minimaler und maximaler RMSE; Angabe der Konvergenz (+ konvergierender RMSE, – divergierender RMSE)

innerhalb und zwischen den Messtagen schwanken kann. Davon sind vor allem die Parameter der Bodenreaktionskraft in anterior-posterior Richtung betroffen. Deren RMSE schwanken bis zu 50 %. Durchschnittlich betrachtet weisen die Differenzen des RMSE geringe prozentuale Abweichungen auf (siehe Tabelle 4.1.3) und können zumindest für einige Parameter reliabel gemessen werden.

	Studie I			Studie II		
	DE_{intra}^p	DE_{inter}^p		DE_{intra}^p	DE_{inter}^p	
	M 1 (%)	M 2 (%)	(%)	M 1 (%)	M 2 (%)	(%)
RMSE	18 ± 17	11 ± 10	12 ± 13	8 ± 9	15 ± 9	10 ± 6

Tabelle 4.1.3:

Betrachtung der Differenz des RMSE DE_{intra}^p und DE_{inter}^p

Abbildung 4.1.3 zeigt den Verlauf der kumulierten Folge der RMSE für die maximale Winkelgeschwindigkeit ω_{max} . Für diesen Parameter wird die erhebliche Schwankung innerhalb eines Tages sichtbar, wobei der RMSE aus fünf Läufen dem aus 100 Läufen in etwa gleicht.

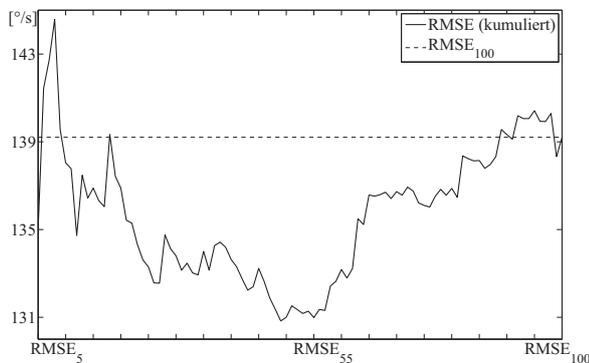


Abbildung 4.1.3:

Beispiel einer kumulierten RMSE-Folge (ω_{max})

4.2 Betrachtung der Gesamtkurve

Zunächst werden die kumulierten Standardabweichungskurven auf deren Konvergenz überprüft. Zu diesem Zweck werden diese kumulierten Kurven jeweils punktweise für $m = 1, \dots, 101$ auf die letzte Standardabweichung SD_{m100}^{kr} normiert. Tabelle 4.2.1 zeigt für jede Kurve k die Anzahl an Probanden, für welche die punktweise betrachteten kumulierten SD-Folgen konvergieren und die entsprechende Anzahl an Läufen, ab wann diese Konvergenz einsetzt. Dabei ist auffällig, dass im Gegensatz zu den Mittelwerten für fast alle Probanden konvergente Standardabweichungskurven vorliegen, wobei keine Systematiken hinsichtlich der entsprechenden Anzahl für die vier untersuchten Gruppierungen vorliegen.

p	Studie I		Studie II	
	Messtag 1 (#)	Messtag 2 (#)	Messtag 1 (#)	Messtag 2 (#)
F_z	14 56 ± 13	14 54 ± 13	19 57 ± 14	19 60 ± 11
F_y	13 58 ± 8	14 60 ± 12	20 64 ± 9	20 61 ± 13
F_x	14 65 ± 12	14 63 ± 8	19 66 ± 9	20 56 ± 11
ω			18 71 ± 6	20 71 ± 7

Tabelle 4.2.1:

Visuelle Überprüfung der Konvergenz der punktweise kumulierten SD-Kurven: Anzahl konvergierender Probanden (mittlere minimale Anzahl \pm Standardabweichung)

Abbildung 4.2.1 zeigt ein Beispiel für derartig konvergierende punktweise betrachtete kumulierte SD-Folgen. Lediglich zwei Probanden entwickelten divergierende SD-Folgen jeweils für die Kurve der Winkelgeschwindigkeit (siehe zum Beispiel Abbildung 4.2.2). Dabei sei angemerkt, dass deren gemittelte Kurven der Winkelgeschwindigkeit ebenso divergieren.

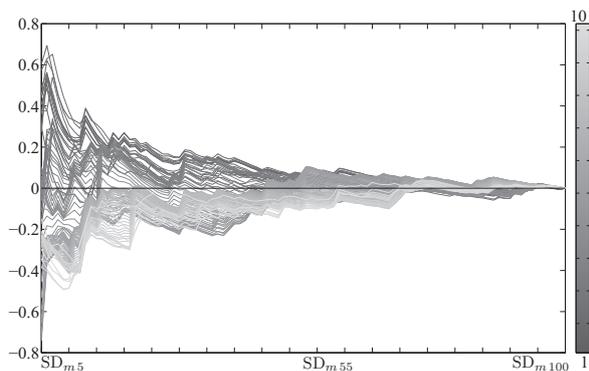


Abbildung 4.2.1:

Beispiel punktweise konvergierender kumulierter SD-Kurven für F_z (Konvergenz ab 50 Werten)

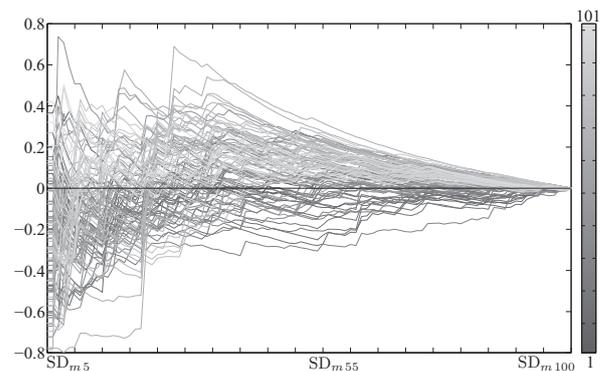


Abbildung 4.2.2:

Beispiel punktweise divergierender kumulierter SD-Kurven für F_z

Bei den vier restlichen divergierenden Kurven handelt es sich jeweils um das gleiche Phänomen (siehe Abbildung 4.2.3). Diese weisen konvergierende Standardabweichungskurven lediglich in einem

bestimmten Intervall auf. Die häufig letzten zehn Mittelwertkurven ändern die Standardabweichung derart extrem, dass keine über die gesamten Läufe konvergierenden Folgen entstehen. Dies tritt bei verschiedenen Probanden und innerhalb der drei Bodenreaktionskräfte auf. Bei zwei dieser Testpersonen divergiert zudem die zugehörige Mittelwertkurve. Bei einem Vergleich der mittleren Anzahl an Versuchen der Standardabweichungskurve für Probanden, deren Mittelwertkurve divergiert, wurde tendenziell eine leicht höhere Anzahl an notwendigen Versuchen festgestellt. Um bis zu 14 Versuche mehr waren notwendig, wenn divergierende Mittelwertkurven vorlagen. Für drei der 14 untersuchten Kurven (F_{ySIM1} , F_{xSIM2} , ω_{SIM1}) konnte diese Tendenz jedoch nicht festgestellt werden.

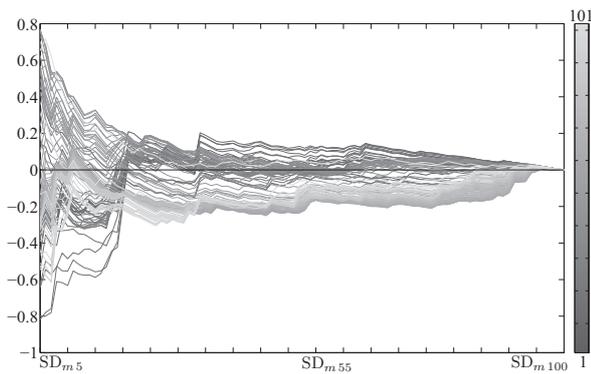


Abbildung 4.2.3:
Beispiel einer teilweise punktweise konvergierenden kumulierten SD-Folge für ω (Konvergenz zwischen 35 und 90 Läufen)

Des Weiteren ist erkennbar, dass die vorliegende Konvergenz ab einer relativ stabilen Anzahl erreicht wird. Es reichen rund 60 bis 65 Versuche aus, um stabile Standardabweichungskurven zu erhalten. Die geringen Standardabweichungen deuten zudem darauf hin, dass diese Anzahl relativ gleich über alle Probanden erreicht wird.

Die Abbildungen 4.2.4 und 4.2.5 zeigen jeweils ein Beispiel für die 96 kumulierten Standardabweichungskurven der vertikalen Bodenreaktionskraft und der Winkelgeschwindigkeit für einen Proband. Die vier Gruppierungen (Geschlecht, Ausdauerlauf, sportliche Aktivität, Laborerfahrung) zeigten keine Auffälligkeiten bezüglich der visuell bestimmten minimalen Anzahl.

Eine weitere Möglichkeit, die Streuungskomponente des Probanden zu betrachten, ist die Analyse der Prädiktionsbänder für die gesamten 100 Versuche eines Probanden pro Kurve. In dieses Intervall sollten bei wiederholter Messung die gemessenen Kurven fallen. Abbildung 4.2.6 und 4.2.7 zeigen jeweils ein Prädiktionsband für die 100 gemessenen Kurven eines Probanden für die vertikale Bodenreaktionskraft und die Winkelgeschwindigkeit. Es ist zu erkennen, dass dieses Prädiktionsband die Gesamtfläche fast aller gemessenen Kurven überdeckt. Je variabler die gemessenen Kurven ausfallen, um so breiter wird das Prädiktionsband und um so größer die Fläche, die das Prädiktionsband überdeckt.

Zum Vergleich innerhalb eines Messtages und zwischen den Messtagen wird die Fläche betrachtet, welche nicht von den beiden involvierten Prädiktionsbändern überdeckt wird. Für die prozentuale Abweichung wird diese Differenzenfläche auf die mittlere Fläche der beiden Prädiktionsbänder normiert. In Tabelle 4.2.2 sind die prozentualen Abweichungen sowohl innerhalb eines Tages ($DAPB_{intra}^{kr}$) als auch zwischen zwei Messtagen ($DAPB_{inter}^{kr}$) aufgelistet. Angegeben sind die mittleren Prozentwerte über alle Probanden und die mittlere Standardabweichung. Es ist erkennbar,

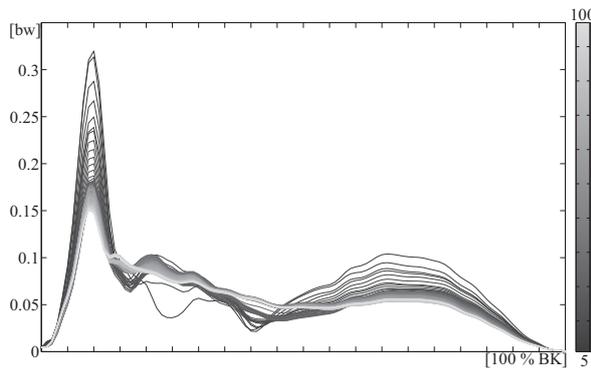


Abbildung 4.2.4:
Beispiel kumulierter Standardabweichungskurven SD_{mn}^{kr} ($k = F_z$)

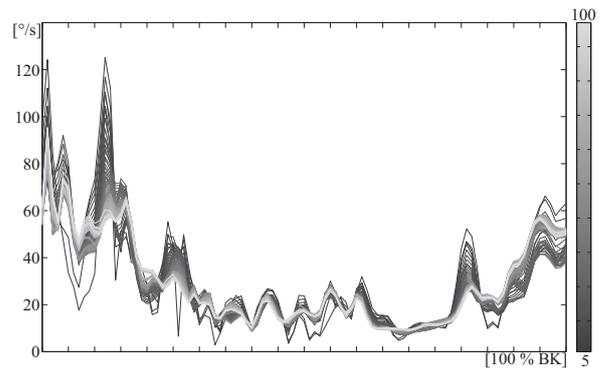


Abbildung 4.2.5:
Beispiel kumulierter Standardabweichungskurven SD_{mn}^{kr} ($k = \omega$)

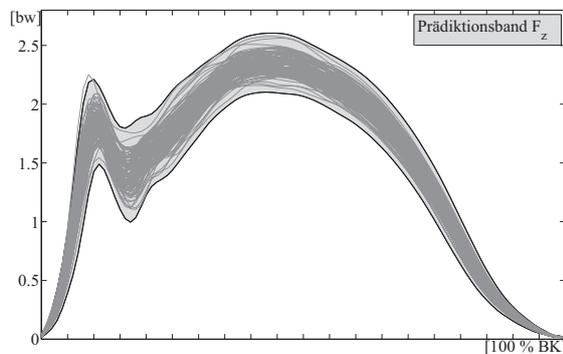


Abbildung 4.2.6:
Prädiktionsband $[PB_{100u}^{kr}(t); PB_{100o}^{kr}(t)]$ eines Probanden r für $k = F_z$ (mit allen F_z -Kurven)

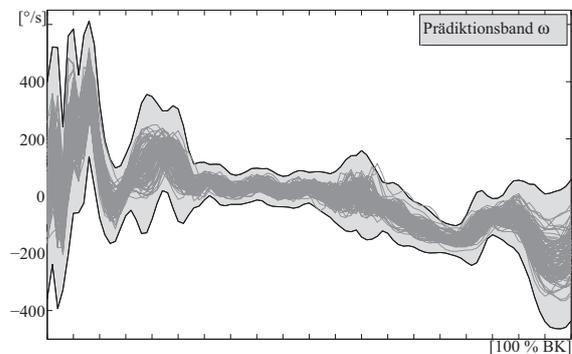


Abbildung 4.2.7:
Prädiktionsband $[PB_{100u}^{kr}(t); PB_{100o}^{kr}(t)]$ eines Probanden r für $k = \omega$ (mit allen Kurven der Winkelgeschwindigkeit)

dass die prozentuale Abweichung innerhalb eines Messtages enorm hoch ist. Für die jeweiligen Gruppierungen hinsichtlich des Geschlechtes, der Laborerfahrung, der sportlichen Aktivität und des Ausdauerlaufens zeigt lediglich die Bodenreaktionskraft in mediolateraler Richtung eine Veränderung von 175 % auf 160 % sowohl für aktive Probanden als auch für jene Testpersonen, die sich gegen das durchgängige Laufen im Labor entschieden hatten.

Die Abbildungen 4.2.8 und 4.2.9 zeigen diese hohen Unterschiede innerhalb eines Messtages beispielhaft für die vertikale Bodenreaktionskraft und die Winkelgeschwindigkeit eines Probanden. Dabei sei angemerkt, dass Werte kleiner Null innerhalb der Prädiktionsbänder der vertikalen Bodenreaktionskraft abgeschnitten wurden, da eine negative vertikale Kraft nicht auftreten kann. Es ist eindeutig zu erkennen, dass die Prädiktionsbänder für die ersten fünf Läufe erheblich größer sind, als die Prädiktionsbänder für alle 100 Versuche.

Die Differenzenflächen zwischen den Messtagen für die Bodenreaktionskräfte sind hingegen relativ

Studie I			
	$DAPB_{intraM1}^{kr}$ (%)	$DAPB_{intraM2}^{kr}$ (%)	$DAPB_{inter}^{kr}$ (%)
F_z	161 ± 15	167 ± 13	22 ± 9
F_y	167 ± 7	173 ± 12	21 ± 7
F_x	168 ± 11	168 ± 23	20 ± 6
Studie II			
	$DAPB_{intraM1}^{kr}$ (%)	$DAPB_{intraM2}^{kr}$ (%)	$DAPB_{inter}^{kr}$ (%)
F_z	160 ± 16	165 ± 13	24 ± 7
F_y	171 ± 11	172 ± 14	24 ± 6
F_x	164 ± 17	171 ± 13	16 ± 5
ω	170 ± 16	175 ± 10	28 ± 12

Tabelle 4.2.2:

Prozentuale Abweichung der Fläche des Prädiktionsbandes bezüglich der mittleren Fläche: mittlere prozentuale Abweichung ± Standardabweichung

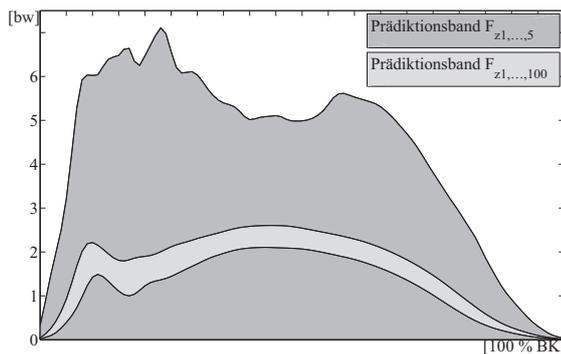


Abbildung 4.2.8:

Prädiktionsband $[PB_{5u}^{kr}(t); PB_{5o}^{kr}(t)]$ und $[PB_{100u}^{kr}(t); PB_{100o}^{kr}(t)]$ eines Probanden für die vertikale Bodenreaktionskraft $k = F_z$

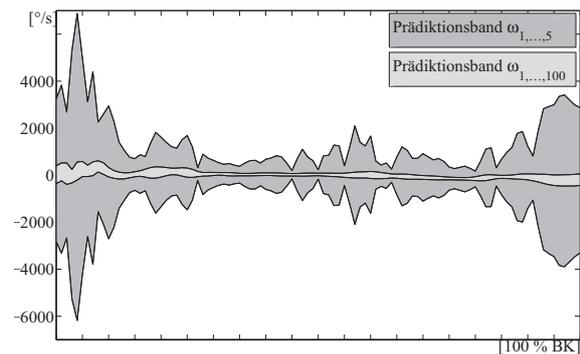
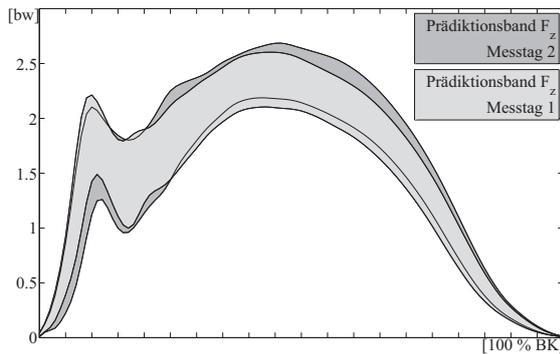


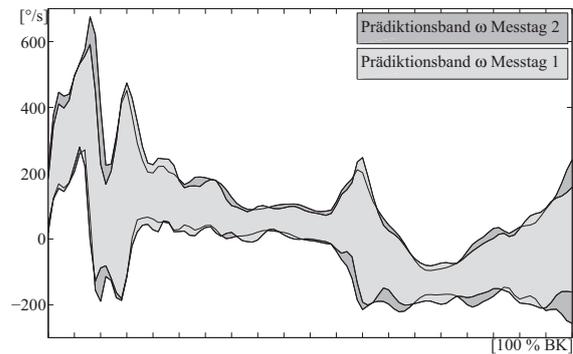
Abbildung 4.2.9:

Prädiktionsband $[PB_{5u}^{kr}(t); PB_{5o}^{kr}(t)]$ und $[PB_{100u}^{kr}(t); PB_{100o}^{kr}(t)]$ eines Probanden für die Winkelgeschwindigkeit $k = \omega$

stabil messbar. Die Abweichungen um 16 % bis 24 % fallen kaum ins Gewicht (siehe auch Abbildung 4.2.10). Für die Winkelgeschwindigkeit wurde eine Abweichung innerhalb von zwei Messtagen von 24 % ermittelt, wobei deren Standardabweichung relativ hoch erscheint. Zwei Probanden wiesen eine prozentuale Abweichung bis 60 % auf. Abbildung 4.2.11 zeigt beispielhaft die Prädiktionsbänder der Winkelgeschwindigkeit für Messtag 1 und Messtag 2. Lediglich für die Bodenreaktionskraft in mediolateraler Richtung wurde ein signifikanter Unterschied von -4% für die Gruppe mit Labinerfahrung ermittelt. Für die Bodenreaktionskraft in anterior-posterior Richtung wurde ein signifikanter Unterschied von -8% für die Gruppe ohne wöchentliches Laufpensum ermittelt. Beide Unterschiede erscheinen im Hinblick auf die tatsächliche Größenordnung der Ausprägung der F_y - und F_x -Kurven als nicht relevant.


Abbildung 4.2.10:

Prädiktionsband $[PB_{100u}^{kr}(t); PB_{100o}^{kr}(t)]$ für Messtag 1 und Messtag 2 eines Probanden für die vertikale Bodenreaktionskraft $k = F_z$


Abbildung 4.2.11:

Prädiktionsband $[PB_{100u}^{kr}(t); PB_{100o}^{kr}(t)]$ für Messtag 1 und Messtag 2 eines Probanden für die Winkelgeschwindigkeit $k = \omega$

4.3 Diskussion der Ergebnisse

Mit Hilfe der Analysen zu Hypothese II war die Frage zu klären, ob die Variabilität eines Probanden endlich ist und vollständig charakterisiert werden kann. Bezüglich der visuellen Betrachtung der Konvergenz konnten für 60 bis 70 % der Probanden konvergente Folgen kumulierter Standardabweichungen registriert werden. Dies sind 10 bis 15 % mehr konvergente Standardabweichungsfolgen als Mittelwertfolgen. In 50 % der divergierenden SD-Folgen ist die Ursache in divergierenden zugehörigen Mittelwertfolgen zu finden. Die Gründe für die Divergenz der anderen Folgen und der damit verbundenen zumeist heteroskedastischen Daten kann nicht genau anhand des Datensatzes geklärt werden. Die möglichen Erklärungsansätze anhand verschiedener Probandengruppen müssen verworfen werden, da über alle diese Gruppierungen unsystematische Verteilungen der Testpersonen mit divergierenden Folgen zu finden sind.

Die hohen Schwankungen der Anzahl, ab wann die kumulierten Folgen der Standardabweichungen konvergieren, deutet auf eine individuell zu bestimmende minimale Anzahl hin, wobei die hohe Schwankung dieser Anzahl zwischen den verschiedenen Messtagen zudem darauf hindeutet, dass es keine einheitliche minimale Anzahl für einen Probanden gibt. Dies ist bedenklich, da bei variierenden Häufigkeiten die Bestimmung dieser Anzahl nicht mit einer Probemessung möglich ist, sondern interaktiv festgelegt werden muss. Die kaum vorhandenen Systematiken der kumulierten SD-Folgen weder bezüglich der Gruppierungen noch des Verlaufes dieser Folgen deuten auf die Dominanz der Individualität jedes Probanden hin. Dies erschwert zudem die Bestimmung einer minimalen Anzahl, da auf kein einheitliches Muster über alle Probanden zurückgegriffen werden kann, und unterstreicht die Notwendigkeit der Bestimmung einer minimalen Anzahl an Versuchen für jeden Proband anhand eines Algorithmus während der Datenaufnahme.

Die Existenz von vollständig konvergierenden Folgen kumulierter Standardabweichungen für diskrete Parameter und deren Häufung zum Teil auf einen Probanden gibt einen Hinweis, dass für einige Probanden die Variabilität durchaus endlich sein kann. Im Hinblick auf eine Verallgemeinerbarkeit sind knapp 4 % zu wenig. Die Anzahl an Probanden mit sowohl an Messtag 1 als auch

an Messtag 2 konvergierenden SD-Folgen ist mit 46 % und 41 % höher als für kumulierte Mittelwertfolgen. Dies sind dennoch relativ wenig Probanden-Parameter-Kombinationen. Dieser Umstand deutet darauf hin, dass nur für wenige Probanden eine reliable Messung der Variabilität zwischen mehreren Messtagen möglich ist.

Die Betrachtung der Folgen der kumulierten Root Mean Square Errors über alle Probanden unterstreicht diese Vermutung. Lediglich 50 % dieser Kurven konvergieren, das heißt, für die Hälfte der diskreten Parameter konnte keine stabile Abweichung zwischen den Probanden gefunden werden. Einige Parameter, vor allem die maximale Kraftanstiegsrate und die Winkelgeschwindigkeit, konnten sowohl innerhalb als auch zwischen den Tagen relativ stabil gemessen werden.

Die hohen absoluten Schwankungen, vor allem für die maximale Kraftanstiegsrate (37 bw/s) und die Winkelgeschwindigkeit (150 °/s), sowie die absoluten Abweichungen eines Probanden innerhalb eines Messtages von 32 bw/s respektive 110 °/s (↗ Tab. B.5, S. 116) werfen die Frage auf, inwiefern in der Literatur angegebene signifikante Unterschiede für diese Parameter bewertet werden müssen. Ein Blick in eigene Vorarbeiten zeigt zudem, dass die Abweichungen, ermittelt in dieser Studie, nicht zufällig sind, wobei Schwankungen für das dortige Probandenkollektiv von 32 bw/s für die maximale Kraftanstiegsrate und 169 °/s für die maximale Winkelgeschwindigkeit anhand von fünf Läufen ermittelt wurden (Oriwol et al., 2011). Diese unterscheiden sich nur unwesentlich von den hier berechneten Abweichungen. Signifikante und für relevant befundene Differenzen der jeweiligen Untersuchungsbedingungen von zum Beispiel 173 °/s (Brauner et al., 2009), 71.8 °/s (Mündermann et al., 2003) und 98 °/s (Milani et al., 1995) müssten daher einer erneuten Relevanzprüfung unterzogen werden. Mit den hier ermittelten Schwankungen fallen die ermittelten Differenzen in den Bewegungsspielraum eines Probanden und können streng genommen nicht der untersuchten Bedingung zugeschrieben werden. Zukünftige Untersuchungen sollten daher stets die Schwankung des Probandenkollektivs im Blick behalten und die Relevanz der signifikanten Unterschiede damit vergleichen und beurteilen (Maiwald et al., 2011).

Schlussendlich muss für diskrete Parameter die Hypothese der Endlichkeit und der Stabilität der Variabilität über alle Probanden abgelehnt werden.

30 bis 40 % divergierende Folgen kumulierter Standardabweichungen, hohe Schwankungen innerhalb der Probanden und zum Teil hohe Differenzen sowohl der ermittelten Anzahl der Konvergenz als auch des RMSE zwischen den Messtagen verdeutlichen die Ablehnung. Als Konsequenz aus der Zurückweisung der Hypothesen I und II für die verwendeten diskreten Parameter muss von einer Entwicklung eines Algorithmus zur Bestimmung einer minimalen Anzahl (Hypothese III) abgesehen werden, da die generelle Verwendung von diskreten Parametern (zumindest des Mittelwertes als statistischem Vertreter) mit dieser Negation in Frage gestellt wird.

Ein Blick in die Literatur bezüglich der Ermittlung einer minimalen Anzahl zeigt (siehe auch Kapitel 1.3), dass häufig diskrete Parameter innerhalb der Untersuchungen verwendet wurden. Die von Diss (2001) und Laroche et al. (2011) vorgeschlagenen Algorithmen unter der Benutzung der Spearman-Brown-Formel mit einem vorgegebenen Reliabilitätslevel sowie einem am Verfahren von Bland & Altman orientierten Modell können daher nicht verwendet werden.

Hinsichtlich der Standardabweichungskurven der gesamten analysierten Zeitreihen ergibt sich ein anderes Bild. Fast alle Probanden weisen konvergente kumulierte punktweise betrachtete SD-

Kurven auf. Dies zeugt von einer Endlichkeit der Variabilität. Die vier Kurven, welche das Phänomen einer Intervall-Konvergenz zeigen, müssen gesondert betrachtet werden. Hierbei stellt sich die Frage, was den Probanden innerhalb der letzten Läufe beeinflusst hat, dass die letzten zehn Kurven einen derart enormen Einfluss auf die Standardabweichung ausüben konnten. Ursachen könnten in einer Erschöpfung des Probanden liegen, wobei kein Proband auch bei direkter Nachfrage explizit über eine Ermüdung geklagt hat. Eine Beeinflussung der Geschwindigkeit oder zumindest niedrigere Geschwindigkeiten innerhalb der Messungen konnten nicht festgestellt werden. Des Weiteren könnten die Probanden einen Motivationsschub zum Ende der Messungen erfahren haben, so dass sich der Laufstil und damit verbunden die Bodenreaktionskraftkurven unbewusst geändert haben. Weitere Ursachen liegen unter Umständen in einer plötzlichen Lärmbelastung in den Laborräumen oder in anderen Einflüssen, welche die Konzentration beeinflusst haben. Dies kann durchaus während der gesamten Messung aufgetreten sein, wobei anhand der Standardabweichungskurven ungewöhnliche Sprünge möglicherweise beobachtet, aber nicht als phänomenale Abweichung deklariert wurden.

Die relativ stabile Anzahl von 60 bis 65 Läufen deutet auf die generelle Möglichkeit hin, eine minimale Abweichung für alle Probanden zu finden. Die lediglich visuelle Betrachtung der Standardabweichungskurven und der Bestimmung der mittleren Anzahl, ab wann die SD-Kurven konvergieren, ist dabei subjektiv behaftet und bedarf eines objektiven Algorithmus.

Die Untersuchung der Variabilität anhand von Prädiktionsbändern zeigt, dass diese innerhalb eines Tages extrem abweicht. Dies untermauert die Annahme, dass mehr als die bisher üblichen durchschnittlichen fünf Läufe pro Proband aufgenommen werden müssen. Dabei ist zu beachten, dass die Verwendung des Bootstrap-Verfahrens auf eine hohe Rechenkapazität des eingesetzten Computers angewiesen ist. Zum einen muss die Anzahl an Koeffizienten adäquat gewählt sein, um anhand der Fouriertransformation eine annähernd korrekte Darstellung der originären Kurven zu garantieren und den Anpassungsfehler zu minimieren. Zum anderen schränkt eine Interpolation der Kurven mit mehr als 101 Datenpunkten die Anzahl der Wiederholungen B innerhalb des Verfahrens ein. Im Sinne der Quantilsbestimmung für die Breite des Prädiktionsbandes ist B der zu optimierende Faktor. Je größer B gewählt wird, um so genauer kann der Mittelwert der entstehenden Stichprobe geschätzt werden. Die iterative Herangehensweise bei der Berechnung des entsprechenden Quantils erhöht zudem den Rechenaufwand. Der entstehende zeitliche Aufwand ist von vornherein nicht abzuschätzen, wobei für die Bestimmung eines Quantils in Abhängigkeit des vorhandenen Arbeitsspeichers zwischen 10 und 40 Sekunden eingeplant werden müssen.

Des Weiteren sollte nach Albers et al. (2007) die Stichprobe, welche für das Bootstrap-Verfahren herangezogen wird, mindestens 10, optimalerweise mindestens 25 Kurven beinhalten. Die Betrachtung der Differenzflächen der sich nicht überlagernden Prädiktionsbänder für 10 und 100 Kurven $DAPB_{10|100}^{kr}$ (↗ Tab. C.5, S. 124) zeigt dennoch deutlich, dass mehr als 10 Versuche aufgezeichnet werden sollten, um ein annähernd gleiches Prädiktionsband und damit verbunden eine stabile Charakterisierung der Variabilität eines Probanden zu erreichen. Die prozentualen Differenzen für 25 und 100 involvierte Kurven $DAPB_{25|100}^{kr}$ demonstriert eine beginnende Stabilisierung der Flächeninhalte. Die Differenz der Prädiktionsbänder zwischen den Messtagen deutet zudem darauf hin, dass sowohl die Bodenreaktionskräfte als auch die Winkelgeschwindigkeit reliabel über mehrere Tage messbar sind.

Die Hypothese II bezüglich der Endlichkeit der Variabilität für die gesamten untersuchten Kurven ist somit bestätigt. Die Konvergenz fast aller Standardabweichungskurven und die gute Abschätzung der Prädiktionsbänder zwischen mehreren Messtagen verifizieren diese Aussage.

Die in Hypothese III zu verwendenden Algorithmen sollten sich an der Variabilität der gesamten Kurven orientieren, da sich Hypothese I über die ausreichende Charakterisierung der Probanden anhand der gesamten Mittelwertkurve zum Teil als haltlos erwiesen hat. Als Grundlage für die Algorithmen dienen die Differenzen der Standardabweichungskurven und die Fläche der Prädiktionsbänder. Die Entwicklung eines Algorithmus, der während der interessierenden Messung applizierbar ist, birgt den Vorteil, dass eine eventuelle Probemessung für den Proband entfallen würde. Daher sollten die im Folgenden vorgestellten Algorithmen auf interaktive Eignung überprüft werden.

Kapitel 5

Resultate zu Hypothese III: Algorithmen zur Bestimmung einer minimalen Anzahl

Ausgehend von den Auswertungen in Kapitel 3 und 4 konnte lediglich Hypothese II für die gesamten Kurven der Bodenreaktionskräfte und der Winkelgeschwindigkeit bestätigt werden. Dieser Umstand lässt für eine Bestimmung der minimal aufzuzeichnenden Anzahl an Messversuchen nicht viel Spielraum, so dass sich die Untersuchungen an den bereits in Abschnitt 4.2 eingeführten Methoden orientieren. Zunächst wird die Variabilität anhand der Fläche der Prädiktionsbänder zur Bestimmung der minimalen Anzahl herangezogen. Eine weitere Methode zur Bestimmung einer Mindestanzahl an Versuchen besteht in der Verwendung der Differenzen der Standardabweichungskurven und einer entsprechenden Populationsdifferenzenkurve.

Die Auswertungen zu Hypothese I und II zeigen, dass keine entscheidenden Unterschiede zwischen den Probanden unterschiedlichen Geschlechtes, mit oder ohne wöchentlichem Laufpensum, mit oder ohne Laborerfahrung oder der Verwendung des Ausdauerlaufens während der Messung ermittelt wurden. Ein plötzlicher Wechsel zu erheblichen relevanten Unterschieden für die Berechnung der minimal notwendigen Versuche für jeden Proband ist daher nicht zu erwarten und es wird im Folgenden auf eine Unterteilung bezüglich der vier Gruppierungen verzichtet.

5.1 Minimale Anzahl unter Verwendung von Prädiktionsbändern

Der Vergleich der Differenzen der Flächen der Prädiktionsbänder innerhalb eines Tages (↗ Tab. 4.2.2) ließ die Vermutung zu, dass mit zunehmender Anzahl das Ausmaß dieser Fläche geringer wird. Die Annahme des exponentiell abnehmenden Verhaltens der kumulierten Folge APB_N^{kr} konnte anhand visueller Überprüfung bestätigt werden (↗ Abb. 5.1.1). Dies ermöglichte die nichtlineare Anpassung von APB_N^{kr} anhand der in Abschnitt 2.5.2 vorgestellten Modelle 1 und 2. In Tabelle 5.1.1 sind jeweils die Wurzeln aus den Fehlervarianzen der Anpassung der Flächen der Prädiktionsbänder APB_N^{kr} sowie der normierten Flächen $APBZ_N^{kr}$ an das vorgegebene Modell dargestellt.

Es ist deutlich zu erkennen, dass die Kurve der Winkelgeschwindigkeit ω anhand der originären Flächeninhalte der Prädiktionsbänder mit keiner der Methoden optimal bestimmt werden konnte, erst nach Normierung wurde die entsprechende Anpassung verbessert. Die Fehlervarianzen für Modell 2 sind durchschnittlich leicht geringer als für Modell 1. Die Anpassung anhand Modell 2 ist daher vorzuziehen und die Bestimmung der minimal zu messenden Anzahl an Versuchen wird anhand des maximalen Krümmungspunktes von Modell 2 ermittelt. Tabelle 5.1.2 gibt einen Überblick über die mittleren Anzahlen für jede gemessene Kurve der Bodenreaktionskraft und der

Studie I		APB_N^{kr}		$APBZ_N^{kr}$	
k		Modell 1	Modell 2	Modell 1	Modell 2
F_z	M 1	6.45 ± 3.00	5.31 ± 2.73	0.17 ± 0.08	0.13 ± 0.05
	M 2	6.71 ± 3.35	6.65 ± 4.99	0.17 ± 0.07	0.15 ± 0.07
F_y	M 1	2.36 ± 0.80	1.81 ± 0.76	0.15 ± 0.06	0.12 ± 0.05
	M 2	3.22 ± 1.33	2.28 ± 1.43	0.21 ± 0.08	0.16 ± 0.08
F_x	M 1	2.15 ± 0.61	1.58 ± 0.87	0.17 ± 0.05	0.13 ± 0.06
	M 2	2.18 ± 0.97	1.47 ± 0.73	0.19 ± 0.10	0.14 ± 0.08
Studie II		APB_N^{kr}		$APBZ_N^{kr}$	
k		Modell 1	Modell 2	Modell 1	Modell 2
F_z	M 1	5.73 ± 3.48	4.98 ± 4.10	0.15 ± 0.09	0.12 ± 0.08
	M 2	6.08 ± 2.76	4.30 ± 2.53	0.18 ± 0.08	0.13 ± 0.07
F_y	M 1	3.00 ± 2.14	2.36 ± 1.46	0.21 ± 0.16	0.17 ± 0.13
	M 2	2.87 ± 1.30	1.38 ± 0.97	0.23 ± 0.11	0.15 ± 0.09
F_x	M 1	1.98 ± 1.18	1.72 ± 1.36	0.19 ± 0.11	0.15 ± 0.10
	M 2	1.99 ± 0.76	1.44 ± 0.65	0.20 ± 0.09	0.15 ± 0.07
ω	M 1	4351.48 ± 1912.96	16875.45 ± 12639.19	0.21 ± 0.10	0.20 ± 0.10
	M 2	4722.24 ± 2355.45	12564.15 ± 6466.29	0.26 ± 0.12	0.18 ± 0.09

Tabelle 5.1.1:

Güte der Anpassung der nichtlinearen Regression anhand der Wurzel aus der Fehlervarianz σ_ε^2 für Studie I und II

Winkelgeschwindigkeit, wobei für ω lediglich die transformierte Kurve verwendet wurde.

Studie I		APB_N^{kr}		$APBZ_N^{kr}$	
k		Messtag 1	Messtag 2	Messtag 1	Messtag 2
F_z		13 ± 2.6	19 ± 5.6	8 ± 0.5	8 ± 0.5
		12 ± 1.1	13 ± 2.7	8 ± 0.4	8 ± 0.5
F_y		11 ± 1.4	12 ± 2.1	8 ± 0.4	8 ± 0.7
Studie II		APB_N^{kr}		$APBZ_N^{kr}$	
k		Messtag 1	Messtag 2	Messtag 1	Messtag 2
F_z		13 ± 3.1	16 ± 3.0	8 ± 0.7	8 ± 0.6
		13 ± 1.9	12 ± 1.2	8 ± 0.7	8 ± 0.6
F_y		12 ± 1.6	12 ± 1.8	8 ± 0.8	8 ± 0.6
				9 ± 0.6	9 ± 0.6
F_x					
ω					

Tabelle 5.1.2:

Minimal zu messende Anzahl an Versuchen bestimmt anhand Krümmungspunkt der angepassten Kurve der Folge der Flächen der Prädiktionsbänder

Für die kumulierte Folge der absoluten Flächen APB_N^{kr} konnte eine mittlere minimale Anzahl von 11 bis 19 Versuchen festgestellt werden. Die geringen Standardabweichungen deuten darauf hin, dass diese Werte relativ stabil für die Probanden ermittelt werden konnten. Ein Blick auf die Mindestanzahl, bestimmt anhand der transformierten Folge der Flächen der Prädiktionsbänder $APBZ_N^{kr}$, zeigt eine sehr stabile Mindestanzahl von 8 Versuchen. Dies offenbart die Abhängigkeit der Berechnungsmethode vom Ausmaß der jeweiligen Folge. Für die absoluten Flächen wurden viel größere Werte verzeichnet als für die transformierten Flächen. Dies wurde durch die Normierung auf die letzte ermittelte absolute Fläche APB_{100}^{kr} bedingt. Die Abbildungen 5.1.1 und 5.1.2 zeigen zwei Beispiele von Folgen der absoluten Fläche der Prädiktionsbänder. Zum einen wird das große Ausmaß dieser Folgen deutlich und zum anderen die durch dieses Ausmaß verbundene trügerische Qualität der Anpassung mittels nichtlinearer Regression.

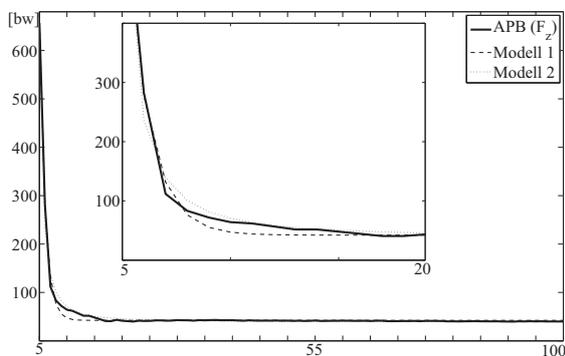


Abbildung 5.1.1:

Anpassung von APB_N^{kr} mittels Modell 1 und Modell 2 für F_z (mit Ausschnitt der ersten 5 bis 20 kumulierten Werte)

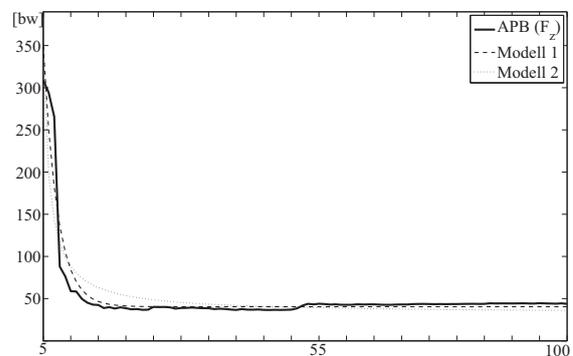


Abbildung 5.1.2:

Anpassung von APB_N^{kr} mittels Modell 1 und Modell 2 für F_z (der vorhergesagte Verlauf der Kurve ist größtenteils nicht gegeben)

Die erste Grafik zeigt eine vergleichsweise gute Qualität der Anpassung vor allem von Modell 2. Bei genauer Betrachtung des vergrößerten Abschnittes ist zu erkennen, dass die Anpassung den interessierenden, tatsächlichen Krümmungspunkt der Folge etwas überschätzt. Abbildung 5.1.2 zeigt zudem einen Probanden, für den der vorhergesagte exponentielle Verlauf der Folge nicht gegeben ist und damit verbunden die Qualität der Anpassung für beide Modelle eher gering ausfällt. Vor allem ab 20 Werten ist eine intervallartige Zunahme der Werte erkennbar. Abbildung 5.1.3 zeigt beispielhaft die transformierten Folgen von APB_N^{kr} für die vertikale Bodenreaktionskraft aller Probanden für Messung 1 (Studie II). Die Annahme des exponentiellen Verhaltens kann ab 15 Läufen für einige Probanden nicht mehr bestätigt werden.

Bei Betrachtung der Differenzen der sich nicht überlagernden Flächen der Prädiktionsbänder $DAPB_N^{kr}$ zweier aufeinander folgender Läufe wird deutlich, dass einige Läufe das absolute Ausmaß der Prädiktionsbänder wesentlich verändern. Lediglich ein Prozent der Probanden weisen keine Läufe auf, welche eine derartige Veränderung verursachen. Abbildung 5.1.4 zeigt die korrespondierende Folge der Differenzen der Fläche der Prädiktionsbänder für den Proband aus Abbildung 5.1.2. Sowohl Lauf 25 als auch Lauf 52 verändern die Spannweite und damit verbunden das Prädiktionsband derart extrem, dass dies in einer Zunahme der Fläche resultiert. Die Grafiken 5.1.5 und 5.1.6 zeigen

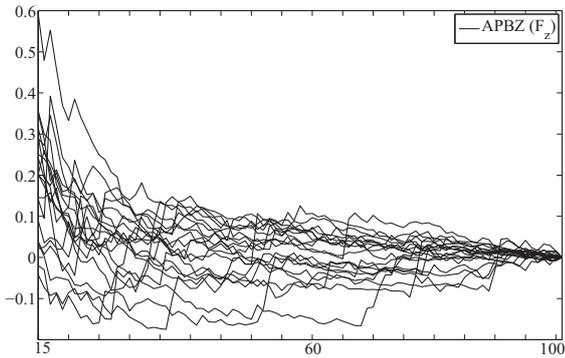


Abbildung 5.1.3:
Kumulierte Folgen von $APBZ_N^{kr}$ für alle Probanden r der vertikalen Bodenreaktionskraft ab 15 Läufen

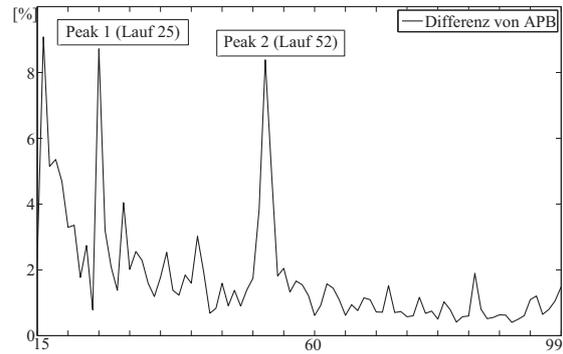


Abbildung 5.1.4:
Folge der Differenz der sich nicht überlagernden Flächen zweier aufeinander folgender Prädiktionsbänder $DAPB_N^{kr}$ ab 15 Läufen für APB_N^{kr} aus Abbildung 5.1.2 (F_z)

die entsprechenden Kurven und ihre Auswirkung auf das Prädiktionsband.

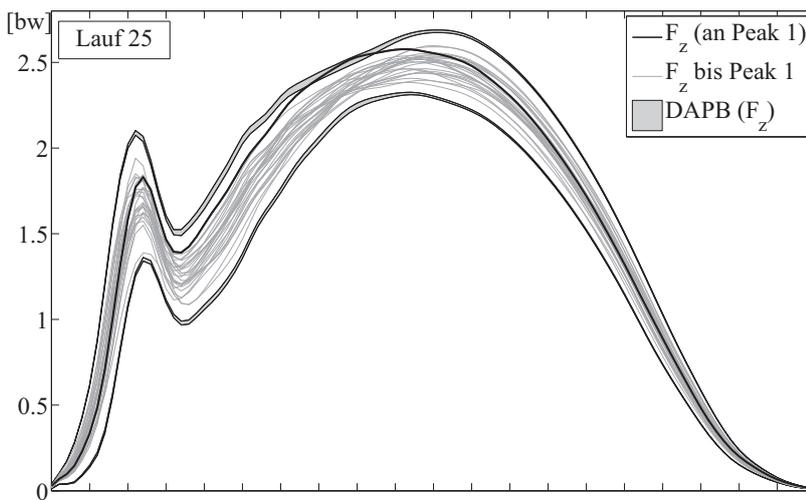


Abbildung 5.1.5:
Auswirkung des Laufs 25 (Peak 1) auf das Prädiktionsband entsprechend der Abbildung 5.1.4

Die grauen Kurven entsprechen den aufgezeichneten Kurven der vertikalen Bodenreaktionskraft der ersten 24/51 Läufe. Die schwarze Kurve entspricht dem Lauf, welcher für den Anstieg in der prozentualen Abweichung der Prädiktionsbänder verantwortlich ist. In beiden Grafiken ist zu erkennen, dass diese beiden Kurven das Ausmaß der Variabilität des Probanden erheblich vergrößern. Derartige Peaks treten für alle Probanden und alle verschiedenen Kurven auf, wobei die jeweils verantwortliche Kurve das Ausmaß der Variabilität erhöht. Es liegt daher nahe, die minimale Anzahl ab dem Lauf zu notieren, welcher eine derartige Veränderung hervorruft. Im Beispiel aus den Abbildungen 5.1.4, 5.1.5 und 5.1.6 müssten 52 Läufe aufgezeichnet werden, um das tatsächliche Ausmaß der Variabilität des betroffenen Probanden vollständig zu charakterisieren.

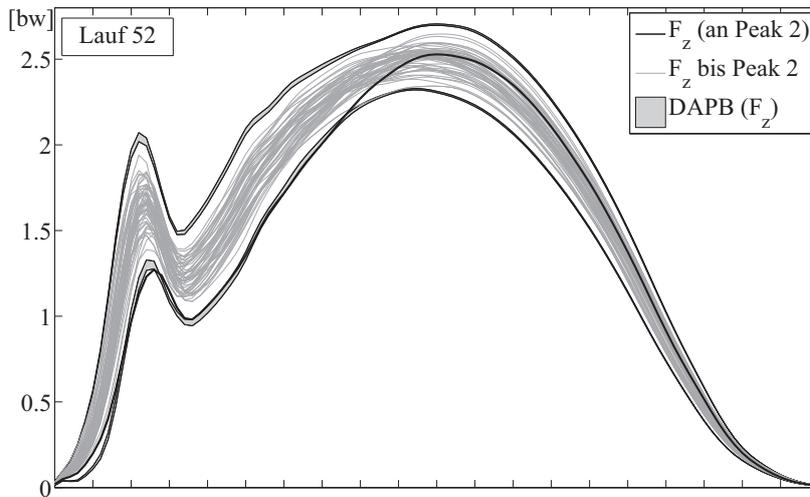


Abbildung 5.1.6:
Auswirkung des Laufs 52
(Peak 2) auf das Prädik-
tionsband entsprechend der
Abbildung 5.1.4

Tabelle 5.1.3 zeigt die mit verschiedenen Schwellenwerten der prozentualen Differenz der Flächen der Prädiktionsbänder bestimmte mittlere minimale Anzahl. Die entsprechenden Schwellen wurden auf 2 %, 3 %, 4 % und 5 % Abweichung festgesetzt. Zusätzlich erfolgte eine visuelle Überprüfung und das Notieren der entsprechenden minimalen Anzahl.

Es ist deutlich zu erkennen, dass sich die Mindestanzahl, ermittelt über den Krümmungspunkt der angepassten Funktion, wesentlich von den anhand der Differenzen von APB_N^{kr} bestimmten minimalen Anzahl an Versuchen unterscheidet. Die hohen Standardabweichungen in Tabelle 5.1.3 spiegeln die Individualität jedes Probanden wider, wobei einige Probanden schon mit wenigen Versuchen ihre tatsächliche Variabilität erreichen und andere Testpersonen den gesamten Pool an aufgezeichneten Läufen benötigen, um ihre Variabilität zu charakterisieren. Im letzten Fall muss die Variabilität dennoch nicht vollständig charakterisiert sein. Zur Überprüfung bedarf es mehr als 100 Versuche. Zudem ist die Abhängigkeit von den entsprechenden vorgegebenen Schwellen zu erkennen. Legt man die visuelle Überprüfung zu Grunde, entspricht die ermittelte minimale Anzahl der für mindestens drei Prozent Abweichung. Die mittlere Anzahl von ungefähr 60 Versuchen gleicht in etwa der minimalen Anzahl, ermittelt anhand der visuellen Überprüfung der Standardabweichungskurven. Die Betrachtung der Differenzen der minimalen Anzahl zwischen mehreren Messtagen zeigt zudem, dass im Mittel stark unterschiedliche Werte, ermittelt mit diesem Ansatz, bestimmt werden konnten. Die hohen Standardabweichungen unterstreichen, dass eine gleichbleibende Anzahl für einen Proband über mehrere Messtage nicht zu erwarten ist (↗ Tab. D.1, S. 125).

5.2 Minimale Anzahl unter Verwendung von Standardabweichungskurven

Der Ansatz zur Bestimmung einer minimal zu messenden Anzahl an Versuchen anhand von Differenzen der Standardabweichungskurven beruht auf der Idee der Existenz einer Populationsdifferenzenkurve (PDSD-Kurve) für die zu untersuchende Grundgesamtheit. Anhand der hier vorliegenden

Studie I						
k		$s = 2$	$s = 3$	$s = 4$	$s = 5$	visuell
F_z	M 1	84 ± 12	60 ± 22	53 ± 25	46 ± 25	63 ± 26
	M 2	78 ± 15	60 ± 23	47 ± 28	33 ± 24	61 ± 22
F_y	M 1	78 ± 13	70 ± 19	63 ± 20	55 ± 26	68 ± 21
	M 2	78 ± 14	58 ± 18	52 ± 18	30 ± 17	60 ± 18
F_x	M 1	84 ± 10	70 ± 16	52 ± 19	37 ± 16	70 ± 16
	M 2	78 ± 17	60 ± 20	50 ± 20	42 ± 21	56 ± 19
Studie II						
k		$s = 2$	$s = 3$	$s = 4$	$s = 5$	visuell
F_z	M 1	70 ± 17	58 ± 23	48 ± 22	42 ± 24	56 ± 24
	M 2	75 ± 14	56 ± 18	48 ± 20	37 ± 20	60 ± 21
F_y	M 1	79 ± 17	63 ± 26	54 ± 26	39 ± 22	65 ± 27
	M 2	73 ± 18	55 ± 22	39 ± 18	26 ± 11	53 ± 24
F_x	M 1	78 ± 14	60 ± 18	52 ± 22	42 ± 21	58 ± 21
	M 2	83 ± 13	69 ± 25	55 ± 23	45 ± 26	65 ± 27
ω	M 1	79 ± 15	63 ± 23	45 ± 20	35 ± 20	64 ± 21
	M 2	74 ± 18	58 ± 23	41 ± 17	30 ± 9	54 ± 22

Tabelle 5.1.3:

Mittlere minimale Anzahl anhand der Differenz der Flächen der Prädiktionsbänder $DAPB_N^{kr}$ und verschiedener vorgegebener Schwellen von 2%, 3%, 4% 5% und visuell

Daten soll diese PDS-D-Kurve zunächst geschätzt und danach kreuzvalidiert werden. Für die Determination dieser Populationsdifferenzenkurve wird das maximale Ausmaß aller Differenzen der Standardabweichungskurven (Abb. 5.2.1) für alle Probanden verwendet (Abb. 5.2.2). Die Verwendung einer mittleren Differenzenkurve würde den Probanden und das tolerierbare Ausmaß der Differenz der Standardabweichungskurve sehr streng und daher nicht repräsentativ charakterisieren. Dies wird häufig durch die letzten kumulierten Standardabweichungskurven hervorgerufen, welche eine Differenzenkurve nahe Null aufweisen (Abb. 5.2.1).

Gemäß der in Kapitel 2.5.2 beschriebenen Prozedur werden alle möglichen Differenzen für die gewählten maximalen Differenzen gebildet und gemittelt (Abb. 5.2.3). Die vorherige Randomisierung der zwei Messungen für jeden Proband in jeweils eine andere Stichprobe und die 100-malige Wiederholung dieser Prozedur ermöglicht die Bestimmung einer repräsentativen Stichprobe von mittleren Differenzenkurven für das Probandenkollektiv. Es entsteht somit eine Kurvenschar, für die mittels Bootstrap-Verfahren ein Prädiktionsband bestimmt wird (siehe Abbildung 5.2.4). Die äußeren Grenzen und der Mittelwert des Prädiktionsbandes kommen als Kandidaten für eine Populationsdifferenzenkurve in Frage. Die untere Grenze des Prädiktionsbandes hat aufgrund zum Teil starker Streuung der 100 mittleren maximalen Differenzenkurven (MDS-D-Kurven) für die Kurven der Bodenreaktionskräfte aus Studie I und der Winkelgeschwindigkeit negative Tendenzen, welche für absolute Differenzen nicht auftreten können. Von einer Verwendung der unteren Grenze des Prädiktionsbandes als PDS-D-Kurve wird daher abgesehen.

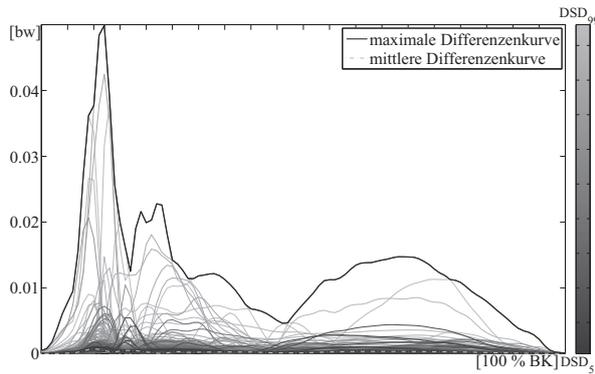


Abbildung 5.2.1: Differenz der Standardabweichungskurven der vertikalen Bodenreaktionskraft für einen Probanden mit maximalem und mittlerem Ausmaß

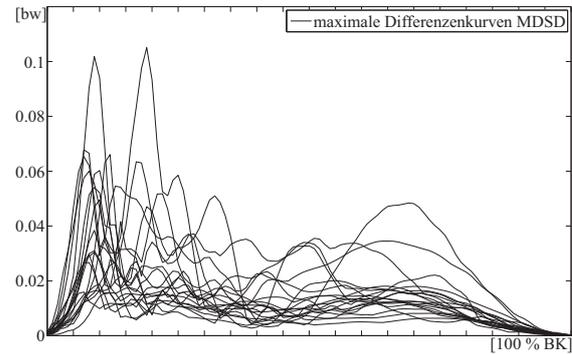


Abbildung 5.2.2: Maximale Differenzkurven aller Probanden nach Randomisierung (F_z)

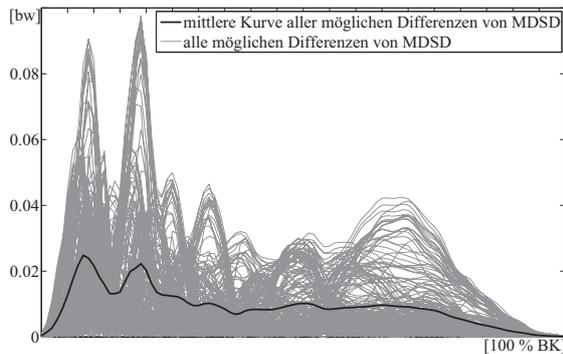


Abbildung 5.2.3: Alle möglichen absoluten Differenzkurven sowie Mittelwert dieser Kurven (F_z)

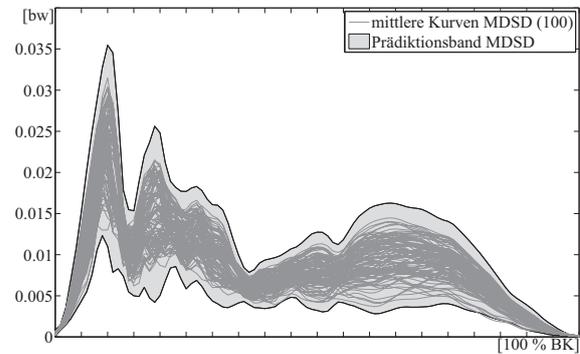


Abbildung 5.2.4: Prädiktionsband dieser entstehenden 100 mittleren maximalen Differenzkurven nach 100 Randomisierungen (F_z)

Abbildung 5.2.5 zeigt beispielhaft die Differenzen der Standardabweichungskurven eines Probanden mit der oberen Grenze und des Mittelwertes des Prädiktionsbandes der mittleren MDS-Kurven. Als minimal zu messende Anzahl für einen Probanden wird die Anzahl an Kurven notiert, ab der erstmals alle Differenzkurven kleiner als die vorgegebene PSD-Kurve sind. Für das Beispiel in Abbildung 5.2.5 ist dies für 20 Läufe der Fall. In Abbildung 5.2.6 sind entsprechend die in den Auswertungen zu Hypothese II bestimmten punktwise kumulierten und transformierten Standardabweichungskurven mit eingezeichneter minimaler Anzahl (=20) für dieses Beispiel dargestellt (↗ auch Abb. 4.2.3). Die Abweichung der Standardabweichungskurven in den letzten zehn Werten würde mit dieser Berechnung der minimalen Anzahl nicht berücksichtigt werden. Die teilweise Konvergenz davor wird durch die minimale Anzahl relativ stabil wiedergegeben.

In Tabelle 5.2.1 sind die mittleren minimal zu messenden Versuche pro Proband angegeben, wobei jeweils die mit derselben Stichprobe ermittelte entsprechende PSD-Kurve verwendet wurde. Auf

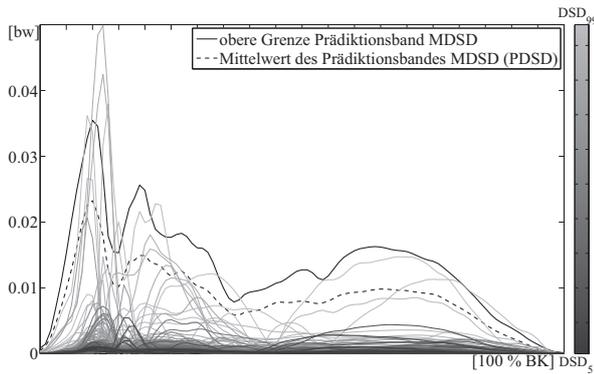


Abbildung 5.2.5:

Alle Differenzenkurven der Standardabweichung DSD für einen Probanden mit der oberen Grenze des Prädiktionsbandes und des Mittelwertes des Prädiktionsbandes (ab 20 Läufen liegen alle Kurven innerhalb der maximalen und der mittleren Grenze) (F_z)

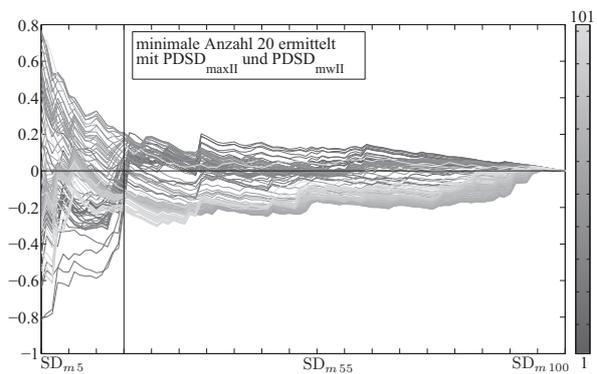


Abbildung 5.2.6:

Punktwise kumulierte und transformierte Standardabweichungskurven für den Proband aus Abbildung 5.2.5 mit anhand PDSD-Kurve ermittelter minimaler Anzahl

eine Unterteilung bezüglich der Randomisierung wird im Folgenden verzichtet. Die Differenzen der minimalen Anzahl bezüglich der beiden entstehenden Stichproben sind nahezu Null⁹. Daher sind in Tabelle 5.2.1 lediglich die Ergebnisse für die erste randomisierte Stichprobe angegeben. Erwartungsgemäß fällt die mittlere Anzahl mit der oberen Grenze des Prädiktionsbandes ($PDSD_{max}^k$) kleiner aus als mit dem Mittelwert des Prädiktionsbandes ($PDSD_{mw}^k$). Die zum Teil hohen Standardabweichungen zeugen von einer individuellen minimal zu messenden Anzahl für jeden Proband. Die minimal auftretende Anzahl schwankt zwischen 0 und 10 Versuchen für $PDSD_{max}^k$ und 6 bis 16 Versuchen für $PDSD_{mw}^k$. Die maximal bestimmte Anzahl reicht von 53 bis 92 für $PDSD_{max}^k$ und entsprechend 68 bis 92 für $PDSD_{mw}^k$. Die hohen mittleren Differenzen und Standardabweichungen der minimalen Werte zwischen den Messtagen hebt zudem das individuelle Verhalten für jeden Probanden hervor (↗ Tab. D.2, S. 126).

Der Verlauf der Populationsdifferenzenkurven hängt dabei enorm vom involvierten Probandenkollektiv ab. Abbildung 5.2.7 zeigt die drei berechneten $PDSD_{mw}^k$ -Kurven für Studie I, Studie II und alle verschiedenen Probanden aus beiden Studien (III). Es ist deutlich zu erkennen, dass das Probandenkollektiv aus Studie I eine höhere PDSD-Kurve aufweist als das Probandenkollektiv aus Studie II. Die Populationsdifferenzenkurve, zu deren Bildung alle unterschiedlichen Probanden herangezogen wurden, reiht sich genau in der Mitte ein. Dies ist aufgrund der Mittelwertbildung der maximalen Differenzenkurven zu erwarten. Die Einflüsse der geringeren Differenzen der Probanden aus Studie II und der hohen Unterschiede aus Studie I sollten sich aufgrund des Wiederholens der Berechnung der mittleren maximalen Differenzenkurven ausgleichen. Die Anwendung der Populationsdifferenzenkurve $PDSD_{II}^k$ auf Studie I und umgekehrt sollte für die minimale Anzahl zu messender Versuche für die Probanden aus Studie I eine Erhöhung und für die Probanden aus

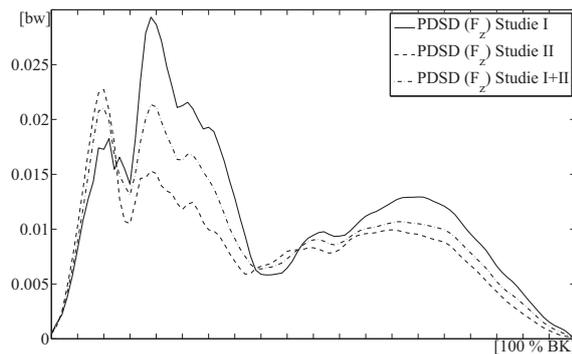
⁹ Ein Unterschied ist aufgrund der 100-maligen Wiederholung der beschriebenen Prozedur nicht zu erwarten, da die Randomisierung jeweils vor der Bildung der mittleren MDS D-Kurve vorgenommen wird. Die Information für jeden Proband bezüglich Messung eins und zwei geht somit mehrmals in die Berechnung ein. Auf lange Sicht werden sich die mittleren MDS D-Kurven und die entsprechenden Prädiktionsbänder angleichen.

		Studie I		Studie II	
k		$PDS D_{max I}^k$	$PDS D_{mw I}^k$	$PDS D_{max II}^k$	$PDS D_{mw II}^k$
F_z	M 1	26 ± 23	35 ± 26	30 ± 21	42 ± 23
	M 2	15 ± 7	24 ± 11	19 ± 9	29 ± 16
F_y	M 1	28 ± 15	48 ± 14	26 ± 15	39 ± 19
	M 2	17 ± 7	34 ± 17	19 ± 6	26 ± 10
F_x	M 1	19 ± 12	31 ± 16	21 ± 17	40 ± 21
	M 2	18 ± 13	31 ± 15	22 ± 19	34 ± 19
ω	M 1			29 ± 18	49 ± 20
	M 2			28 ± 18	39 ± 18

Tabelle 5.2.1:

Mittlere minimale Anzahl ermittelt anhand oberer Grenze und Mittelwert des Prädiktionsbandes der mittleren MDS-D-Kurve als Populationsdifferenzkurve (Mittelwert \pm Standardabweichung)

Studie II eine Verringerung der entsprechenden Anzahl bedeuten.


Abbildung 5.2.7:

Populationsdifferenzkurven: $PDS D_{mw I}^k$, $PDS D_{mw II}^k$, $PDS D_{mw III}^k$

Tabelle 5.2.2 zeigt die mittleren Differenzen der ermittelten minimalen Anzahl pro Proband, wobei jeweils die Differenz der originär ermittelten Anzahl zur kreuzvalidierten Anzahl berechnet wurde. Für Studie I ergab sich somit eine leicht höhere Anzahl (negative Differenz) und für Studie II deuten die leicht positiven Differenzen auf eine Verringerung der minimal zu messenden Anzahl.

Neben dem Vergleich der minimalen Anzahl, ermittelt anhand diverser Populationsdifferenzkurven, ist die Gegenüberstellung zur minimalen Anzahl, berechnet mittels der kumulierten Differenz der Flächen der Prädiktionsbänder, eine zusätzliche Information. Tabelle 5.2.3 beinhaltet die Differenzen bezüglich der entsprechenden Vergleiche, wobei die Differenzenbildung gemäß $\#PDS D - \#DAPB$ berechnet wurde.

Es wird deutlich, dass die minimale Anzahl, welche durch die Populationsdifferenzkurve bestimmt wurde, zum Teil wesentlich kleiner ist als für die Differenzen der Flächen der Prädiktionsbänder (negative Differenz). Die Betrachtung der Differenz für den Schwellenwert $s = 5$ zeigt teilweise geringe Unterschiede bei relativ hohen Standardabweichungen. Für andere Populationsdifferenzkurven ändert sich diese Abweichung. Für die PDS-D-Kurve aus beiden Studien (III) fallen die

		Studie I		Studie II	
k		$PDS D_{max II}^k$	$PDS D_{mw II}^k$	$PDS D_{max I}^k$	$PDS D_{mw I}^k$
F_z	M 1	$- 7 \pm 19$	$- 5 \pm 8$	2 ± 8	13 ± 23
	M 2	$- 7 \pm 9$	$- 9 \pm 13$	4 ± 9	$- 1 \pm 16$
F_y	M 1	$- 4 \pm 15$	-12 ± 18	2 ± 22	0 ± 15
	M 2	-11 ± 13	-10 ± 15	$- 2 \pm 14$	$- 1 \pm 17$
F_x	M 1	$- 8 \pm 14$	$- 8 \pm 13$	7 ± 7	9 ± 9
	M 2	$- 1 \pm 17$	0 ± 12	16 ± 22	13 ± 18

Tabelle 5.2.2:

Kreuzvalidierung der Ermittlung der minimal zu messenden Anzahl anhand der entgegengesetzten Populationsdifferenzenkurve

Unterschiede für diesen Schwellenwert kleiner aus, werden somit unterschätzt. Die Anzahl der Probanden, deren Differenz der minimalen Anzahl für beide Berechnungsmethoden um maximal neun schwankt, bleibt ungefähr gleich. Weisen bei Studie II 15 von 20 Probanden (Messtag 1) für die vertikale Bodenreaktionskraft ungefähr die gleiche Anzahl auf, ändert sich dies bei Anwendung von $PDS D_{mw III}^k$ auf 13 von 18 Probanden¹⁰ (↗ Tab. D.3, S. 126). Die hohen Standardabweichungen für den Schwellenwert $s = 5$ deuten darauf hin, dass die relevanten Peaks der Differenzen der Flächen der Prädiktionsbänder nicht für jeden Probanden gefunden werden. Bei einer erneuten Berechnung der Populationsdifferenzenkurve unter Verwendung weiterer Probanden liegt die Vermutung nahe, dass sich die Ergebnisse wiederum ändern werden.

In einem letzten Schritt werden die sich nicht überlappenden Flächen der Prädiktionsbänder für alle 100 Kurven mit der für die mittels Populationsdifferenzenkurve bestimmten minimalen Anzahl verglichen. Tabelle 5.2.4 gibt einen Überblick über die mittlere prozentuale Abweichung aus diesen beiden Prädiktionsbändern. Dabei sei angemerkt, dass der Vergleich ab einer minimalen Anzahl größer 20 vorgenommen wurde, da die Berechnung der Prädiktionsbänder für kleinere Stichprobengrößen anfälliger ist (Albers et al., 2007, S. 398). Probanden, welche eine kleinere minimale Anzahl aufwiesen, wurden aus diesem Vergleich daher ausgeschlossen. Verwendet wurde die mittels $PDS D_{mw I}^k$ oder $PDS D_{mw II}^k$ bestimmte minimale Anzahl jeweils für die erste randomisierte Stichprobe. Es ist deutlich zu erkennen, dass die prozentualen Abweichungen wesentlich geringer sind als die Unterschiede, welche innerhalb eines Tages festgestellt werden konnten (↗ Tab. 4.2.2). Des Weiteren fallen diese zudem geringer aus als die Unterschiede zwischen den Tagen. Ein Vergleich dieser prozentualen Differenzen zwischen APB_{25}^{kr} und APB_{100}^{kr} (↗ Tab. C.6, S. 124) zeigt zudem eine Abnahme der prozentualen Abweichungen für die mittels PDS D-Kurve bestimmten minimal zu messenden Anzahl an Versuchen gegenüber dieser Differenz. Dies deutet auf eine Verbesserung der Übereinstimmung der Prädiktionsbänder und damit verbunden auf eine im Mittel verbesserte Charakterisierung der Variabilität eines Probanden.

Die ermittelte minimale Anzahl für die drei Bodenreaktionskräfte pro Proband ist relativ stabil

¹⁰ Der Vollständigkeit halber muss erwähnt werden, dass bei Ausschluss der Probanden aus Studie II, welche an beiden Studien teilgenommen haben, sich eine Anzahl von 14 von 18 mit $PDS D_{mw II}^k$ und 13 von 18 mit $PDS D_{mw III}^k$ ergibt.

Studie I						
k		$s = 2$	$s = 3$	$s = 4$	$s = 5$	visuell
F_z	M 1	-49 ± 29	-25 ± 27	-18 ± 28	-10 ± 21	-27 ± 28
	M 2	-54 ± 18	-36 ± 26	-22 ± 30	-8 ± 24	-36 ± 26
F_y	M 1	-30 ± 22	-22 ± 30	-15 ± 27	-7 ± 32	-20 ± 31
	M 2	-45 ± 24	-24 ± 19	-18 ± 16	4 ± 20	-26 ± 21
F_x	M 1	-52 ± 20	-39 ± 25	-20 ± 20	-5 ± 15	-38 ± 25
	M 2	-47 ± 23	-29 ± 28	-19 ± 28	-11 ± 28	-25 ± 27

Studie II						
k		$s = 2$	$s = 3$	$s = 4$	$s = 5$	visuell
F_z	M 1	-29 ± 25	-17 ± 28	-7 ± 20	-1 ± 23	-15 ± 29
	M 2	-46 ± 19	-26 ± 19	-19 ± 19	-8 ± 21	-31 ± 24
F_y	M 1	-39 ± 26	-24 ± 31	-15 ± 28	0 ± 24	-25 ± 31
	M 2	-47 ± 18	-29 ± 23	-13 ± 23	0 ± 15	-27 ± 25
F_x	M 1	-38 ± 21	-19 ± 21	-12 ± 24	-1 ± 26	-17 ± 23
	M 2	-49 ± 23	-35 ± 33	-21 ± 28	-12 ± 29	-32 ± 31
ω	M 1	-30 ± 21	-14 ± 20	4 ± 21	14 ± 25	-15 ± 21
	M 2	-36 ± 23	-19 ± 22	-3 ± 9	9 ± 18	-16 ± 22

Tabelle 5.2.3:

Differenz $\#_{PDSD} - \#_{DAPB}$ mittels $PDSD_{mw}^k$ und der entsprechenden eingeführten Schwellen (Mittelwert \pm Standardabweichung)

	Studie I		Studie II	
p	Messtag 1 (%)	Messtag 2 (%)	Messtag 1 (%)	Messtag 2 (%)
F_z	18 ± 10	21 ± 12	13 ± 6	16 ± 8
F_y	11 ± 4	17 ± 10	14 ± 7	18 ± 9
F_x	16 ± 15	15 ± 11	12 ± 6	15 ± 14
ω			10 ± 5	14 ± 6

Tabelle 5.2.4:

Differenz der Flächen der Prädiktionsbänder $APB_{\#_{PDSD}}^{kr}$ für die mittels Populationsdifferenzenkurve ermittelte minimale Anzahl $\#_{PDSD}$ und APB_{100}^{kr} für alle 100 involvierten Kurven

und weicht um neun bis 18 Versuche voneinander ab (\nearrow Tab. D.4, S. 127). Die Schwankung dieser Anzahl innerhalb eines Probanden sinkt anhand der Verwendung der Populationsdifferenzenkurve aus beiden Studien.

5.3 Diskussion der Ergebnisse

Der Kern der Auswertungen zu Hypothese III setzt sich mit der Entwicklung eines Algorithmus zur Bestimmung der minimal notwendigen zu messenden Anzahl an Versuchen im Labor für jeden Proband auseinander. Der erste Algorithmus bediente sich des Flächeninhaltes der Prädiktionsbänder der kumulierten Kurven der Bodenreaktionskraft und der Winkelgeschwindigkeit.

Die nichtlineare Regression der kumulierten Folge der Flächeninhalte durch eine exponentiell abfallende Funktion und die Bestimmung des Krümmungspunktes des entsprechend angepassten Modells zeigen auf den ersten Blick relativ stabile Messungen. Die Abhängigkeit vom Ausmaß, dem maximalen Wert der jeweiligen eingehenden Ausgangsfolge, deutet auf die Möglichkeit der Manipulierbarkeit des Algorithmus. Die Veränderung des maximalen ersten Wertes der Folge hat Auswirkungen auf die letztlich zu bestimmende minimale Anzahl. Verwendet man die normierte Kurve der Fläche des Prädiktionsbandes APB_N^{kr} müssen lediglich acht Versuche aufgezeichnet werden. Mit einer weiteren kleinen Anpassung des Ausmaßes könnte man die weit verbreiteten fünf Versuche bestätigen. Das Potential der Steuerung auf ein bestimmtes Ergebnis ist eine Einschränkung der Verwendbarkeit dieser Herangehensweise.

Bei der genauen Betrachtung der Qualität der Anpassung lenkt das große Ausmaß der kumulierten Folge der Prädiktionsbänder von der tatsächlichen Güte der Anpassung ab. Die Anwendung der Fehlervarianz als Maß für die Güte der Anpassung beinhaltet keine Wertung derselben. Im Vergleich zweier Modelle ist es lediglich möglich eine Aussage zu treffen, welches Modell weniger Fehler produziert, so dass eine visuelle Überprüfung Vorteile birgt. Dabei wird für einige Probanden der tatsächliche Krümmungspunkt über- oder unterschätzt. Des Weiteren muss für andere Probanden die Annahme des exponentiellen Abnehmens der Folge abgelehnt werden. Diese Folgen zeigen keine tatsächliche Konvergenz (↗ Abb. 5.1.3). Die verwendeten Modelle 1 und 2 weisen jedoch ein derartiges exponentiell stetig abnehmendes Verhalten auf. In einigen Parameterkombinationen existiert darüber hinaus ein Grenzwert für $x \rightarrow \infty$ oder zumindest eine sehr hohe Wiederholungszahl:

$$\lim_{x \rightarrow \infty} \underbrace{e^{(-\gamma x + \delta)}}_{\rightarrow 0} + \eta = \eta \quad \text{für } \gamma < 0 \quad (\text{Modell 1})$$

und

$$\lim_{x \rightarrow \infty} \alpha + \underbrace{\beta x^\gamma}_{\rightarrow 0} = \alpha \quad \text{für } \beta > 0, \gamma < 0 \quad (\text{Modell 2}).$$

Der Umstand, dass die Folgen APB_N^{kr} zum Teil keine abnehmenden Werte aufweisen, stellt die Anwendung der nichtlinearen Regression anhand der Modelle 1 und 2 in Frage. Es ist somit nicht möglich, ein einheitliches Modell für alle Probanden zu finden, und schließt die Ermittlung einer minimalen Anzahl anhand APB_N^{kr} -Folgen aus.

Die Anfälligkeit des Bootstrap-Verfahrens für kleine Stichproben (Albers et al., 2007, S. 398) unterstreicht zudem, dass die Folgen der APB_N^{kr} erst ab mindestens $N = 10$ betrachtet werden sollten. Bei Wegfallen der extremen Maxima dieser Folgen wird deutlich, dass die Annahme des Vorliegens eines einheitlichen Modells nicht gegeben ist (↗ Abb. 5.1.3).

Das Ausweichen auf die Differenzen der sich nicht überlagernden Flächen des Prädiktionsbandes $DAPB_N^{kr}$ und die Bestimmung der minimalen Anzahl anhand entscheidender Peaks erscheint daher

sinnvoll. Diese Herangehensweise funktioniert lediglich, wenn die entsprechenden Kurven, für welche das Prädiktionsband bestimmt wurde, nicht gleichmäßig stets am Rand der jeweiligen Spannweite der vorangegangenen Streuung der Kurven liegen. Aufgrund der Annahme der zufälligen Streuung ist dieser Ausnahmefall jedoch nicht zu erwarten.

Die Betrachtung der Peaks innerhalb der gebildeten Differenzenfolge der APB_N^{kr} offenbart, dass Anzahl, Höhe und der Zeitpunkt des Auftretens sehr individuell für jeden Probanden in Erscheinung treten.

Die Überlegung der Bildung eines Algorithmus, welcher während einer Labormessung einsetzbar ist, kann auf Basis dieser Daten nicht entwickelt werden.

Die Vorhersage der entsprechenden Peaks durch die während der Messung gewonnenen Daten ist bisher nicht möglich, da die vermutlich stochastischen Einflussgrößen zunächst nicht bekannt sind und das stochastische Modell oder der stochastische Prozess ohne Kenntnis dieser Größen nicht aufgestellt werden kann. Eine vollständige Charakterisierung der Variabilität eines Probanden ist daher ausschließlich post-hoc möglich. Die technischen Einschränkungen bezüglich der zeitlichen Intensität des Verfahrens begrenzt die Anwendung während der Datenaufnahme im Labor (↗ S. 69 f.).

Die Möglichkeit der Bestimmung einer minimalen Anzahl pro Proband im Labor vor der entscheidenden Studie muss aufgrund der hohen Differenzen der ermittelten Mindestanzahl innerhalb der Messtage ausgeschlossen werden. Dieser Umstand unterstreicht die starke Individualität bezüglich der Streuung der involvierten Kurven für jede Testperson. Eine minimale Anzahl von im Mittel 60 Versuchen ist im ökonomischen Sinn nicht mehr tragbar. Für eine Labormessung mit 60 Läufen wird pro Bedingung mindestens eine halbe Stunde Messzeit benötigt, wobei mehrere Untersuchungsbedingungen diese zudem erhöhen.

Der zweite Ansatz der Ermittlung einer minimal zu messenden Anzahl an Versuchen beruht auf der Betrachtung der kumulierten Standardabweichungskurven und deren Differenzen zwischen zwei aufeinander folgenden Messungen. Die Annahme der Existenz einer Populationsdifferenzenkurve wird durch die Vermutung forciert, dass alle Probanden ähnliche Standardabweichungskurven und ähnliche Differenzen dieser Kurven aufweisen. Für die vertikale Bodenreaktionskraft zeigen alle Probanden im Bereich der ersten Kraftspitze eine höhere Standardabweichung als zu Beginn oder am Ende der Messung. Dies deutet darauf hin, dass eine derartige gemeinsame Kurve existieren könnte.

Die zweite Annahme, dass die Stichprobe eine Grundgesamtheit repräsentiert, was im Sinne der Berechnung einer Populationsdifferenzenkurve gegeben sein muss, ist schwieriger zu überprüfen. In dieser Untersuchung wurden originär Läufer gemessen, welche über eine gewisse Distanz pro Woche laufen sollten. Dies konnte in Studie II nicht ausschließlich übernommen werden. Sechs der 20 Probanden waren zum Zeitpunkt der Messung sportlich inaktiv. Das Ausmaß der maximalen Differenzenkurve für diese sechs Probanden unterschied sich jedoch nicht wesentlich von den weiteren Testpersonen, so dass Läufer als Grundgesamtheit angesehen werden können. Die Unterscheidung der Populationsdifferenzenkurven beider Studien lässt vermuten, dass wesentlich mehr Probanden für die Bildung einer PDS-D-Kurve benötigt werden, als bisher zur Verfügung standen. Die Tatsache, dass die PDS-D-Kurve aus beiden Studien eine Art Mittelwert aus beiden Kurven bildet,

unterstreicht diese These. Fraglich ist zudem, inwieweit sich die Populationsdifferenzenkurven ändern, wenn kein einziger Läufer dem Probandenkollektiv angehören würde, und ob entsprechend die PDS-D-Kurven angepasst oder eigens erstellt werden müssten.

Die Verwendung der maximalen Differenzenkurve als Ausgangspunkt für die Entwicklung einer Populationsdifferenzenkurve ist zunächst eine eher konservative Herangehensweise, wobei der Schritt der Differenzenbildung dem Vorgehen für jeden einzelnen Probanden entspricht. An dieser Stelle entstehen zwei Möglichkeiten, welche das hier vorgestellte Verfahren optimieren könnten. Zum einen können andere Ausgangswerte die Ermittlung der minimalen Anzahl beeinflussen. Denkbar wären der Median oder andere Quantile der entsprechenden Differenzenkurven des Probanden. Zum anderen ist die Wiederholungszahl, wie oft die Prozedur der Randomisierung der Auswahl der entsprechenden maximalen Differenzenkurve erfolgt, optional. Die hier gewählten 100 Wiederholungen können durchaus noch optimiert werden, wobei ein $n > 100$ eine genauere mittlere Differenzenkurve bedingt und damit das Prädiktionsband dieser und die entsprechende PDS-D-Kurve genauer geschätzt werden können. Jedoch müssen an dieser Stelle die Grenzen der technischen Umsetzbarkeit des Bootstrap-Verfahrens hinsichtlich der Anzahl der Wiederholungen und der zeitintensiven Bestimmung des interessierenden Quantils für die Breite des Prädiktionsbandes berücksichtigt werden (↗ S. 69 f.). Zudem kann die Bildung aller möglichen Differenzenkurven pro Proband durch ein anderes Verfahren ersetzt werden. Die Berechnung der mittleren maximalen Differenzenkurve ist ferner möglich, führt jedoch zu einer Kurve mit größerem Ausmaß, welche höhere Differenzen toleriert und in einer kleineren Anzahl zu messender Versuche mündet. Die optimale Balance zwischen allen möglichen Einstellungen und der tatsächlichen Charakterisierung der PDS-D-Kurve muss ferner gefunden werden.

Die letztlich bestimmte minimale Anzahl an zu messenden Versuchen mit $PDS-D_{mw}^k$ oder $PDS-D_{max}^k$ ergibt nach visueller Überprüfung gemäß Abbildung 5.2.6 und der Einschätzung des Verlaufes der punktwise kumulierten Standardabweichungen durchaus akzeptable Werte. Diese Auffassung ist subjektiv behaftet. Ohne Angabe einer derartigen Grenze ist die Einstellung viel konservativer. Bei Begutachtung der punktwise kumulierten Standardabweichungen ohne eingetragene minimale Anzahl ergaben sich höhere minimal zu messende Anzahlen (↗ Tab. 4.2.1). Eine mögliche Begründung könnte in den Urteilsheuristiken aus der Psychologie liegen. Diese Strategien werden von Personen in dem Moment verwendet, wenn ihnen zur Urteilsfindung nicht genügend Informationen zur Verfügung stehen. Durch Angabe eines Ankers wird das entsprechende Urteil an diesem Anker ausgerichtet (Fischer & Wiswede, 2002). Dabei kommt es zum so genannten Ankereffekt, der „ungenügenden Anpassung von Einschätzungen, die von einem (irrelevanten) Startwert ausgehen“ (Frey, 2007, S. 330). Diese „mentale Abkürzung“ führt dazu, dass sich die getroffene Entscheidung an diesem Anker ausrichtet (Aronson, Wilson & Akert, 2004, S. 80). Im ersten Fall, der Betrachtung der punktwise kumulierten Standardabweichungen ohne eingezeichnete minimale Anzahl, wird als Anker der letzte Wert (=Null) der Grafik verwendet. Im zweiten Fall entspricht der Anker gerade der minimalen Anzahl. In beiden Fällen richtet sich die Entscheidung an diesen Ankern aus. Diese Beeinflussung ist nach Frey (2007, S. 214) eine „unbewusste Denkheuristik, die nur schwer rational kontrollierbar ist“, da selbst das Wissen, dass es diesen Effekt gibt, und die Mahnung vor der Verwendung dieser Ankerwerte keinen Einfluss auf das jeweilige Urteil ausüben (Aronson et al., 2004, S. 82).

Die hohen Schwankungen sowohl zwischen den Probanden als auch zwischen den Messtagen zeu-

gen wiederum von der Individualität jedes einzelnen Probanden. Dies unterstreicht die Wichtigkeit der Implementierung eines Entscheidungssystems, welches während der Messung die Variabilität des Probanden bewertet. Die relativ geringen Schwankungen der minimalen Anzahl innerhalb eines Probanden für alle drei Bodenreaktionskraftkurven deutet zudem auf eine generelle minimale Anzahl unabhängig von der betrachteten Kurve.

Der Algorithmus anhand der Verwendung der Differenzen der Standardabweichungskurven scheint ein möglicher Lösungsansatz zu sein.

Die Abhängigkeit der PDSK-Kurve vom jeweiligen Probandenkollektiv macht eine durchgängige Erneuerung dieser Kurve unumgänglich, wobei die Frage offen bleibt, wie sich diese Kurve mit zunehmendem Probandenkollektiv ändert und wie repräsentativ die fiktive Populationsdifferenzkurve aus z. B. 100 verschiedenen Teilnehmern sein wird und wie diese durch unterschiedliche Untersuchungsbedingungen beeinflusst wird.

Die Auswertungen bezüglich der Differenz der minimalen Anzahl, ermittelt anhand der PDSK-Kurve, und den Peaks der Differenzen der prozentualen Abweichung der Prädiktionsbänder zeigt zudem, dass die Variabilität für einige Probanden perfekt und für andere Testpersonen übermäßig oder zu wenig charakterisiert wird. Inwiefern eine globale PDSK-Kurve dieses Bild ändern würde, bleibt offen. Die Tatsache, dass eine prozentuale Veränderung um mindestens 5 % in etwa festgestellt werden kann, eröffnet die Diskussion, ob eine Abweichung von 5 % die Variabilität des Probanden dennoch ausreichend charakterisiert. Im Sinne der Statistik wird eine Irrtumswahrscheinlichkeit von 5 % durchaus als akzeptabel angesehen, so dass eine prozentuale Abweichung um 5 % eine Art Irrtumswahrscheinlichkeit dieses Algorithmus darstellen könnte. Dies impliziert, dass Interventionen, welche überprüft werden sollen, eine Veränderung der prozentualen Abweichung über 5 % pro Proband erreichen sollten.

Kapitel 6

Diskussion und Ausblick

„Nun sag, wie hast du's mit der Messung im Labor?“¹¹

So oder ähnlich lautet die sich anhand dieser Arbeit ergebende Gretchenfrage bezüglich biomechanischer Auswertungen und Studien des Laufsports, welche unter Laborbedingungen durchgeführt werden. Die Reflexion der Ergebnisse bezüglich der drei aufgestellten Hypothesen ist in Tabelle 6.1 dargestellt.

	Hypothese	Diskrete Parameter	Gesamtkurve
HI	Der Mittelwert kann als ausreichende und stabile Charakterisierung des Probanden angesehen werden.	✗	F_z, F_y ✓ F_x, ω ✗
II	Die Variabilität eines Probanden kann mit einer endlichen Anzahl an Wiederholungen vollständig und stabil charakterisiert werden.	✗	✓
III	Unter der Annahme der Bestätigung von Hypothese I und/oder Hypothese II kann die endliche Anzahl an Wiederholungen mit einem einheitlichen Algorithmus bestimmt werden.	–	APB_N^{kr} ✗ $PDSD^k$ ✓

Tabelle 6.1:

Darstellung der Ergebnisse zu den aufgestellten Hypothesen I, II und III (✓ - Nachweis erbracht; ✗ - Nachweis nicht erbracht)

Es wird deutlich, dass die postulierte ausreichende und stabile Charakterisierung eines Probanden anhand des Mittelwertes (HI) sowie der vollständigen und soliden Beschreibung der Variabilität (HII) für diskrete Parameter nicht nachgewiesen werden konnte. Die Entwicklung eines Algorithmus für diskrete Parameter erwies sich daher als hinfällig.

Für gesamte Kurven ergab sich ein anderes Bild. Die Probanden konnten anhand der mittleren vertikalen Bodenreaktionskräfte sowie der Bodenreaktionskräfte in anterior-posterior Richtung stabil und ausreichend charakterisiert werden. Für die Bodenreaktionskräfte in mediolateraler Richtung

¹¹ Originaltext aus Goethes Faust: „Nun sag, wie hast du's mit der Religion?“ (Goethe, 1990, Vers 3416)

und die Kurve der Winkelgeschwindigkeit wurde dies nicht bestätigt. Die Auswertung zu Hypothese II verifizierte die Möglichkeit der Charakterisierung der Variabilität eines Probanden.

Die Bestimmung der minimalen Anzahl und Wiederholungen durch einen einheitlichen Algorithmus (HIII) konnte damit anhand der Variabilität erfolgen, wobei ein Verfahren erzeugt werden musste, welches während der Messung des Probanden applizierbar ist. Als Methoden wurden die sich nicht überlagernden Flächen der benachbarten Prädiktionsbänder APB_N^{kr} herangezogen sowie die Differenzen der Standardabweichung DSD_{mn}^{kr} , welche mittels einer Populationsdifferenzenkurve $PDSD^k$ zu beurteilen waren. Prädiktionsbänder konnten für die Entwicklung eines Algorithmus, welcher die Variabilität des Probanden während der Aufzeichnung der Messversuche ausreichend charakterisieren kann, nicht genutzt werden. Der Algorithmus bezüglich der Verwendung einer Populationsdifferenzenkurve hat Potential für einen möglichen Einsatz während der Messung eines Probanden im Labor.

Die hier eingesetzten Methoden bezüglich der Entwicklung eines Algorithmus zur Bestimmung einer minimalen Anzahl an Messversuchen im Labor stellen lediglich einen Bruchteil der gesamten Vielfalt an möglichen zur Verfügung stehenden Modellen dar. Zudem ist das Verfahren anhand einer Populationsdifferenzenkurve momentan nicht ausreichend validiert. Für den tatsächlichen Einsatz im Labor müssen die entsprechenden Punkte aus Kapitel 5.3 hinsichtlich der Existenz und Wahl einer PDSD-Kurve, der Verwendung der maximalen Differenzenkurven der Standardabweichungskurven sowie des Einsatzes des Bootstrap-Verfahrens geklärt und optimiert werden (↗ S. 69 f., S. 83 f.). Letzteres wird bereits bei der Bestimmung der Prädiktionsbänder der gesamten Kurven eingesetzt. Bezüglich der Anzahl an Koeffizienten für die Fourierentwicklung und der Rechenkapazität, insbesondere bei der iterativen Bestimmung des entsprechenden Quantils, bedarf es einer Optimierung dieses Verfahrens. Die Einschränkung, dass mindestens zehn, besser 25 Versuche für die Bestimmung der Bootstrap-Stichprobe verwendet werden sollten, beeinflusst die Entscheidung, wie viele Versuche pro Proband aufzuzeichnen sind.

Die einfachen statistischen Hilfsmittel, wie kumulierte Mittelwerte und Standardabweichungen sowie deren Konvergenzbetrachtung, scheinen daher eher geeignet, um die methodischen Aspekte zu beurteilen. Dabei sei erwähnt, dass die Berechnung der Standardabweichungen in Abhängigkeit vom Mittelwert erfolgt. Dies eröffnet die Frage, ob Mittelwert und Standardabweichung als funktionale Einheit betrachtet werden müssen. Wenn der Mittelwert keine adäquate Charakterisierung des Probanden darstellt, inwiefern kann dann davon ausgegangen werden, dass der Standardabweichung diese Funktion zukommen könnte? Der Einfluss des nicht adäquat verwendeten Mittelwertes auf die Ergebnisse bezüglich Hypothese II kann nicht abgeschätzt werden. Wenn divergierende Mittelwerte einen direkten Einfluss auf die Konvergenz der Standardabweichungskurven ausüben, müssten die Probanden mit divergierenden Mittelwertkurven ausschließlich divergierende SD-Kurven zeigen. Dies ist in den vorliegenden Daten nicht der Fall. Trends in den Einzelwerten sowie abnehmende oder zunehmende Streuung beeinflussen die Konvergenz relativ unabhängig voneinander, wobei die Auswirkung divergenter Mittelwertkurven nicht negiert werden kann. Bei Betrachtung des einfachen Schwankungsintervalls ($MW \pm SD$) wird diese gegenseitige Beeinflussung deutlich. Die alleinige Betrachtung der Streuung ist ohne die Betrachtung des Mittelwertes nicht aussagekräftig, so dass Mittelwert und Standardabweichung als Einheit aufgefasst werden sollten. Selbiges gilt für die gesamten Kurven. Streng genommen ist lediglich für die vertikalen Bodenreaktionskräfte und die Kräfte in anterior-posterior Richtung die Entwicklung eines Algorithmus zur Bestimmung einer

minimalen Anzahl sinnvoll, da deren Mittelwertkurven zum Großteil konvergieren.

Die lediglich visuell bestimmte minimale Anzahl an Versuchen anhand der jeweiligen kumulierten Kurven diskreter Mittelwerte und gemittelter Gesamtkurven birgt zwei wesentliche Vorbehalte. Die Entscheidung des konvergenten Verhaltens sowie die letztlich festgelegte minimale Anzahl sind subjektiv behaftet und werden durch den letzten bestimmten Wert beeinflusst (↗ Ankereffekt S. 84 f.). Die dabei getroffene Annahme, dass dieser letzte Wert jeweils das Optimum darstellt, wird zwar durch die hohe Anzahl an Messversuchen diktiert, ist aber nicht ohne weiteres begründbar. Dieses Optimum kann wesentlich eher, sehr viel später als bei 100 Messversuchen oder gar nicht auftreten. Vor allem bei der Betrachtung der Konvergenz gesamter Kurven entsteht der Eindruck, dass der auf Null normalisierte letzte Wert der optimale Wert sei. Die Angabe der minimalen Anzahl an Versuchen orientiert sich somit an diesem letzten Wert, wobei die Überprüfung durch mehrere Untersucher dies nicht verändern würde und die subjektive Entscheidung erhalten bliebe. Dies verdeutlicht zudem die Notwendigkeit eines objektiven Algorithmus und hebt die Angabe der tatsächlichen prozentualen Abweichung zwischen dem Mittelwert von fünf und 100 Messversuchen hervor, welche sich für konvergente diskrete Parameter im Mittel auf rund fünf bis sechs Prozent und für konvergente gesamte Kurven auf 14 bis 19 Prozent beläuft. Die absoluten Differenzen eines Probanden innerhalb eines Messtages und die damit begangenen Fehler können unter Umständen genau so groß sein, wie die mittleren Differenzen für verschiedene Untersuchungsbedingungen (↗ S. 68 f.).

Vor allem die prozentualen Abweichungen für diskrete Parameter der Bodenreaktionskräfte üben einen Einfluss auf die Berechnung von weiteren biomechanischen Parametern aus, welche direkt oder indirekt anhand der drei verschiedenen Kräfte berechnet werden. Bodenreaktionskräfte werden für diverse kinematische Berechnungen u. a. von Drehmomenten, Körperschwerpunktsberechnungen und Kräften an verschiedenen Gelenken sowie zur Synchronisation mit kinematischen Aufnahmen verwendet. Wenn die im Labor bestimmten Bodenreaktionskräfte einen Probanden nicht adäquat repräsentieren, verursacht dies Konsequenzen innerhalb der kinematischen Betrachtung unter Laborbedingungen. Die Fehler, die während der Messung von Bodenreaktionskräften in Kauf genommen werden, beeinflussen somit weitere biomechanische Parameter, welche zur Determination z. B. des Abrollvorgangs beim Ausdauerlaufen herangezogen werden.

Die Diskrepanzen für die Winkelgeschwindigkeit und das Pronationsausmaß lassen die Vermutung offen, dass derartige Unterschiede für andere Gelenkwinkel zu erwarten sind. Dies könnte sich auf Ganganalysen und sonstige biomechanische oder medizinische Verfahren auswirken, welche kinematische Messmethoden verwenden. Die Untersuchung der tatsächlichen Unterschiede für kinematografisch aufgenommene Daten sollten in weiteren Studien überprüft werden.

In der in Kapitel 1.1 vorgestellten Literatur wurden fast ausschließlich Mittelwerte diskreter Parameter verwendet. Vor dem Hintergrund der Ergebnisse aus der vorliegenden Arbeit bezüglich Hypothese I und der fehlerhaften Verwendung des Mittelwertes als Charakterisierung des Probanden müssen streng genommen Ergebnisse und Interpretationen vergangener Studien in Frage gestellt werden. Die Auswirkungen dieser irrtümlichen Verwendung des Mittelwertes sind weitreichend und beeinflussen zunächst die zukünftigen Auswertungen und die Entwicklung von weiteren Studien vor allem in Bezug auf die Entwicklung neuer Methoden.

Die interne Haltung einer Arbeitsgruppe gegenüber diesen Ergebnissen für diskrete Parameter,

der Ablehnung bezüglich Hypothese I, kann in nur zwei Einstellungen resultieren - Ignorieren dieser Forschungsergebnisse oder aktives Gegenwirken. Ignoranz ist im Sinne von Wissenschaft und Forschung nicht zu empfehlen, sogar abzulehnen. Die Differenzen innerhalb und zwischen den Messungen deuten auf erhebliche Messfehler hin, falls weniger Messversuche für jeden Proband aufgezählt werden, als ihn tatsächlich repräsentieren. Für einige Probanden reichten 100 Messversuche nicht aus, um eine Konvergenz des Mittelwertes zu erreichen. Dem inkonsistenten Verhalten von konvergenten und divergenten kumulierten Mittelwertkurven folgt, dass es nicht möglich ist, ein konvergentes Probandenkollektiv vorab einer Messung zusammenzustellen, zumal die Auswahl der Probanden entsprechend der zu untersuchenden Grundgesamtheit erfolgen und sich inhaltlich an der gestellten Aufgabe orientieren sollte. Bei konsequenter Weiterführung der Messung der üblichen fünf Versuche im Labor ergibt sich ein Messfehler, der weder kontrolliert noch dessen Größe richtig eingeschätzt werden kann. Diese Auswertungen offenbaren, dass Fehler bei der Mittelwertbildung fern der tatsächlichen Wirklichkeit entstehen.

Die Auswirkungen dieser Messfehler pflanzen sich bis zum Einsatz statistischer Hilfsmittel und Testverfahren fort. Dies sei an einem Beispiel demonstriert¹². Fiktiv soll der Einfluss einer unabhängigen Untersuchungsbedingung mit zwei Ausprägungen auf eine abhängige Variable z. B. auf die vertikale Bodenreaktionskraft analysiert werden. Nach der erfolgreich abgeschlossenen Messung im Labor beginnt die statistische Auswertung, welche die Mittelwertbildung der fünf berechneten diskreten Parameter z. B. der maximalen Kraftanstiegsrate und den Einsatz des T-Tests für zwei unabhängige Stichproben umfasst. Angenommen, es existiert ein den Probanden repräsentierender Mittelwert, welcher mit einer endlichen Anzahl größer fünf gemessen werden kann. Bei der Verwendung von fünf Versuchen entsteht dann automatisch ein Fehler. Der gemessene Mittelwert für jeden Proband r wird mit $\overline{\mu}_r = \mu_r + E(\varepsilon_r)$ angegeben, wobei μ_r dem tatsächlichen (wahren) Mittelwert und $E(\varepsilon_r) \neq 0$ dem mittleren Messfehler entspricht¹³.

Entsprechend dem Testwert des T-Tests für unabhängige Stichproben

$$T_{intra} = \frac{\overline{MW}_1 - \overline{MW}_2}{\sqrt{\frac{(n_1 - 1)\overline{SD}_1^2 + (n_2 - 1)\overline{SD}_2^2}{n_1 + n_2 - 2}} / \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (6.1)$$

müssen Stichprobenmittelwerte \overline{MW}_1 , \overline{MW}_2 und Varianzen \overline{SD}_1^2 , \overline{SD}_2^2 gebildet werden. Im Fall eines Offsets (alle Probanden generieren denselben Messfehler $E(\varepsilon_r)$) ergeben sich $\overline{MW}_1 = MW_1 + E(\varepsilon_1)$, $\overline{MW}_2 = MW_2 + E(\varepsilon_2)$, $\overline{SD}_1^2 = SD_1^2$ und $\overline{SD}_2^2 = SD_2^2$ (↗ Gleichungen S. 129). Ein Offset hat somit auf den Testwert aus Gleichung 6.1 und auf das entsprechende Testergebnis keine Auswirkung. Lediglich das Ausmaß der diskreten Parameter ist veränderlich. Fehler, welche zum Beispiel durch eine zu niedrige Filterfrequenz entstehen, können den Ausgang der statistischen Untersuchung somit nicht verfälschen.

Im Fall, dass die Messfehler $E(\varepsilon_r) \neq 0$ für alle Probanden unterschiedlich ausfallen, ist der Einfluss erheblicher Natur. Mit den entsprechenden Mittelwerten und Stichprobenvarianzen (↗ Gleichungen

¹² Die statistischen Hilfskonstruktionen und Berechnungen sind im Anhang E ab Seite 129 notiert.

¹³ Siehe auch Gleichung 1.2.1 (S. 9) zur Interpretation des aufgenommenen Signals S_T anhand des wahren Wertes und entsprechender Messfehler.

S. 130) ändert sich der Testwert wie folgt:

$$T_{intra} = \frac{(MW_1 - MW_2) + (E(\varepsilon_1) - E(\varepsilon_2))}{\sqrt{\frac{(n_1 - 1)(SD_1^2 + SD(\varepsilon_1)^2) + (n_2 - 1)(SD_2^2 + SD(\varepsilon_2)^2)}{n_1 + n_2 - 2}} / \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}. \quad (6.2)$$

Die Kombination aus mittleren Fehlern und Fehlervarianzen kann zu drei Ausgängen führen, wobei auf eine genaue Zusammensetzung der verschiedenen Fälle verzichtet und lediglich deren Auswirkung auf den Testwert eruiert wird¹⁴:

1. Der Testwert wird unterschätzt und ist somit kleiner als der Testwert, welcher anhand des wahren Wertes berechnet wird: Die Nullhypothese wird in diesem Fall „später“ abgelehnt. Falls die Nullhypothese zutrifft, wird an dieser Stelle kein Fehler begangen. Ist die Nullhypothese in Wirklichkeit unwahr, dann wird der Fehler zweiter Art (β) größer (\nearrow Tab. E.1). Dies tritt gerade dann ein, wenn die mittlere Fehlerdifferenz zwischen den Gruppen erheblich kleiner ist als die Streuung der entsprechenden Fehler.
2. Der Testwert wird überschätzt und ist daher größer als der wahre Testwert: Die Nullhypothese wird „eher“ abgelehnt. Entspricht die Alternativhypothese der Realität, dann hat der Fehler keinen Einfluss auf das Ergebnis. Trifft hingegen die Nullhypothese zu und müsste das Ergebnis eigentlich nicht signifikant ausfallen, dann wird durch den Messfehler eine Erhöhung des α -Fehlers begangen. Dieser Umstand tritt vor allem dann ein, wenn die mittlere Fehlerdifferenz wesentlich größer ausfällt als die summierte Fehlervarianz.
3. Der dritte mögliche Ausgang besteht in gleichbleibenden Testwerten.

Durch Beibehaltung der ursprünglichen Messprozedur ist somit die Wahrscheinlichkeit sehr hoch, dass der begangene Fehler den Ausgang der statistischen Auswertung beeinflusst und damit Eigenschaften der vorliegenden Grundgesamtheit unter Umständen falsch widerspiegelt. Interventionsmöglichkeiten sind aufgrund der Unkenntnis der tatsächlich begangenen Fehler rar gesät. Der Erhöhung des α -Fehlers kann mit einer Korrektur der Irrtumswahrscheinlichkeit entgegengewirkt werden. Dies minimiert zunächst die Wahrscheinlichkeit einer falsch-positiven Entscheidung. Eine konservativere Beurteilung des statistischen Tests bedingt gleichzeitig eine Verschlechterung der Power. Im ersten Fall der möglichen Einflüsse würde eine Verringerung von α eine Signifikanz bei einem höheren Testwert verursachen. Bei einer Unterschätzung des Testwertes würde dies sogar einen größeren Fehler verursachen als unter gewohnten Umständen. Die Belassung des α -Fehlers auf dem ursprünglichen Niveau nährt jedoch die Zweifel am Ausgang und der Aussage des Testes und damit verbunden an den Erkenntnissen im Rahmen biomechanischer Messungen. Die zusätzliche Beeinflussung der begangenen Messfehler bezüglich der Verteilung der Daten unterstreicht zudem die Befürwortung der Verwendung von nichtparametrischen Verfahren. Deren eher konservative Beurteilung der Differenzen anhand Rangbildung zwischen und innerhalb der Gruppen könnte die fehlerhaften Einflüsse minimieren. Die Kontrolle der entsprechenden Fehlerarten ist unter Verwendung von nichtparametrischen Verfahren jedoch nicht mehr gegeben.

Diese Diskrepanzen verdeutlichen, dass von einer Verwendung des Mittelwertes diskreter Parameter

¹⁴ Folgende Möglichkeiten und deren Kombinationen aus den einzelnen mittleren Fehlern und Fehlervarianzen sind denkbar: $E(\varepsilon_1)/E(\varepsilon_2)$ sind größer oder kleiner als Null und $SD(\varepsilon_1)^2/SD(\varepsilon_2)^2$ sind größer oder kleiner als Null. Diese Optionen beeinflussen die mittlere Fehlerdifferenz und die Fehlervarianzsumme.

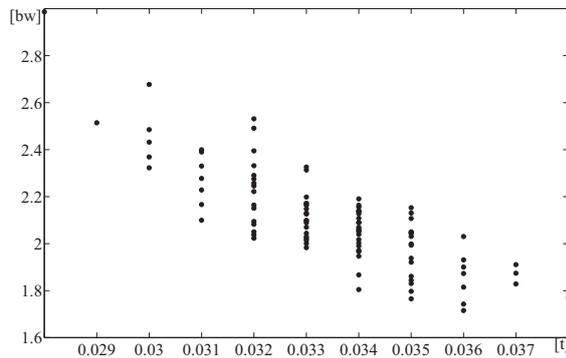
abgesehen werden sollte, da der Fehler sowie dessen unbekanntes Ausmaß zum Teil unkontrollierbar und dessen Auswirkungen auf weitere biomechanische Kenngrößen nicht überprüfbar sind. Von einem Offset, welches keinen Einfluss auf das Ergebnis der statistischen Auswertung hätte, kann aufgrund der großen Streuung der absoluten Differenzen nicht ausgegangen werden. 5 % Abweichung erscheinen im empirischen Kontext als vertretbar, da sie innerhalb der Schwankung des Probandenkollektivs liegen (RMSE). Dessen Konsequenzen für die Beurteilung von Untersuchungsbedingungen sind hingegen immens. Hierbei sei zusätzlich erwähnt, dass unter Akzeptanz der möglichen Fehler die mittleren zu bewertenden Unterschiede stets anhand der Schwankung innerhalb der Probanden gemessen mittels Root Mean Square Error bewertet werden sollten (Maiwald et al., 2011). Dies minimiert zudem die Wahrscheinlichkeit einer falschen statistischen Entscheidung und relativiert die tatsächlichen Effekte, wobei die Beeinflussung von kinematischen Messungen nicht eingedämmt werden kann. Diese Fehler werden bis zur Verwendung von statistischen Hilfsmitteln weiter getragen.

Die obigen Punkte zeigen, dass die Nichtbeachtung der Ergebnisse von Hypothese I und II bezüglich diskreter Parameter fatale Folgen haben kann. Den bisherigen und zukünftigen Ergebnissen biomechanischer Studien unter der Messung im Labor, welche anhand diskreter Parameter diskutiert wurden, muss mit Skepsis begegnet werden. Die von vornherein akzeptierte Irrtumswahrscheinlichkeit in empirischen Wissenschaften könnte um ein Vielfaches größer sein als angenommen, so dass der statistische Ausgang eher einem Würfeln gleichkommt als tatsächlicher Erkenntnisgewinnung. Damit verbunden wird ein aktives Gegenwirken bezüglich dieser Missstände erforderlich und es sollten neue Impulse bezüglich messtechnischer Umsetzung und statistischer Auswertung daraus resultieren.

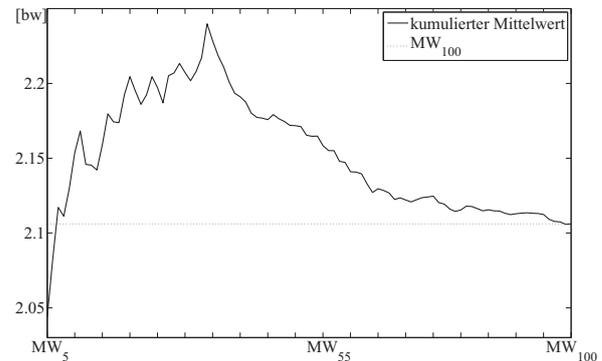
Die Informationsreduktion auf einen einzigen Wert resultiert in einem Informationsverlust, welcher durch eine mehrdimensionale Betrachtung der diskreten Parameter oder durch die Analyse der gesamten Kurven behoben werden könnte. Die Beurteilung der diskreten Parameter ohne ihre zeitliche Komponente führt zu einer Verzerrung der tatsächlichen Information, welche zum Beispiel die erste Kraftspitze beinhaltet. Tritt F_{z_I} zeitlich später auf, ist dies häufig mit einer Reduktion der Kraft verbunden (siehe Abbildung 6.1). Für dieses Beispiel ergibt sich für die Folge kumulierter Mittelwerte von F_{z_I} eine annähernd divergente Folge (Abb. 6.2).

Für die Auswertung zwei- oder mehrdimensionaler metrischer Daten wartet die herkömmliche Statistik lediglich mit Korrelationen, anderen Zusammenhangsmaßen und multiplen Regressionsanalysen auf. Methoden, in denen die abhängigen Variablen gleichwertig in die Auswertungen eingehen, sind rar gesät. Eine Möglichkeit könnte in der Weiterentwicklung der „point cluster technique“ liegen, welche von Andriacchi et al. (1998) publiziert wurde. Bei dieser Methode werden alle relevanten Daten in einer Matrix betrachtet, deren Kovarianzmatrix sowie zugehörige Eigenwerte und Eigenvektoren berechnet. Diese ermöglichen eine Aussage und Bewertung über die Schwankung innerhalb dieser Kovarianzmatrix.

Eine weitere Möglichkeit, der Informationsreduktion entgegenzuwirken, besteht in der Verwendung der gesamten Kurve. Die Ergebnisse bezüglich Hypothese I und II für Gesamtkurven zeigen, dass die Informationen, welche die gesamte Kurve liefert, durchaus zur Beschreibung eines Probanden herangezogen werden kann. Für gesamte Kurven sind bisher keine Auswertungsmöglichkeiten etabliert, so dass die Daten zunächst lediglich deskriptiv ausgewertet werden können. Die Entwicklung adäquater statistischer Methoden für diese Auswertung sollte im Interesse der ganzen biomecha-


Abbildung 6.1:

Beispiel der Einzelwerte für die erste Kraftspitze mit zeitlicher Dimension


Abbildung 6.2:

Zugehörige Folge kumulierter Mittelwerte (keine Stabilisierung)

nischen Wissenschaftsgesellschaft liegen. Bis dahin wird sich zunächst eine „Deskriptive Ära“ zur Analyse von gesamten Kurven verbreiten, in der sich Forscher beschreibender Methoden bedienen müssen. Dies bedeutet zudem, dass lediglich Aussagen über Probandenkollektive getroffen und keine generellen für die Grundgesamtheit geltenden Thesen aufgestellt werden können.

Ein möglicher Ansatz zur Beschreibung gesamter Kurven liegt in der Auswertung der Prädiktionsbänder und damit der Beschreibung der Gesamtschwankung. Die sich nicht überlappenden Flächen von zwei Prädiktionsbändern, welchen unterschiedliche Untersuchungsbedingungen zugrunde liegen, könnten in die statistische Auswertung aufgenommen werden. Bezogen auf die Ausführungen zum menschlichen dynamischen System und die positiven Auswirkungen von Variabilität (↗ S. 8 f.) besteht in der Verwendung von statistischen Parametern, welche sich auf die Schwankung eines Probanden beziehen, eine Möglichkeit der Erweiterung der zu verwendenden Methoden. Die Einschätzung und Bewertung dieser Variabilität ist dabei von großer Bedeutung. Ein schmaleres Prädiktionsband für vertikale Bodenreaktionskräfte ist mit einem kleineren Anteil am gesamten Kraftimpuls und damit mit weniger Variabilität verbunden. Im Hinblick auf die Übertragung auf Gelenkwinkel und damit verbundene mögliche schmalere Prädiktionsbänder ist unter Umständen eine Restriktion des tatsächlichen Bewegungsausmaßes und der entstehenden Variabilität impliziert. Bedingungen, welche diese Variabilität einschränken, könnten im Sinne einer positiv zu bewertenden Variabilität als eher „schlechter“ für den Probanden eingestuft werden (Goldberger et al., 1990; Stam et al., 1994; van Emmerik & van Wegen, 2000). Dies bedarf weiterer Untersuchungen, inwiefern das Ausmaß der Prädiktionsbänder in Bezug zur Bewertung der Variabilität eines Probanden zu setzen ist.

Des Weiteren sei zu beachten, dass generell eine Einzelfallanalyse einer Mittelwertsüberprüfung vorgezogen werden sollte. Im Hinblick auf die oben genannten Probleme bezüglich der getroffenen Annahmen über α - und β -Fehler und der Aufnahme der minimal notwendigen Messversuche treten Fehler lediglich für divergierende Probanden auf. Die separate Analyse für jeden Proband minimiert letztendlich die Missinterpretation, obwohl die Fehler nicht vollständig eliminiert werden können. Die minimal zu verwendende Anzahl an aufzuzeichnenden Messversuchen muss dabei im Hinterkopf behalten werden, da diese gewährleistet die gesamte Variabilität abzubilden. Wird diese über- oder

unterschätzt, werden im Fall der Untersuchung von Prädiktionsbändern wiederum Fehler begangen. Die Auswertung dieser entstehenden Unterschiede für verschiedene Untersuchungsbedingungen muss weiter erforscht und deren technische Anforderungen verbessert werden.

Aussagen mittels Einzelfallanalysen über das gesamte Probandenkollektiv können anhand Dichotomisierung der Ergebnisse und damit verbunden mittels χ^2 -Tests getroffen werden. Diese setzen eine große Stichprobengröße voraus, wobei dies kontroversen Ergebnissen, welche häufig aufgrund geringer Stichprobengrößen bemängelt werden, entgegenwirken würde.

Die Aufstockung der Stichprobengröße ist in statistischer Hinsicht optimal, um die Ergebnisse zu untermauern. Eine vorgelagerte Stichprobenplanung würde die Glaubhaftigkeit der angewendeten Verfahren zudem stützen. In diesem Zusammenhang treten zwei neue Probleme auf - die hohe Anzahl an zu messenden Versuchen und die nicht vorhandenen Verfahren zur Stichprobenplanung. Erst nach der Entwicklung neuer Bewertungsmethoden können zusätzliche statistische Instrumente zum Beispiel zur Beschreibung von Effektstärken oder zur Stichprobenplanung konstruiert werden.

Die in der hier dargestellten Literatur verwendeten 6 ± 3 Messversuche (\nearrow Tab. 1.3.1, S. 11) konnten mit keiner der in dieser Arbeit angewendeten Methoden bestätigt werden. Maximal wurden 60 bis 70 Versuche (vor allem anhand visueller Überprüfung) ermittelt. Die anhand Populationsdifferenzkurve im Abgleich mit den Peaks der Differenzen der sich nicht überlagernden Prädiktionsbänder ermittelte Anzahl belief sich auf rund 34 Versuche. Aus ökonomischer Sicht sind Studien mit einer mittleren minimalen Anzahl von bis zu 35 bis 60 Versuchen bei mehr als 20 Probanden und mehr als 2 Untersuchungsbedingungen kaum durchführbar. Dies unterstreicht u. a. die Forderung, die Studienanzahl auf wesentliche Inhalte zu beschränken (Weihe, 2004).

Die bisherigen Messprotokolle zur Messung von Bodenreaktionskräften im Labor vor allem in Bezug auf die Anzahl der Versuche und auf die spätere Datenauswertung müssen aufgrund dieser Auswertungen als fehlerbehaftet gelten. Die mögliche Beeinflussung der minimal zu verwendenden Anzahl und der Schwankung des Probanden durch Ablenkung oder Motivationsschübe ist zudem eine Einschränkung der Messung im Labor. Dies macht die Erforschung alternativer Methoden zur Datenaufnahme sowie zur Weiterentwicklung und Verbesserung bestehender Methoden unumgänglich. Falls eine Messung im Labor nicht vermieden werden kann, sollte die Verwendung der Gesamtkurven wie weiter oben beschrieben forciert werden. Hierfür fehlen sowohl die zu betrachtenden Parameter als auch die statistischen Auswertungsmethoden, welche in zukünftigen Projekten entwickelt werden sollten.

Bisher wurde der Charakter des Ausdauerlaufens als zyklische Bewegung nicht beachtet und der Aspekt der Zerlegung dieser Bewegung bisher nicht erörtert. Die Laufbewegung per se stellt einen zyklischen Bewegungsablauf dar, der während der herkömmlichen Messung im Labor nicht vollständig aufgenommen werden kann und lediglich in etwa jeder 20. Schritt in die Auswertung eingeht. Die Bewertung anhand von Gesamtkurven nähert sich der zyklischen Bewegung, jedoch werden die ungefähr 19 Schritte, welche die Bewegung komplettieren, vollkommen ignoriert. Es liegt die Vermutung nahe, dass eine zyklische Bewegung einen zyklischen Mittelwert generiert und dies die Berechnung von diskreten Parametern und deren Mittelwert sowie die Berechnung einer mittleren Gesamtkurve der einzelnen Schritte ad absurdum führen würde. Die Abbildung und Auswertung dieser zyklischen Bewegung erfordert andere Messprozeduren sowie weitere Bewertungsmethoden. Eine Möglichkeit besteht in der Verwendung von Laufbändern. Diese können die Bewegung über

einen längeren Zeitraum aufzeichnen. Die Bewegung wird hingegen in eine etwas andere Form abgewandelt. Dies wird durch die Tatsache begründet, dass sich zum einen der Untergrund bewegt und zum anderen dadurch andere Muskeln und neuronale Verbindungen hergestellt werden müssen. Häufig lehnen sich die Probanden mehr nach vorn, die Schrittlänge verkürzt sich und die Schrittfrequenz wird erhöht (Wank et al., 1998; Fellin, Manal & Davis, 2010). Diese Untersuchungen verglichen jedoch die Messung im Labor anhand der herkömmlichen Messprozedur mit der auf einem Laufband. Ein Vergleich zwischen Laufen in der freien Natur und auf einem Laufband wurde bisher nicht erbracht. Somit ist unklar, inwiefern dieser gegenüber der Labormessung modifizierte Lauf mit dem im Feld übereinstimmt. Eine Annäherung an die tatsächliche Laufbewegung lässt sich vermuten und sollte in zukünftigen Untersuchungen näher betrachtet werden.

Diese Überlegungen stützen zudem die Vermutung, dass generell die Messung und Auswertung von zyklischen Bewegungen oder Datenreihen, welche über einen längeren Zeitraum aufgenommen werden, anhand von einzelnen Aufnahmeverfahren und statistischen Bewertungen keine adäquate Methode darstellen. Zeitreihen sollten mit gesonderten Bewertungsmethoden ausgewertet werden.

In der Entwicklung von Messgeräten und entsprechenden Auswertungsverfahren, welche in der freien Natur eingesetzt werden können, besteht eine weitere Alternative zu herkömmlichen Messprotokollen. Diesem Forschungsgebiet wurde bisher sehr wenig Aufmerksamkeit geschenkt. Meardon, Hamill und Derrick (2011) verwenden einen Beschleunigungsaufnehmer an der Tibia, um Aussagen über einen Lauf auf einer 300 m Bahn treffen zu können. Wenige Methoden zur Auswertung von Zeitreihen existieren bereits in der Bewegungswissenschaft, wie zum Beispiel der Lyapunov Exponent (Bruijn et al., 2009a, 2009b) oder der scaling Exponent α einer trendbereinigenden Fluktuationsanalyse (Nakayama et al., 2010), wie ihn auch Meardon et al. (2011) verwenden. Bei der Entwicklung und Etablierung von Auswertungsmethoden und Messgeräten sowie der Bewertung der entsprechenden Ergebnisse besteht durchaus Nachholbedarf. Es ist in Bezug auf die Analyse der Variabilität von Zeitreihen des Ausdauerlaufens ein eher chaotisches Verhalten zu erwarten. Meardon et al. (2011) zeigten, dass die Schrittfrequenz über die Zeit nicht abgeschätzt werden kann und einen vielmehr zufälligen Charakter aufweist. In den Untersuchungen von Jordan, Challis und Newell (2006) wurden Laufgeschwindigkeiten fern der bevorzugten als beeinflussender Faktor für die Komplexität des menschlichen Systems bewertet, wobei diese eine höhere Vorhersagbarkeit für unübliche Geschwindigkeiten aufwiesen. Dies lässt vermuten, dass sich die Bewertung der Güte einer „gesunden“ Variabilität eher chaotischen Systemen nähert und zudem diesbezüglich eingeschätzt werden muss. Nähert sich die Variabilität für verschiedene Probanden eher periodischen Prozessen, könnte dies möglicherweise auf z. B. weniger Flexibilität des neuromuskulären Systems hinweisen und unter Umständen eine Einschränkung darstellen. Weitere Untersuchungen sind an dieser Stelle unumgänglich, um den Einfluss und die damit verbundene Bewertung der Variabilität zu diskutieren und mögliche Erklärungsansätze zu eruieren.

Des Weiteren werden häufig verschiedene Untersuchungsbedingungen z. B. diverse Schuhmodelle während einer Studie analysiert. Die Bewertung von diesbezüglichen Unterschieden wirft die Frage auf, ob der Schuh oder eine andere Bedingung aktiv die Bewegung des Probanden beeinflussen kann oder ob die Probanden bezogen auf das große Schwankungsspektrum, gemessen anhand RMSE, selbst den größeren Einfluss auf ihr Bewegungsmuster ausüben. Diese Überlegung macht alternative Auswertungsmethoden fern der quantitativen und metrischen Auswertung attraktiv. Die Bewertung von Passform und Komfort könnte aufgrund der konzeptionellen Probleme bisher ver-

wendeter Messprozeduren verstärkt in den Vordergrund rücken. Diverse Arbeiten wurden in dieser Richtung schon unternommen (Schubert, Oriwol & Sterzing, 2011; Schubert, 2010; Kunde, 2010) und könnten in der Entwicklung eines Komfort- oder Passformfragebogens münden. Die bisherigen Auswertungen vor allem von Schubert (2010) zeigen, dass die befragten Personen prozentual mehr Wert auf Passform und Bequemlichkeit legen. Dies bestärkt die Nachfrage nach einer Bewertung von Komfort und Passform. Neben der Weiterführung dieser Untersuchungen liegt in qualitativen Auswertungsmethoden eine weitere Option zur Bewertung von unterschiedlichen Schuhbedingungen, wobei die Bewertung durch den Probanden in den Vordergrund gestellt wird (Tenenbaum & Driscoll, 2005).

Die Begrenzung der Studienanzahl und damit einhergehend die Beschränkung auf wichtige Inhalte, die Erhöhung der Studiendauer aufgrund einer erhöhten Anzahl an aufzuzeichnenden Messversuchen, einer nicht verwendbaren herkömmlichen Methode sowie die notwendige Entwicklung neuer Mess- und Auswerteprozeduren hat weitreichende Konsequenzen innerhalb des Mikro- und Makrokosmos der biomechanischen Wissenschaftsgesellschaft (siehe Abbildungen 6.3 und 6.4).

Die Konsequenzen für die beteiligten Wissenschaftsräume müssen in Bezug zu deren entsprechenden Einflussfaktoren betrachtet werden. Zu diesen zählen in Bezug auf den Mikrokosmos einer einzelnen Arbeitsgruppe die aus universitärer Sicht erwünschte Einwerbung von Drittmitteln und Haltung von externen Geldgebern sowie die Universität selbst als Arbeitgeber. Der Makrokosmos bezieht sich auf die biomechanische Wissenschaftsgesellschaft, welche national und international außerhalb der eigenen Arbeitsgruppe agiert und wissenschaftlich tätig ist. Veröffentlichungen innerhalb dieser Gemeinschaft tragen zu wissenschaftlicher Erkenntnis bei. Die damit verbundene Reputation erhöht das Ansehen und den möglichen Einfluss der Arbeitsgruppe innerhalb dieses Makrokosmos. Somit müssen die Auswirkungen dieser Arbeit bezüglich der Charakterisierung des Probanden anhand des Mittelwertes und deren Bedeutung für den Ausgang wissenschaftlicher Studien innerhalb des Mikro- und Makrokosmos der Bewegungswissenschaft diskutiert werden.

Mikrokosmos Die ökonomische Begrenzung der Studienanzahl impliziert, dass die Forschungsansätze und Studieninhalte einer Selektion unterzogen werden müssen. Dabei gilt es „wichtige“ von „unwichtigen“ Inhalten zu trennen. Dies ist in Bezug auf Forschung und Entwicklung eine enorme Einschränkung, minimiert diversifizierte Forschungsfelder und veranlasst den Forscher, notwendige, jedoch „kleine“ Forschungsfragen zu ignorieren. Dies kann zudem die Erhöhung der Komplexität des Studienaufbaus bewirken, wobei die entsprechenden Methoden ferner entwickelt werden oder zu viele Voraussetzungen erfüllt sein müssen, so dass von der Steigerung der Vielschichtigkeit von Studien abzuraten ist.

Einer Verringerung der Studienanzahl folgt die Minimierung der Anzahl an Veröffentlichungen. Dies impliziert wiederum die Beschneidung der Geldmittel, welche durch die Universität der Arbeitsgruppe zur Verfügung gestellt werden und an die Anzahl an Publikationen gekoppelt ist. Dieses System, anhand der Menge von Publikationen und Zitierungen den Gehalt der Wissenschaft zu messen, ist nicht adäquat gewählt. Die Methode der Szientometrie untersucht lediglich die Quantität von Veröffentlichungen nicht aber ihre Qualität. Lediglich die Menge an Zitationen ist für die Messung wichtig, nicht aber ihre inhaltliche Verwendung. Wird eine Referenz im negativen Sinne gebraucht und deutet somit auf eine qualitativ weniger wertvolle Publikation, ist dies zunächst für die Berechnung von zum Beispiel Impact-Faktoren für Zeitschriften oder Hirschfaktoren für Au-

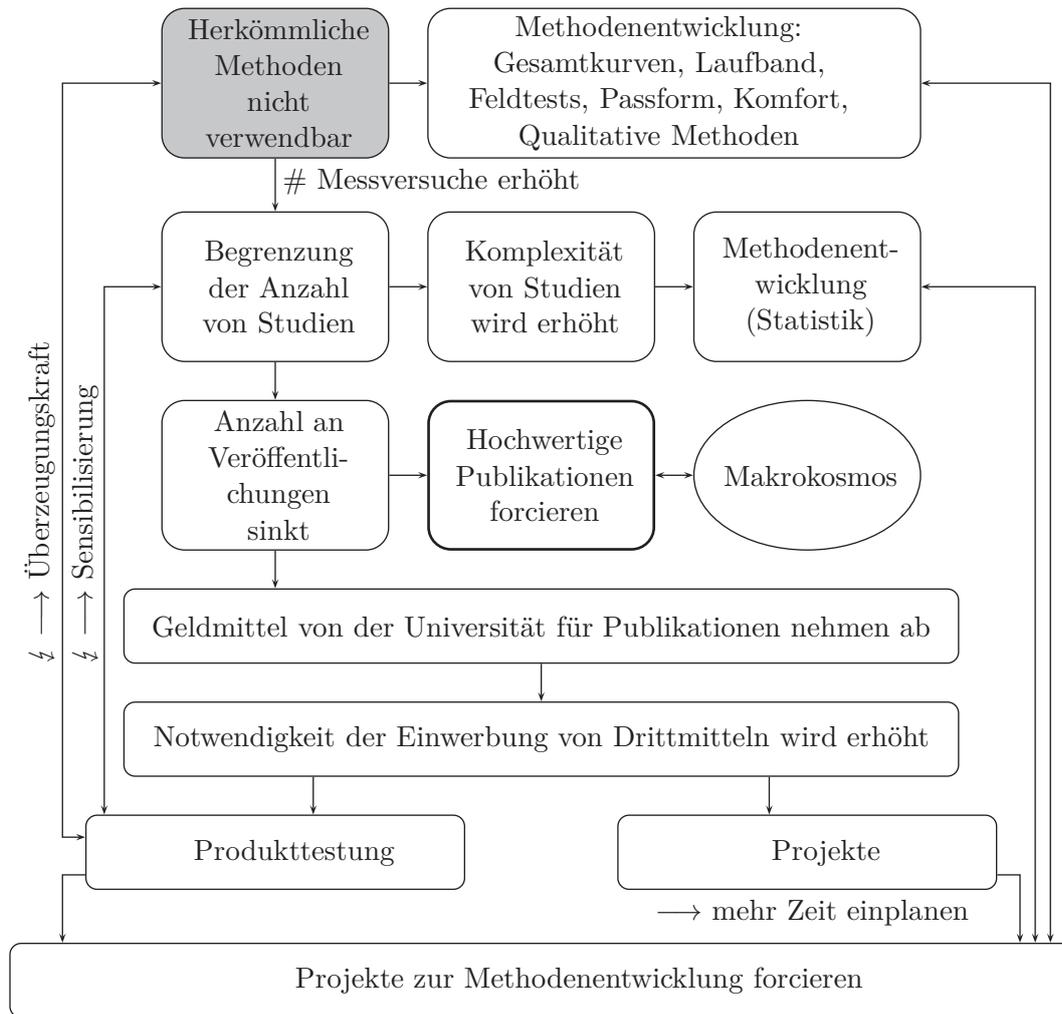


Abbildung 6.3: Auswirkung der Ergebnisse hinsichtlich fehlerhafter herkömmlicher Methoden auf den Mikrokosmos der biomechanischen Wissenschaftsgesellschaft

toren unerheblich. Dieses fehlerbehaftete System könnte den Forscher aufgrund der Kopplung an wissenschaftliche Gelder animieren, Studien trotz methodischer Fehler zu publizieren.

Eine Möglichkeit, sich anderweitig zu finanzieren, liegt in der Einwerbung von Drittmitteln. Diese können zum einen über Produkttestung und Entwicklung für externe Firmen generiert werden. In der Antragstellung von Projekten über Forschungsgesellschaften oder über kompetitive nationale und internationale Förderprogramme liegt die zweite Variante der Drittmittelinwerbung. Nach den obigen Erkenntnissen sollte generell auf eine erhöhte Studiendauer geachtet und dies in der Antragstellung zeitlich eingeplant werden. Innerhalb einer neuen Antragstellung ist dies ohne weiteres möglich, wobei die Sensibilisierung bereits gewonnener Kooperationspartner bezüglich der neuen Erkenntnisse schwieriger ist. Hier bedarf es einer enormen Überzeugungskraft, nicht nur in Hinsicht auf die methodische Fehlentwicklung der letzten Jahre, sondern auch auf die Beschränkung der

Studienanzahl und damit einhergehend der Einschränkung auf ausgewählte Forschungsinhalte. Es gilt den Industriepartner für Forschung und Entwicklung zu begeistern und zu verdeutlichen, dass Investitionen in eben diese, vor allem innerhalb einer universitären Einrichtung, zudem die Generierung von Wissen beinhalten. Empirische Wissenschaft beruht auf Versuch und Irrtum. Wichtig ist, die neuen Erkenntnisse nicht zu ignorieren und für die weitere Entwicklung zu nutzen, wodurch Fortschritt und Verbesserungen hinsichtlich der Produktentwicklung und Testung vorangebracht werden.

Ein Imageschaden, wie er durch eine bewusste Falschdarstellung der Wirklichkeit entstehen kann, ist durch diese Arbeit nicht zu erwarten. Fehlinterpretationen dürfen nicht bewusst begangen und Studienergebnisse müssen nach bestem Wissen und Gewissen erklärt werden. Die Erkenntnisse dieser Arbeit sollten als Chance verstanden werden, die eigene Forschung zu verbessern und neue Standards und Methoden zu entwickeln. Die neuen Entwicklungen, Methoden und Messprozeduren haben Potential, sowohl für die involvierte Arbeitsgruppe eine Vorreiterrolle einzunehmen als auch für die Antragstellung von Projekten durch Innovation zu bestehen. Dabei sollten Projekte, welche sich der methodischen Weiterentwicklung der Messung des Ausdauerlaufens entweder im Labor oder unter realistischeren Bedingungen widmen, zudem forciert werden.

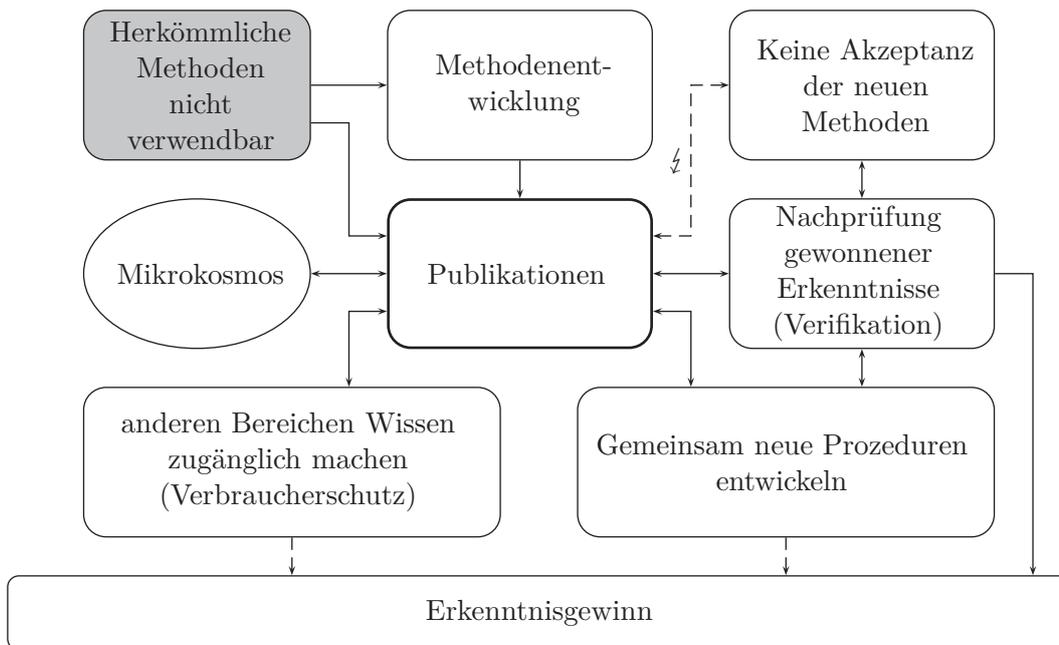


Abbildung 6.4:

Auswirkung der Ergebnisse hinsichtlich fehlerhafter herkömmlicher Methoden auf den Makrokosmos der biomechanischen Wissenschaftsgesellschaft

Makrokosmos Generell sollten die Informationen bezüglich der Erkenntnisse in dieser Arbeit sowie zukünftiger Entwicklungen neuer Methoden der nationalen und internationalen biomechanischen Wissenschaftsgesellschaft mittels Publikationen zur Verfügung gestellt werden. Publikationen sind das Verbindungsglied zwischen einzelnen Arbeitsgruppen und der gesamten biomechanischen

Wissenschaftsgesellschaft. Vor allem die geforderte Abkehr von herkömmlichen Auswertungsmethoden bezogen auf die Mittelwertbildung diskreter Parameter der Bodenreaktionskräfte und der Winkelgeschwindigkeit sollte publiziert werden, um gemeinsam Lösungen und neue Prozeduren entwickeln zu können. Dabei ist aus erkenntnistheoretischer Sicht zu beachten, dass die gewonnenen Erkenntnisse überprüft werden müssen. Fehlerfreiheit kann in einem empirischen Kontext nicht gewährleistet werden. Erst die Verifikation der Resultate durch viele Arbeitsgruppen ermöglicht Erkenntnisgewinn.

Zudem sollten alle Ergebnisse veröffentlicht werden und der „publication bias“ verringert werden. Neben den für die Entwicklung der aufgestellten Theorie „positiven“ vor allem signifikanten Ausgängen muss die Publikation von nicht signifikanten oder „negativen“ Ergebnissen forciert werden. Lediglich die gesamte Ergebnislage kann Hinweise auf das Zutreffen einer Hypothese geben. Dabei sind neben Forschenden auch Reviewer aufgefordert diesem Trend entgegenzuwirken. Dazu zählt zudem die Offenheit gegenüber Neuerungen und Publikationen mit ähnlichem Inhalt. Eine Verifizierung der Ergebnisse ist erst dann möglich, wenn andere Arbeitsgruppen von diesen Ergebnissen erfahren. Werden Publikationen von vornherein abgelehnt, weil die Methodik nicht der üblichen und etablierten Vorgehensweise entspricht oder weil ähnliche Publikationen bereits existieren, dann wird der Schritt in Richtung Erkenntnisgewinn von vornherein torpediert. Dies hat wiederum Auswirkungen auf die einzelnen Arbeitsgruppen. Können Studien aus diversen Gründen nicht veröffentlicht werden, gerät der Forscher in den szientometrischen Teufelskreis. Ist die Wahrscheinlichkeit höher, mit einer etablierten, aber möglicherweise inkorrekten Methodik Ergebnisse veröffentlichen zu können, befindet sich der Forscher in einem Dilemma und steht vor den Optionen der wissenschaftlichen Interpretation der Daten anhand „korrekter“ Methoden und des finanziellen Überlebens. Dem Umstand, dass der Forscher gezwungen wird, zwischen diesen Optionen zu wählen, muss vehement entgegengewirkt werden. Dies sollte im Interesse der globalen biomechanischen Wissenschaftsgesellschaft liegen. Ein Dialog zwischen verschiedenen Arbeitsgruppen ist an dieser Stelle unumgänglich. Es gilt, gemeinsam Lösungen zu finden und zu verifizieren. Vor dieser Wahl sollten zudem die Industriepartner der eigenen Arbeitsgruppe und externe Firmen, welche die in der Biomechanik etablierten Prozeduren verwenden, geschützt werden. Dies kann wiederum lediglich mit Publikationen, öffentlichen Dialogen und einer mikrokosmenübergreifenden Zusammenarbeit und Projekten geschehen. Letztlich müssen die biomechanische Wissenschaftsgesellschaft sowie die Verbraucher und Konsumenten der jeweiligen Ergebnisse vor Missinterpretationen geschützt werden.

In der Konsequenz dieser Arbeit muss festgestellt werden, dass sich die biomechanische Wissenschaftsgesellschaft an einem Scheideweg befindet. Die bisherigen Erkenntnisse biomechanischer Studien, welche Bodenreaktionskräfte und Winkelgeschwindigkeiten in einer Laborsituation verwenden, werden in Frage gestellt. Labormessungen sollten aufgrund des momentan lediglich zerlegenden Charakters der Laufbewegung vermieden und Messgeräte entwickelt werden, welche die zyklische Bewegung des Laufens abbilden. Neue Methoden müssen gefunden werden, um eine korrekte statistische Auswertung zu garantieren und damit verbundene Fehlinterpretationen zu vermeiden. Diese Verfahren sollten zudem die Bewertung der Variabilität eines Probanden in einem neuen Kontext betrachten. Forschergruppen müssen zusammenarbeiten, um die neuen Hinweise zu verifizieren, gemeinsame Prozeduren für die Auswertung biomechanischer Fragestellungen entwickeln und in wissenschaftstheoretischer Sicht neue Erkenntnis generieren sowie neue Theorien aufstellen.

Die in dieser Arbeit aufgestellten methodischen Kritikpunkte einer Messung des Ausdauerlaufens im Labor könnten eine Erklärung für diverse Diskrepanzen innerhalb der am Anfang vorgestellten Veröffentlichungen aus dem Bereich der Biomechanik liefern. Es ist durchaus vorstellbar, dass die beginnende Diskussion über methodische Aspekte im Zusammenhang mit biomechanischen Messverfahren die verschiedenen Forscher dazu animiert, Resultate und Interpretationen vorangegangener Studien zu überdenken (Frederick, 2011). Die biomechanische Wissenschaftsgesellschaft könnte im Sinne von Sokrates erneut am Anfang wissenschaftlicher Erkenntnis stehen. Ihr obliegt mit der Einsicht, dass sie möglicherweise nichts weiß, die einmalige Chance neues Wissen zu generieren, Defizite zu beheben, global zusammenzuarbeiten und neue Wege zu beschreiten.

Literaturverzeichnis

- Albers, S., Klapper, D. & Konradt, U. (2007). *Methodik der empirischen Forschung* (2. Aufl.). Wiesbaden: Gabler Verlag.
- Andriacchi, T. P., Alexander, E. J., Toney, M. K., Dyrby, C. & Sum, J. (1998). A point cluster method for in vivo motion analysis: applied to a study of knee kinematics. *Journal of Biomechanical Engineering*, 120(6), 743–749.
- Aronson, E., Wilson, T. D. & Akert, R. M. (2004). *Sozialpsychologie* (4. Aufl.). München: Pearson Studium.
- Atkinson, G. & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, 26(4), 217–238.
- Axmann, D., Groten, M., Rupp, F. & Scheideler, L. (2006). Wozu „Mittelwert \pm Standardabweichung“? Und Verwandtes. *Deutsche Zahnärztliche Zeitschrift*, 61(4), 214–219.
- Bates, B., Osternig, L., Sawhill, J. & James, S. (1983). An assessment of subject variability, subject-shoe interaction, and the evaluation of running shoes using ground reaction force data. *Journal of Biomechanics*, 16(3), 181–191.
- Bates, B. T., James, C. R. & Dufek, J. S. (2004). Single-Subject Analysis. In N. Stergiou (Hrsg.), *Innovative Analyses of Human Movement* (S. 3–28). Champaign, IL: Human Kinetics.
- Belli, A., Lacour, J., Komi, P., Candau, R. & Denis, C. (1995). Mechanical step variability during treadmill running. *European Journal of Applied Physiology and Occupational Physiology*, 70, 510–517.
- Brauner, T., Sterzing, T., Gras, N. & Milani, T. L. (2009). Small changes in the varus alignment of running shoes allow gradual pronation control. *Footwear Science*, 1(2), 103–110.
- Bronstein, I. N., Semendjajew, K. A., Musiol, G. & Mühlig, H. (2008). *Taschenbuch der Mathematik* (7. Aufl.). Frankfurt am Main: Verlag Harri Deutsch.
- Bruijn, S. M., Dieën, J. H. van, Meijer, O. G. & Beek, P. J. (2009a). Is slow walking more stable? *Journal of Biomechanics*, 42(10), 1506–1512.
- Bruijn, S. M., Dieën, J. H. van, Meijer, O. G. & Beek, P. J. (2009b). Statistical precision and sensitivity of measures of dynamic gait stability. *Journal of Neuroscience Methods*, 178(2), 327–333.
- Butler, R. J., Davis, I. S. & Hamill, J. (2006). Interaction of arch type and footwear on running mechanics. *American Journal of Sports Medicine*, 34(12), 1998–2005.
- Butler, R. J., Hamill, J. & Davis, I. (2007). Effect of footwear on high and low arched runners' mechanics during a prolonged run. *Gait and Posture*, 26(2), 219–225.

- Cavanagh, P. R. (1990). *Biomechanics of Distance Running*. Champaign, IL: Human Kinetics.
- Cavanagh, P. R. & LaFortune, M. A. (1980). Ground reaction forces in distance running. *Journal of Biomechanics*, 13(5), 397–406.
- Chen, H., Nigg, B. M. & de Koning, J. (1994). Relationship between plantar pressure distribution under the foot and insole comfort. *Clinical Biomechanics*, 9(6), 335–341.
- Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112(1), 155–159.
- Collins, J. J., Imhoff, T. T. & Grigg, P. (1996). Noise-enhanced information transmission in rat SA1 cutaneous mechanoreceptors via aperiodic stochastic resonance. *Journal of Neurophysiology*, 76(1), 642–645.
- Croce, U. D., Leardini, A., Chiari, L. & Cappozzo, A. (2005). Human movement analysis using stereophotogrammetry. Part 4: assessment of anatomical landmark misplacement and its effects on joint kinematics. *Gait and Posture*, 21(2), 226–237.
- Czado, C. & Schmidt, T. (2011). *Mathematische Statistik. Statistik und ihre Anwendungen* (1. Aufl.). Springer.
- DeVita, P. & Bates, B. T. (1988). Intraday reliability of ground reaction force data. *Human Movement Science*, 7(1), 73–85.
- Digby, C. J., Lake, M. J. & Lees, A. (2005). High-speed non-invasive measurement of tibial rotation during the impact phase of running. *Ergonomics*, 48(11-14), 1623–1637.
- Diss, C. E. (2001). The reliability of kinetic and kinematic variables used to analyse normal running gait. *Gait and Posture*, 14(2), 98–103.
- Dittmar, B. (2008). Altes und Neues über Isoperimetrie. *Scientia Halensis - Das Uni-Magazin der MLU Halle*. (URL: <http://www.unimagazin.uni-halle.de/index.php?id=459>)
- Eberhard, K. (1999). *Einführung in die Erkenntnis- und Wissenschaftstheorie. Geschichte und Praxis der konkurrierenden Erkenntniswege*. Kohlhammer.
- Eslami, M., Begon, M., Hinse, S., Sadeghi, H., Popov, P. & Allard, P. (2009). Effect of foot orthoses on magnitude and timing of rearfoot and tibial motions, ground reaction force and knee moment during running. *Journal of Science and Medicine in Sport*, 12(6), 679–684.
- Fahrmeir, L., Künstler, R., Pigeot, I. & Tutz, G. (2007). *Statistik. Der Weg zur Datenverarbeitung* (6. Aufl.). Berlin: Springer Verlag.
- Fellin, R. E., Manal, K. & Davis, I. S. (2010). Comparison of lower extremity kinematic curves during overground and treadmill running. *Journal of Applied Biomechanics*, 26(4), 407–414.
- Ferrandis, R., García, A. C., Ramiro, J., Hoyos, J. V. & Vera, P. (1994). Rearfoot motion and torsion in running: the effects of upper vamp stabilizers. *Journal of Applied Biomechanics*, 10(1), 28–42.
- Fischer, E. P. (1996). *Einstein. Ein Genie und sein überfordertes Publikum*. Heidelberg: Springer Verlag.
- Fischer, L. & Wiswede, G. (2002). *Grundlagen der Sozialpsychologie*. München: Oldenbourg.
- Frederick, E. C. (2011). Editorial. Starting Over. *Footwear Science*, 3(2), 69–70.

- Frey, U. (2007). *Der blinde Fleck. Kognitive Fehler in der Wissenschaft und ihre evolutionsbiologischen Grundlagen*. Heusenstamm: ontos verlag.
- Goethe, J. W. von. (1990). *Faust. Der Tragödie erster Teil*. Leipzig: Verlag Philipp Reclam jun.
- Goldberger, A. L., Rigney, D. R. & West, B. J. (1990). Chaos and Fractals in Human Physiology. *Scientific American*, 37–41.
- Grau, S. & Horstmann, T. (2007). Entwicklung eines Stabilitätslaufschuhs zur Prävention von Achillessehnenbeschwerden - Nike Air Cesium. *Sport-Orthopädie - Sport-Traumatologie - Sports Orthopaedics and Traumatology*, 23(3), 179–184.
- Hamill, J., Bates, B. T. & Holt, K. G. (1992). Timing of lower extremity joint actions during treadmill running. *Medicine and Science in Sports and Exercise*, 24(7), 807–813.
- Hamill, J. & McNiven, S. L. (1990). Reliability of selected ground reaction force parameters during walking. *Human Movement Science*, 9(2), 117–131.
- Hardin, E. C., van den Bogert, A. J. & Hamill, J. (2004). Kinematic adaptations during running: effects of footwear, surface, and duration. *Medicine and Science in Sports and Exercise*, 36(5), 838–844.
- Heidenfelder, J., Sterzing, T. & Milani, T. L. (2010). Systematically modified crash-pad reduces impact shock in running shoes. *Footwear Science*, 2(2), 85–91.
- Hennig, E. M., Valiant, G. A. & Liu, Q. (1996). Biomechanical Variables and the Perception of Cushioning for Running in Various Types of Footwear. *Journal of Applied Biomechanics*, 12, 143–150.
- Hering, E., Martin, R. & Stohrer, M. (2009). *Taschenbuch der Mathematik und Physik* (5., aktualisierte u. erw. Aufl.). Springer.
- Hilgers, R.-D., Bauer, P. & Scheiber, V. (2007). *Einführung in die Medizinische Statistik* (2., verb. u. überarb. Aufl.). Berlin: Springer.
- Hintermann, B. & Nigg, B. M. (1998). Pronation in Runners: Implications for Injuries. *Sports Medicine*, 26(3), 169–176.
- Hreljac, A., Marshall, R. N. & Hume, P. A. (2000). Evaluation of lower extremity overuse injury potential in runners. *Medicine and Science in Sports and Exercise*, 32(9), 1635–1641.
- Huss, W. (1991). *Die Karthager*. Frankfurt am Main: Büchergilde Gutenberg.
- James, C. R. (2004). Considerations of Movement Variability in Biomechanics Research. In N. Stergiou (Hrsg.), *Innovative Analyses of Human Movement* (S. 29–62). Champaign, IL: Human Kinetics.
- Jordan, K., Challis, J. H. & Newell, K. M. (2006). Long range correlations in the stride interval of running. *Gait and Posture*, 24(1), 120–125.
- Kadaba, M. P., Ramakrishnan, H. K., Wootten, M. E., Gaine, J., Gorton, G. & Cochran, G. V. (1989). Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. *Journal of Orthopaedic Research*, 7(6), 849–860.
- Karamanidis, K., Arampatzis, A. & Brüggemann, G.-P. (2004). Reproducibility of electromyogra-

- phy and ground reaction force during various running techniques. *Gait and Posture*, 19(2), 115–123.
- Keuth, H. (2000). *Die Philosophie Karl Poppers*. Tübingen: Mohr Siebeck.
- Knapik, J. J., Trone, D. W., I.Swedler, D., Villasenor, A., Bullock, S. H., Schmied, E., Bockelman, T., Han, P. & Jones, B. H. (2010). Injury Reduction Effectiveness of Assigning Running Shoes Based on Plantar Shape in Marine Corps Basic Training. *American Journal of Sports Medicine*, 38(9), 1759–1767.
- Korhonen, M. T., Suominen, H., Viitasalo, J. T., Liikavainio, T., Alen, M. & Mero, A. A. (2010). Variability and symmetry of force platform variables in maximum-speed running in young and older athletes. *Journal of Applied Biomechanics*, 26(3), 357–366.
- Krämer, W. (1991). *So lügt man mit Statistik*. Frankfurt/Main: Campus Verlag.
- Kunde, S. (2010). *Bewegungswissenschaftliche Objektivierbarkeit von funktionaler Passform am Beispiel des Laufschuhs*. Dissertation, Technische Universität Chemnitz. (URL: <http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa-61839>)
- Kurz, M. J. & Stergiou, N. (2003). The spanning set indicates that variability during the stance period of running is affected by footwear. *Gait and Posture*, 17(2), 132–135.
- Lafortune, M. A., Cavanagh, P. R., Sommer, H. J. & Kalenak, A. (1994). Foot inversion-eversion and knee kinematics during walking. *Journal of Orthopaedic Research*, 12(3), 412–420.
- Laroche, D., Duval, A., Morisset, C., Beis, J.-N., d'Athis, P., Maillefert, J.-F. & Ornetti, P. (2011). Test-retest reliability of 3D kinematic gait variables in hip osteoarthritis patients. *Osteoarthritis and Cartilage*, 19(2), 194–199.
- Lenhoff, M. W., Santner, T. J., Otis, J. C., Peterson, M. G., Williams, B. J. & Backus, S. I. (1999). Bootstrap prediction and confidence bands: a superior statistical method for analysis of gait data. *Gait and Posture*, 9(1), 10–17.
- Long, J. T., Eastwood, D. C., Graf, A. R., Smith, P. A. & Harris, G. F. (2010). Repeatability and sources of variability in multi-center assessment of segmental foot kinematics in normal adults. *Gait and Posture*, 31(1), 32–36.
- Maiwald, C., Axmann, D. & Grau, S. (2011). Measurement error in footwear research biomechanics. *Footwear Science*, 3(2), 117–124.
- McClay, I. & Manal, K. (1998). A comparison of three-dimensional lower extremity kinematics during running between excessive pronators and normals. *Clinical Biomechanics*, 13(3), 195–203.
- Meardon, S. A., Hamill, J. & Derrick, T. R. (2011). Running injury and stride time variability over a prolonged run. *Gait and Posture*, 33(1), 36–40.
- Milani, T. L., Hennig, E. M. & Lafortune, M. A. (1997). Perceptual and biomechanical variables for running in identical shoe constructions with varying midsole hardness. *Clinical Biomechanics*, 12(5), 294–300.
- Milani, T. L., Schnabel, G. & Hennig, E. M. (1995). Rearfoot motion and pressure distribution patterns during running in shoes with varus and valgus wedges. *Journal of Applied Biomechanics*,

- 11(2), 177–187.
- Mündermann, A., Nigg, B. M., Humble, R. N. & Stefanyshyn, D. J. (2003). Foot orthotics affect lower extremity kinematics and kinetics during running. *Clin Biomech (Bristol, Avon)*, 18(3), 254–262.
- Moosbrugger, H. (2007). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer Medizin Verlag.
- Nakayama, Y., Kudo, K. & Ohtsuki, T. (2010). Variability and fluctuation in running gait cycle of trained runners and non-runners. *Gait and Posture*, 31(3), 331–335.
- Newell, K. M. & Corcos, D. M. (Hrsg.). (1993). *Variability and Motor Control*. Champaign, IL: Human Kinetics.
- Newell, K. M. & Slifkin, A. B. (1998). The Nature of Movement Variability. In J. P. Piek (Hrsg.), *Motor Behavior and Human Skill* (S. 143–160). Champaign, IL: Human Kinetics.
- Nigg, B. M. (1986). *Biomechanics of running shoes*. Champaign, IL: Human Kinetics.
- Nigg, B. M. (2001). The role of impact forces and foot pronation: a new paradigm. *Clinical Journal of Sport Medicine*, 11(1), 2–9.
- Nigg, B. M. & Bahlsen, H. A. (1988). Influence of Heel Flare and Midsole Construction on Pronation, Supination, and Impact Forces for Heel-Toe-Running. *International Journal of Sport Biomechanics*, 4, 205–219.
- Nigg, B. M., Bahlsen, H. A., Luethi, S. M. & Stokes, S. (1987). The influence of running velocity and midsole hardness on external impact forces in heel-toe running. *Journal of Biomechanics*, 20(10), 951–959.
- Nigg, B. M. & Morlock, M. (1987). The influence of lateral heel flare of running shoes on pronation and impact forces. *Medicine and Science in Sports and Exercise*, 19(3), 294–302.
- Nigg, B. M., Stergiou, P., Cole, G., Stefanyshyn, D., Mündermann, A. & Humble, N. (2003). Effect of shoe inserts on kinematics, center of pressure, and leg joint moments during running. *Medicine and Science in Sports and Exercise*, 35(2), 314–319.
- O'Connor, C. M., Thorpe, S. K., O'Malley, M. J. & Vaughan, C. L. (2007). Automatic detection of gait events using kinematic data. *Gait and Posture*, 25(3), 469–474.
- Oriwol, D., Sterzing, T. & Milani, T. L. (2011). The position of medial dual density midsole elements in running shoes does not influence biomechanical variables. *Footwear Science*, 3(2), 107–116.
- Perry, S. D. & Lafortune, M. A. (1995). Influences of inversion/eversion of the foot upon impact loading during locomotion. *Clinical Biomechanics*, 10(5), 253–257.
- Pohl, M. B. & Buckley, J. G. (2008). Changes in foot and shank coupling due to alterations in foot strike pattern during running. *Clinical Biomechanics*, 23(3), 334–341.
- Pohl, M. B., Lloyd, C. & Ferber, R. (2010). Can the reliability of three-dimensional running kinematics be improved using functional joint methodology? *Gait and Posture*, 32(4), 559–563.
- Queen, R. M., Gross, M. T. & Liu, H.-Y. (2006). Repeatability of lower extremity kinetics and

- kinematics for standardized and self-selected running speeds. *Gait and Posture*, 23(3), 282–287.
- Ratkowsky, D. A. (1989). *Handbook of Nonlinear Regression Models* (Bd. 107). New York: Marcel Dekker Inc.
- Richards, C., Magin, P. & Callister, R. (2009). Is your prescription of distance running shoes evidence-based? *British Medical Journal*, 43(3), 159–162.
- Schiepek, G. (Hrsg.). (2004). *Neurobiologie der Psychotherapie*. Stuttgart: Schattauer Verlag.
- Schubert, C. (2010). *Subjektive Bewertung von Laufschuheigenschaften im Geschlechtervergleich. Eine bundesweite Fragebogenerhebung an 4500 Laufsportlern*. Unveröffentlichte Magisterarbeit, Technische Universität Chemnitz.
- Schubert, C., Oriwol, D. & Sterzing, T. (2011). Gender and age related requirements of running shoes: a questionnaire on 4501 runners. In *Proceedings of the Tenth Footwear Biomechanics Symposium (Tübingen, Germany)* (Bd. 3, S. S148–S150). Footwear Science.
- Seber, G. A. F. & Wild, C. J. (2003). *Nonlinear Regression*. Hoboken: Wiley-Interscience.
- Sheskin, D. J. (2007). *Handbook of Parametric and Nonparametric Statistical Procedures: Fourth Edition* (4. Aufl.). Boca Raton: Chapman & Hall/CRC.
- Siegler, R., DeLoache, J. & Eisenberg, N. (2008). *Entwicklungspsychologie im Kindes- und Jugendalter*. Heidelberg: Spektrum Akademischer Verlag.
- Speckmann, E.-J., Heschler, J. & Köhling, R. (Hrsg.). (2008). *Physiologie* (5. Aufl.). München: Urban & Fischer.
- Stacoff, A., Reinschmidt, C., Nigg, B. M., van den Bogert, A. J., Lundberg, A., Denoth, J. & Stüssi, E. (2000). Effects of foot orthoses on skeletal motion during running. *Clin Biomech (Bristol, Avon)*, 15(1), 54–64.
- Stacoff, A., Reinschmidt, C., Nigg, B. M., van den Bogert, A. J., Lundberg, A., Denoth, J. & Stüssi, E. (2001). Effects of shoe sole construction on skeletal motion during running. *Medicine and Science in Sports and Exercise*, 33(2), 311–319.
- Stam, K. J., Tavy, D. L., Jelles, B., Achtereekte, H. A., Slaets, J. P. & Keunen, R. W. (1994). Non-linear dynamical analysis of multichannel EEG: clinical applications in dementia and Parkinson's disease. *Brain Topography*, 7(2), 141–150.
- Stergiou, N. (2004). *Innovative Analyses of Human Movement*. Champaign, IL: Human Kinetics.
- Stergiou, N. & Scott, M. M. (2005). Baseline measures are altered in biomechanical studies. *Journal of Biomechanics*, 38(1), 175–178.
- Steyer, R. & Eid, M. (2001). *Messen und Testen: Mit Übungen und Lösungen* (2., korr. Aufl.). Berlin: Springer.
- Tenenbaum, G. & Driscoll, M. (2005). *Methods of Research in Sport Sciences: Quantitative and Qualitative Approaches*. Aachen: Meyer & Meyer Fachverlag und Buchhandel GmbH.
- van Emmerik, R. E. A. & van Wegen, E. E. H. (2000). On variability and stability in human movement. *Journal of Applied Biomechanics*, 16(4), 394–406.

- van Gheluwe, B., Tielemans, R. & Roosen, P. (1995). The influence of heel counter rigidity on rearfoot motion during running. *Journal of Applied Biomechanics*, 11(1), 47–67.
- Venter, A. & Maxwell, S. E. (2000). Issues in the Use and Application of Multiple Regression Analysis. In H. E. Tinsley & S. D. Brown (Hrsg.), *Handbook of Applied Multivariate Statistics and Mathematical Modeling* (S. 151–182). San Diego: Academic Press.
- Wank, V., Frick, U. & Schmidbleicher, D. (1998). Kinematics and electromyography of lower limb muscles in overground and treadmill running. *International Journal of Sports Medicine*, 19(7), 455–461.
- Webster, K. E., McClelland, J. A., Wittwer, J. E., Tecklenburg, K. & Feller, J. A. (2010). Three dimensional motion analysis of within and between day repeatability of tibial rotation during pivoting. *Knee*, 17(5), 329–333.
- Weierstraß, K. (1927). *Mathematische Werke von Karl Weierstrass. Vorlesungen über Variationsrechnung (bearbeitet von Rudolf Rothe)*. Leipzig: Akademische Verlagsgesellschaft M. B. H.
- Weihe, W. (2004). Von der Wahrscheinlichkeit des Irrtums. *Deutsches Ärzteblatt*, 101(13), A 834–838.
- Wellhöfer, P. R. (1984). *Grundstudium Sozialwissenschaftliche Methoden und Arbeitsweisen*. Stuttgart: Ferdinand Enke Verlag.
- White, R., Agouris, I., Selbie, R. D. & Kirkpatrick, M. (1999). The variability of force platform data in normal and cerebral palsy gait. *Clinical Biomechanics*, 14(3), 185–192.
- Winter, D. A. (2005). *Biomechanics and Motor Control of Human Movement* (3. Aufl.). John Wiley & Sons.
- Zimmermann, K. (2010). *Karthago. Aufstieg und Fall einer Großmacht*. Stuttgart: Conrad Theiss Verlag.

Anhang A

Ergänzende Informationen zu bewegungswissenschaftlichen Studien

Überblick über Standardisierungen und verwendete Messgeräte ausgewählter biomechanischer Studien

Autor	n	v [m/s]	Art	FU/ LB	#	Messtechnik (Aufnahmefrequenz)	Filter (Art, nung, cut-off- Frequenz)	Ord- cut-off- Frequenz)	Labor/ Feld
Orivoli et al. (2011)	16	[3.4;3.6]	Lauf	FU	5	Kraftmessplatte (1000 Hz) Beschl.-aufnehmer (1000 Hz) Elektrogoniometer (1000 Hz)	BuWo, 3, 100		Labor
Heidenfelder et al. (2010)	20	[3.4;3.6]	Lauf	FU	5	Kraftmessplatte (1000 Hz) Beschl.-aufnehmer (1000 Hz) Elektrogoniometer (1000 Hz)	BuWo, 3, 100		Labor
Brauner et al. (2009)	10, 11, 17	[3.4;3.6]	Lauf	FU	5	Kraftmessplatte (1000 Hz) Beschl.-aufnehmer (1000 Hz) Elektrogoniometer (1000 Hz) Fragebogen	BuWo, 3, 50		Labor/ Feld
Eslami et al. (2009)	11	170 Schritte / Minute	Lauf	FU	10	Kraftmessplatte (960 Hz) Kinematographie (60 Hz)	BuWo, 4, 8		Labor
Pohl und Buckley (2008)	12	indiv.	Lauf	FU	5	Kraftmessplatte Kinematographie (240 Hz)	BuWo, 4, 12		Labor
Butler et al. (2007)	12	indiv.	Lauf	LB	5	Kinematographie (120 Hz) Beschl.-aufnehmer (1080 Hz)	BuWo, 2, 8		Labor
Butler et al. (2006)	40	[3.4;3.7]	Lauf	FU	5	Kraftmessplatte (1080 Hz) Kinematographie (120 Hz) Beschl.-aufnehmer (1080 Hz)	BuWo, 2, 50 BuWo, 2, 8		Labor

Fortsetzung auf der nächsten Seite

Überblick über Standardisierungen und verwendete Messgeräte ausgewählter biomechanischer Studien								
Autor	n	v [m/s]	Art	FU/ LB	#	Messtechnik (Aufnahmefrequenz)	Filter (Art, Ordnung, cut-off-Frequ.)	Labor/ Feld
Digby et al. (2005)	9	[3.2;3.5]	Lauf	FU	5 – 7	Kraftmessplatte Kinematographie (1000 Hz)	BuWo, 4, 40	Labor
Hardin et al. (2004)	12, 10	3.4	Lauf	LB	10	Kinematographie (200 Hz) Metabolische Daten	BuWo, -, 12	Labor
Mündermann et al. (2003)	21	[3.8;4.2]	Lauf	FU	12	Kraftmessplatte (2400 Hz) Kinematographie (240 Hz)	BuWo, 2, 50 BuWo, 2, 12	Labor
Nigg et al. (2003)	15	[3.8;4.2]	Lauf	FU	5	Kraftmessplatte (1200 Hz) Kinematographie (240 Hz) Druckverteilung	BuWo, 4, 100 BuWo, 4, 12	Labor
Stacoff et al. (2001)	5	[2.5;3.0]	Lauf	FU	3	Kraftmessplatte Kinematographie (200 Hz)		Labor
Stacoff et al. (2000)	5	[2.5;3.0]	Lauf	FU	3	Kraftmessplatte Kinematographie (200 Hz)	BuWo, 4, 10	Labor
McClay und Manal (1998)	18	3.35	Lauf	LB	5	Kinematographie (200 Hz)	BuWo, 2, 8	Labor
Milani et al. (1997)	27	[3.2;3.4]	Lauf	FU	9	Kraftmessplatte (1000 Hz) Druckverteilung Beschl.-aufnehmer Elektrogoniometer Fragebogen	-, -, 10	Labor
Hennig et al. (1996)	14	[3.7;3.9]	Lauf	LB FU	8	Kraftmessplatte (1000 Hz) Fragebogen Druckverteilung		Labor

Fortsetzung auf der nächsten Seite

Überblick über Standardisierungen und verwendete Messgeräte ausgewählter biomechanischer Studien									
Autor	n	v [m/s]	Art	FU/ LB	#	Messtechnik (Aufnahmefrequenz)	Filter (Art, Ordnung, cut-off-Frequ.)	Labor/ Feld	
van Gheluwe et al. (1995)	30	3.8	Lauf	LB	3	Kinematographie (60 Hz)		Labor	
Milani et al. (1995)	20	[3.4;3.6]	Lauf	FU	8	Kraftmessplatte (1000 Hz) Druckverteilung Beschl.-aufnehmer Elektrogoniometer		Labor	
Perry und Lafortune (1995)	10	[3.6;4.0] [1.4;1.6]	Lauf Gang	FU	5 5	Kraftmessplatte (1000 Hz) Beschl.-aufnehmer Kinematographie (60 Hz)	BuWo, 2, 100 BuWo, 2, 100	Labor	
Chen et al. (1994)	14	indiv.	Gang Lauf	LB	-	Druckverteilung mittels Innen- sohle Fragebogen		Labor	
Ferrandis et al. (1994)	10	3.57	Lauf	LB	3	Kinematographie		Labor	
Lafortune et al. (1994)	5	1.2	Gang	FU	2	Kinematographie (100 Hz)		Labor	
Hamill et al. (1992)	12	konstant	Lauf	LB	5	Kinematographie (200 Hz)		Labor	
Nigg und Bahlsen (1988)	14	[3.7;4.3]	Lauf	FU	1	Kraftmessplatte (1020 Hz) Kinematographie (100 Hz)		Labor	

Tabelle A.1:

Überblick über Standardisierungen und verwendete Messgeräte ausgewählter biomechanischer Studien
(n = Anzahl Probanden; v = Geschwindigkeit; FU = fester Untergrund, LB = Laufband; # = Anzahl Läufe; Beschl.-aufnehmer = Beschleunigungsaufnehmer; Frequ. = Frequenz; BuWo = Butterworth-Filter)

Anhang B

Ergänzende Informationen zu den Resultaten von Hypothese I

<i>Visuelle Überprüfung der Konvergenz des Mittelwertes diskreter Parameter</i>								
p	Studie I				Studie II			
	Messtag 1 (#)		Messtag 2 (#)		Messtag 1 (#)		Messtag 2 (#)	
F_{zI}	4	(63 ± 15)	8	(61 ± 23)	11	(76 ± 12)	12	(64 ± 18)
F_{zA}	6	(58 ± 15)	8	(61 ± 16)	6	(68 ± 12)	14	(57 ± 19)
FR_{zI}	6	(68 ± 19)	6	(65 ± 12)	11	(71 ± 14)	13	(65 ± 21)
F_{yB}	7	(71 ± 16)	7	(61 ± 18)	8	(69 ± 8)	14	(51 ± 20)
F_{yA}	8	(66 ± 12)	9	(49 ± 17)	4	(68 ± 15)	11	(58 ± 19)
FR_{yB}	7	(56 ± 21)	8	(73 ± 12)	11	(69 ± 18)	9	(71 ± 14)
FR_{yA}	5	(70 ± 12)	10	(65 ± 15)	10	(63 ± 13)	8	(64 ± 16)
ω_{max}					10	(63 ± 15)	9	(60 ± 19)
φ_{max}					11	(68 ± 11)	9	(58 ± 19)

Tabelle B.1:

Visuelle Überprüfung der Konvergenz der kumulierten Mittelwerte: Anzahl an Probanden mit konvergierenden Mittelwertkurven (mittlere minimale Anzahl ± Standardabweichung)

<i>Mittelwertkriterium für diskrete Parameter</i>						
Studie I	Messtag 1 (#)			Messtag 2 (#)		
p	$q = 0.1$	$q = 0.25$	$q = 0.5$	$q = 0.1$	$q = 0.25$	$q = 0.5$
F_{z_I}	59 ± 7	28 ± 8	6 ± 6	60 ± 19	29 ± 23	17 ± 15
F_{z_A}	58 ± 23	40 ± 15	14 ± 10	50 ± 15	33 ± 18	13 ± 15
FR_{z_I}	54 ± 18	17 ± 14	4 ± 8	50 ± 10	32 ± 18	8 ± 12
F_{y_B}	60 ± 17	37 ± 13	7 ± 7	55 ± 15	25 ± 21	8 ± 10
F_{y_A}	53 ± 15	30 ± 17	11 ± 13	78 ± 5	35 ± 21	14 ± 16
FR_{y_B}	61 ± 13	17 ± 14	7 ± 9	71 ± 11	51 ± 9	23 ± 16
FR_{y_A}	53 ± 20	41 ± 15	11 ± 14	60 ± 10	33 ± 20	16 ± 13
Studie II	Messtag 1 (#)			Messtag 2 (#)		
p	$q = 0.1$	$q = 0.25$	$q = 0.5$	$q = 0.1$	$q = 0.25$	$q = 0.5$
F_{z_I}	67 ± 12	34 ± 21	10 ± 13	49 ± 18	31 ± 17	7 ± 9
F_{z_A}	64 ± 8	35 ± 16	17 ± 14	50 ± 15	31 ± 16	11 ± 12
FR_{z_I}	63 ± 18	40 ± 15	12 ± 15	53 ± 19	34 ± 18	7 ± 10
F_{y_B}	58 ± 8	28 ± 17	11 ± 11	49 ± 14	18 ± 15	6 ± 7
F_{y_A}	48 ± 12	31 ± 15	6 ± 10	51 ± 16	20 ± 19	5 ± 5
FR_{y_B}	59 ± 9	28 ± 19	9 ± 14	57 ± 13	29 ± 16	12 ± 11
FR_{y_A}	52 ± 15	24 ± 12	9 ± 7	52 ± 18	23 ± 16	4 ± 6
ω_{max}	50 ± 13	25 ± 12	6 ± 6	55 ± 10	23 ± 16	10 ± 9
φ_{max}	58 ± 8	28 ± 14	6 ± 9	47 ± 20	17 ± 12	6 ± 7

Tabelle B.2:

Mit Mittelwertkriterium bestimmte minimale Anzahl für $q = 0.1, 0.25, 0.5$: mittlere Anzahl über Probanden mit visueller Konvergenz ± deren Standardabweichung

<i>Abhängigkeit des Mittelwertkriteriums von der Gesamtanzahl (diskrete Parameter)</i>				
p	Studie I		Studie II	
	Messtag 1 (%)	Messtag 2 (%)	Messtag 1 (%)	Messtag 2 (%)
F_{zI}	64 ± 5	65 ± 3	66 ± 5	62 ± 3
F_{zA}	68 ± 4	66 ± 4	70 ± 5	64 ± 3
FR_{zI}	65 ± 4	69 ± 4	67 ± 5	63 ± 4
F_{yB}	66 ± 5	63 ± 2	64 ± 5	60 ± 3
F_{yA}	66 ± 3	61 ± 3	64 ± 4	59 ± 2
FR_{yB}	64 ± 4	66 ± 5	65 ± 4	65 ± 4
FR_{yA}	68 ± 5	64 ± 4	63 ± 3	64 ± 2
ω_{max}			63 ± 3	59 ± 4
φ_{max}			66 ± 4	63 ± 3

Tabelle B.3:

Abhängigkeit des Mittelwertkriteriums von der Gesamtanzahl für $q = 0.1$: mittlerer prozentualer Anteil an der Gesamtanzahl ± Standardabweichung

<i>Differenz innerhalb eines Messtages für Probanden mit visueller Konvergenz (diskrete Parameter)</i>					
p	Studie I		Studie II		
	Messtag 1 (%)	Messtag 2 (%)	Messtag 1 (%)	Messtag 2 (%)	
F_{zI}	11 (5 ± 4)	10 (5 ± 4)	8 (4 ± 3)	8 (3 ± 2)	
F_{zA}	5 (2 ± 2)	3 (2 ± 1)	2 (1 ± 1)	3 (1 ± 1)	
FR_{zI}	9 (5 ± 4)	9 (4 ± 2)	12 (6 ± 5)	25 (6 ± 6)	
F_{yB}	-11 (-6 ± 4)	-12 (-3 ± 4)	-11 (-6 ± 4)	-21 (-4 ± 6)	
F_{yA}	10 (4 ± 3)	14 (6 ± 4)	8 (4 ± 2)	11 (4 ± 3)	
FR_{yB}	9 (6 ± 3)	32 (14 ± 10)	10 (5 ± 3)	17 (9 ± 5)	
FR_{yA}	16 (5 ± 6)	19 (7 ± 6)	28 (8 ± 8)	16 (4 ± 4)	
ω_{max}			18 (5 ± 5)	26 (12 ± 8)	
φ_{max}			18 (6 ± 5)	15 (9 ± 4)	

Tabelle B.4:

Differenz D_{intra}^{pr} für Probanden mit visueller Konvergenz: maximale prozentuale Abweichung innerhalb aller Probanden (mittlere Differenz ± mittlere Standardabweichung der Differenz über alle Probanden)

Absolute Abweichungen innerhalb und zwischen den Messtagen aller Probanden (diskrete Parameter)

Studie I						
p	Messtag 1		Messtag 2		Messtag 1/2	
F_{z_I}	0.22	(0.08 ± 0.06)	0.25	(0.10 ± 0.07)	0.24	(0.06 ± 0.07)
F_{z_A}	0.13	(0.06 ± 0.04)	0.14	(0.04 ± 0.04)	0.11	(0.04 ± 0.03)
FR_{z_I}	29.00	(6.44 ± 7.22)	32.02	(9.80 ± 10.95)	14.51	(4.36 ± 4.74)
F_{y_B}	0.058	(0.025 ± 0.017)	0.041	(0.018 ± 0.014)	0.032	(0.012 ± 0.011)
F_{y_A}	0.056	(0.019 ± 0.014)	0.048	(0.015 ± 0.012)	0.030	(0.010 ± 0.008)
FR_{y_B}	19.37	(3.16 ± 4.96)	7.22	(2.58 ± 2.30)	12.86	(3.08 ± 3.20)
FR_{y_A}	3.86	(0.90 ± 1.10)	3.61	(0.99 ± 1.02)	2.23	(0.79 ± 0.71)
Studie II						
p	Messtag 1		Messtag 2		Messtag 1/2	
F_{z_I}	0.14	(0.06 ± 0.04)	0.15	(0.06 ± 0.04)	0.17	(0.06 ± 0.04)
F_{z_A}	0.27	(0.06 ± 0.06)	0.07	(0.03 ± 0.02)	0.08	(0.03 ± 0.02)
FR_{z_I}	14.99	(6.02 ± 4.74)	32.27	(5.61 ± 6.89)	13.26	(5.59 ± 3.82)
F_{y_B}	0.043	(0.019 ± 0.013)	0.001	(0.058 ± 0.017)	0.048	(0.016 ± 0.013)
F_{y_A}	0.043	(0.015 ± 0.013)	0.033	(0.012 ± 0.008)	0.026	(0.009 ± 0.007)
FR_{y_B}	11.70	(1.71 ± 2.53)	3.55	(1.68 ± 0.95)	3.08	(1.13 ± 0.90)
FR_{y_A}	4.00	(0.78 ± 1.01)	1.45	(0.48 ± 0.44)	1.22	(0.40 ± 0.33)
ω_{max}	62.23	(22.14 ± 19.70)	110.00	(33.82 ± 27.84)	53.57	(18.85 ± 13.35)
φ_{max}	2.81	(1.03 ± 0.76)	2.78	(1.29 ± 0.66)	4.62	(1.15 ± 1.28)

Tabelle B.5:

Differenz von MW_{100}^{pr} und MW_5^{pr} : maximale Abweichung innerhalb aller Probanden (mittlere Differenz ± mittlere Standardabweichung der Differenz über alle Probanden); (F_{z_I} [bw]; F_{z_A} [bw]; FR_{z_I} [bw/s]; F_{y_B} [bw]; F_{y_A} [bw]; FR_{y_B} [bw/s]; FR_{y_A} [bw/s]; ω_{max} [°/s]; φ_{max} [°])

<i>Visuelle Überprüfung der Konvergenz punktweise kumulierter Mittelwertkurven</i>								
k	Studie I				Studie II			
	Messtag 1 (#)		Messtag 2 (#)		Messtag 1 (#)		Messtag 2 (#)	
F_z	11	66 ± 9	9	57 ± 10	10	73 ± 10	17	71 ± 9
F_y	9	69 ± 13	9	69 ± 13	13	71 ± 7	14	68 ± 10
F_x	9	68 ± 11	11	73 ± 6	15	69 ± 8	16	65 ± 8
ω					7	74 ± 12	10	79 ± 7

Tabelle B.6:

Visuelle Überprüfung der Konvergenz der punktweise kumulierten Mittelwertkurven: Anzahl konvergierender Mittelwertkurven (mittlere minimale Anzahl \pm Standardabweichung)

<i>Mittelwertkriterium für gesamte Kurven</i>							
k	Studie I			Studie II			
	Messtag 1 (#)			Messtag 2 (#)			
	$q = 0.1$	$q = 0.25$	$q = 0.5$	$q = 0.1$	$q = 0.25$	$q = 0.5$	
F_z	83 ± 4	66 ± 6	47 ± 8	84 ± 3	67 ± 7	44 ± 9	
F_y	83 ± 3	66 ± 7	44 ± 8	84 ± 3	66 ± 6	46 ± 7	
F_x	84 ± 4	62 ± 7	39 ± 11	84 ± 3	66 ± 5	40 ± 9	
	Studie I			Studie II			
	Messtag 1 (#)			Messtag 2 (#)			
k	$q = 0.1$	$q = 0.25$	$q = 0.5$	$q = 0.1$	$q = 0.25$	$q = 0.5$	
F_z	82 ± 4	63 ± 5	37 ± 7	82 ± 5	63 ± 6	38 ± 10	
F_y	82 ± 3	65 ± 6	41 ± 10	83 ± 3	62 ± 9	37 ± 12	
F_x	83 ± 2	62 ± 8	37 ± 10	82 ± 5	56 ± 7	30 ± 11	
ω	85 ± 2	63 ± 8	39 ± 8	85 ± 2	68 ± 4	44 ± 7	

Tabelle B.7:

Mit Mittelwertkriterium bestimmte minimale Anzahl für $q = 0.1, 0.25, 0.5$: mittlere Anzahl über alle Probanden mit visueller Konvergenz \pm Standardabweichung

Abhängigkeit des Mittelwertkriteriums von der Gesamtanzahl (gesamte Kurve)

k	Studie I		Studie II	
	Messtag 1 (%)	Messtag 2 (%)	Messtag 1 (%)	Messtag 2 (%)
F_z	66 ± 3	64 ± 4	64 ± 3	61 ± 3
F_y	84 ± 6	84 ± 7	83 ± 7	80 ± 7
F_x	65 ± 2	63 ± 3	62 ± 3	59 ± 1
ω			67 ± 2	66 ± 3

Tabelle B.8:

Abhängigkeit des Mittelwertkriteriums von der Gesamtanzahl für $q = 0.25$: mittlerer prozentualer Anteil an der Gesamtanzahl ± Standardabweichung

Differenzen innerhalb eines Messtages für Probanden mit visueller Konvergenz (gesamte Kurven)

k	Studie I		Studie II	
	Messung 1 (%)	Messung 2 (%)	Messung 1 (%)	Messung 2 (%)
F_z	8 (4 ± 2)	7 (4 ± 1)	7 (3 ± 1)	4 (3 ± 1)
F_y	24 (9 ± 3)	23 (12 ± 6)	19 (11 ± 5)	25 (8 ± 5)
F_x	103 (43 ± 31)	38 (25 ± 11)	82 (34 ± 17)	85 (31 ± 19)
ω			34 (23 ± 9)	34 (21 ± 9)

Tabelle B.9:

Differenz DA_{intra}^{kr} für Probanden mit visueller Konvergenz: maximale prozentuale Abweichung (mittlere Differenz ± mittlere Standardabweichung der Differenz)

Differenz zwischen den Messtagen für Probanden mit zwei konvergenten Mittelwertkurven

k	Studie I (%)	Studie II (%)
F_z	5 (3 ± 2)	5 (3 ± 2)
F_y	11 (7 ± 3)	18 (9 ± 4)
F_x	54 (28 ± 16)	35 (19 ± 6)
ω		27 (18 ± 7)

Tabelle B.10:

Differenz DA_{inter}^{kr} zwischen zwei Messtagen der Probanden mit konvergenten Mittelwertkurven: maximale prozentuale Abweichung (mittlere prozentuale Abweichung ± Standardabweichung)

Anhang C

Ergänzende Informationen zu den Resultaten von Hypothese II

<i>Visuelle Überprüfung der Konvergenz der Standardabweichungen</i>								
p	Studie I				Studie II			
	Messtag 1 (#)		Messtag 2 (#)		Messtag 1 (#)		Messtag 2 (#)	
F_{z_I}	10	(59 ± 24)	10	(57 ± 14)	16	(53 ± 20)	14	(62 ± 11)
F_{z_A}	10	(50 ± 12)	10	(46 ± 24)	7	(54 ± 17)	14	(52 ± 18)
FR_{z_I}	8	(69 ± 14)	10	(65 ± 11)	14	(61 ± 15)	14	(52 ± 17)
F_{y_B}	9	(39 ± 23)	13	(58 ± 22)	13	(62 ± 20)	13	(59 ± 18)
F_{y_A}	10	(47 ± 19)	8	(63 ± 5)	13	(62 ± 16)	17	(59 ± 26)
FR_{y_B}	7	(61 ± 15)	7	(41 ± 15)	9	(52 ± 19)	13	(54 ± 23)
FR_{y_A}	6	(53 ± 14)	10	(59 ± 19)	10	(49 ± 22)	14	(54 ± 23)
ω_{max}					12	(44 ± 19)	10	(58 ± 17)
φ_{max}					15	(44 ± 23)	13	(58 ± 27)

Tabelle C.1:

Visuelle Überprüfung der Konvergenz der kumulierten Standardabweichungen: Anzahl an Probanden mit konvergierenden Standardabweichungen (mittlere minimale Anzahl ± Standardabweichung)

Visuelle Überprüfung der Konvergenz der RMSE-Folgen

Studie I									
p		Messtag 1			Messtag 2				
		$RMSE_{100}^p$	Min	Max	$RMSE_{100}^p$	Min	Max		
F_{z_I}	[bw]	0.37	0.34	0.37	–	0.38	0.37	0.41	+
F_{z_A}	[bw]	0.19	0.18	0.20	+	0.17	0.17	0.19	–
FR_{z_I}	[bw/s]	32.55	27.77	35.71	–	35.19	33.14	37.52	+
F_{y_B}	[bw]	0.09	0.08	0.09	+	0.09	0.09	0.10	+
F_{y_A}	[bw]	0.06	0.05	0.06	+	0.06	0.06	0.08	+
FR_{y_B}	[bw/s]	12.43	8.49	12.50	+	9.94	9.53	10.23	–
FR_{y_A}	[bw/s]	6.88	3.88	7.08	–	4.76	4.76	5.90	–

Studie II									
p		Messtag 1			Messtag 2				
		$RMSE_{100}^p$	Min	Max	$RMSE_{100}^p$	Min	Max		
F_{z_I}	[bw]	0.35	0.35	0.4	–	0.33	0.33	0.43	–
F_{z_A}	[bw]	0.20	0.18	0.2	+	0.16	0.15	0.17	–
FR_{z_I}	[bw/s]	29.03	26.5	30.84	–	29.31	28.85	32.48	+
F_{y_B}	[bw]	0.09	0.08	0.09	+	0.08	0.08	0.1	–
F_{y_A}	[bw]	0.06	0.06	0.07	–	0.05	0.05	0.06	+
FR_{y_B}	[bw/s]	9.52	7.04	9.52	+	8.33	7.96	9.26	+
FR_{y_A}	[bw/s]	3.98	3.31	4.16	–	3.47	3.32	4.23	–
ω_{max}	[°/s]	139.21	130.83	144.58	–	132.17	130.93	149.39	–
φ_{max}	[°]	4.07	3.68	4.29	–	3.93	3.25	4.18	+

Tabelle C.2:

Betrachtung des kumulierten RMSE: $RMSE_{100}^p$ / minimale Ausprägung / maximale Ausprägung; Angabe der Konvergenz (+ konvergierende RMSE-Folge, – divergierende RMSE-Folge)

<i>Prozentuale Differenz des RMSE innerhalb und zwischen den Messtagen</i>						
p	Studie I			Studie II		
	DE_{intra}^p		DE_{inter}^p	DE_{intra}^p		DE_{inter}^p
	M 1 (%)	M 2 (%)	(%)	M 1 (%)	M 2 (%)	(%)
F_{z_I}	6	5	3	11	26	7
F_{z_A}	4	10	8	11	5	20
FR_{z_I}	9	1	8	4	4	1
F_{y_B}	7	11	2	4	30	11
F_{y_A}	20	27	3	6	12	16
FR_{y_B}	29	1	22	20	10	13
FR_{y_A}	51	19	37	12	20	14
ω_{max}				3	7	5
φ_{max}				4	19	3

Tabelle C.3:

Betrachtung der prozentualen Differenz des RMSE DE_{intra}^p innerhalb eines Messtages und DE_{inter}^p zwischen den Messtagen

<i>Aufteilung der divergenten und konvergenten Mittelwertkurven sowie Standardabweichungskurven in einer Vierfeldertafel</i>								
	Studie I		M 2 (%)		Studie II		M 2 (%)	
	M 1 (%)	M 2 (%)	M 1 (%)	M 2 (%)	M 1 (%)	M 2 (%)	M 1 (%)	M 2 (%)
	Standardabweichungen				Standardabweichungen			
	(+)	(-)	(+)	(-)	(+)	(-)	(+)	(-)
MW (+)	27	17	32	23	29	16	38	17
MW (-)	35	21	38	7	32	23	30	15

Tabelle C.4:

Prozentualer Anteil konvergenter und divergenter Standardabweichungskurven entsprechend ihrer zugehörigen konvergenten oder divergenten Mittelwerte ((+) - Konvergenz; (-) - Divergenz)

Abweichung der sich nicht überlagernden Flächen des Prädiktionsbandes für 10 und 100 Läufe

	Studie I		Studie II	
	$DAPB_{10\ 100\ M1}^{kr}$ (%)	$DAPB_{10\ 100\ M2}^{kr}$ (%)	$DAPB_{10\ 100\ M1}^{kr}$ (%)	$DAPB_{10\ 100\ M2}^{kr}$ (%)
F_z	59 ± 19	67 ± 28	54 ± 22	58 ± 17
F_y	49 ± 13	61 ± 18	62 ± 27	67 ± 17
F_x	54 ± 15	58 ± 22	55 ± 17	58 ± 20
ω			69 ± 19	76 ± 17

Tabelle C.5:

Prozentuale Abweichung der Flächen des Prädiktionsbandes für 10 Läufe und für 100 Läufe: mittlere prozentuale Abweichung ± Standardabweichung

Abweichung der sich nicht überlagernden Flächen des Prädiktionsbandes für 25 und 100 Läufe

	Studie I		Studie II	
	$DAPB_{25\ 100\ M1}^{kr}$ (%)	$DAPB_{25\ 100\ M2}^{kr}$ (%)	$DAPB_{25\ 100\ M1}^{kr}$ (%)	$DAPB_{25\ 100\ M2}^{kr}$ (%)
F_z	24 ± 7	27 ± 11	18 ± 5	18 ± 5
F_y	22 ± 6	26 ± 10	19 ± 6	18 ± 7
F_x	19 ± 6	21 ± 9	18 ± 6	16 ± 5
ω			18 ± 4	21 ± 7

Tabelle C.6:

Prozentuale Abweichung der Flächen des Prädiktionsbandes für 25 Läufe und für 100 Läufe: mittlere prozentuale Abweichung ± Standardabweichung

Anhang D

Ergänzende Informationen zu den Resultaten von Hypothese III

Differenz der minimalen Anzahl zwischen den Messtagen für $DAPB_N^{kr}$ und verschiedener vorgegebener Schwellen

Studie I					
k	$s = 2$	$s = 3$	$s = 4$	$s = 5$	visuell
F_z	15 ± 14	27 ± 21	31 ± 21	28 ± 21	30 ± 23
F_y	16 ± 11	24 ± 15	25 ± 16	33 ± 21	20 ± 25
F_x	17 ± 13	23 ± 16	23 ± 16	28 ± 14	25 ± 15

Studie II					
k	$s = 2$	$s = 3$	$s = 4$	$s = 5$	visuell
F_z	20 ± 13	27 ± 18	28 ± 15	32 ± 20	25 ± 18
F_y	23 ± 14	29 ± 17	29 ± 16	20 ± 19	30 ± 20
F_x	16 ± 13	25 ± 19	25 ± 20	24 ± 20	26 ± 21
ω	19 ± 12	24 ± 15	22 ± 13	12 ± 12	25 ± 16

Tabelle D.1:

Mittlere Differenz zwischen den Messtagen der ermittelten minimalen Anzahl anhand der Differenz der Flächen der Prädiktionsbänder $DAPB_N^{kr}$ und verschiedener vorgegebener Schwellen von 2%, 3%, 4% 5% und visuell (Mittelwert \pm Standardabweichung)

Differenz der minimalen Anzahl zwischen den Messtagen anhand verschiedener Populationsdifferenzenkurven

k	Studie I		Studie II	
	$PDSD_{max I}^k$	$PDSD_{mw I}^k$	$PDSD_{max II}^k$	$PDSD_{mw II}^k$
F_z	15 ± 15	18 ± 14	17 ± 19	24 ± 20
F_y	13 ± 11	18 ± 11	10 ± 11	16 ± 15
F_x	15 ± 9	17 ± 11	14 ± 18	24 ± 21
ω			15 ± 20	25 ± 19

Tabelle D.2:

Mittlere Differenz der minimalen Anzahl ermittelt mit $PDSD_{max}^k$ und $PDSD_{mw}^k$. (Mittelwert ± Standardabweichung)

Anzahl an Probanden mit annähernd gleicher Differenz am Schwellenwert $s = 5$

k	Studie I		Studie II	
	Messtag 1 (#)	Messtag 2 (#)	Messtag 1 (#)	Messtag 2 (#)
F_z	11 (11)	9 (10)	15 (13)	10 (7)
F_y	4 (4)	7 (7)	13 (9)	15 (14)
F_x	7 (8)	7 (8)	6 (9)	8 (7)
ω			11	10

Mittlere Differenzen am Schwellenwert $s = 5$ für alle Probanden mit $PDSD_{mw III}^k$

k	Studie I		Studie II	
	Messtag 1 (#)	Messtag 2 (#)	Messtag 1 (#)	Messtag 2 (#)
F_z	-8 ± 14	- 8 ± 24	- 5 ± 21	-17 ± 19
F_y	-7 ± 35	3 ± 17	- 6 ± 20	- 1 ± 16
F_x	-7 ± 14	-14 ± 24	-14 ± 30	-22 ± 25

Tabelle D.3:

Oben: Anzahl ermittelt für originäre PDSD-Kurve (und für $PDSD_{mw III}^k$); Unten: Mittlere Differenz $\#PDSD - \#DAPB$ mittels $PDSD_{mw III}^k$ (Mittelwert ± Standardabweichung)

Mittlere Standardabweichung über die entsprechende ermittelte Anzahl der drei Bodenreaktionskräfte innerhalb eines Probanden

	Messtag 1		Messtag 2	
	$PDSD_{max}^k$	$PDSD_{mw}^k$	$PDSD_{max}^k$	$PDSD_{mw}^k$
Studie I	13 [0;26]	14 [0;30]	8 [2;31]	10 [1;31]
Studie II	11 [0;30]	18 [2;42]	8 [1;34]	12 [3;33]
Studie I & II	11 [0;32]	16 [0;42]	6 [2;23]	9 [1;20]

Tabelle D.4:

Standardabweichung der mittleren Anzahl der drei Bodenreaktionskräfte innerhalb eines Probanden: mittlere Standardabweichung [Minimum ; Maximum]

Anhang E

Ergänzende Informationen zu den Inhalten der Diskussion

Bestimmung des Stichprobenmittelwertes und der Stichprobenvarianz unter Annahme gleicher Fehler für alle Probanden (Offset)

$$\overline{MW}_1 = \frac{1}{n_1} \sum_{r=1}^{n_1} (\overline{\mu}_r) = \frac{1}{n_1} \sum_{r=1}^{n_1} (\mu_r + E(\varepsilon_r)) = MW_1 + E(\varepsilon_1) \quad (\text{E.1})$$

$$\overline{MW}_2 = MW_2 + E(\varepsilon_2) \quad (\text{E.2})$$

Unter der Annahme, dass die eingegangenen Fehler für alle Probanden gleich sind, entsprechen sich sowohl die einzelnen mittleren Fehler $E(\varepsilon_r)$ als auch die Fehler, welche bei der Bildung des arithmetischen Mittels für die beiden Gruppen $E(\varepsilon_1)$ und $E(\varepsilon_2)$ entstehen. Somit ergibt sich für die Varianz

$$\begin{aligned} \overline{SD}_1^2 &= \frac{1}{n_1 - 1} \sum_{r=1}^{n_1} ((\mu_r + E(\varepsilon_r)) - (MW_1 + E(\varepsilon_1)))^2 \\ &= \frac{1}{n_1 - 1} \sum_{r=1}^{n_1} (\mu_r - MW_1)^2 \\ &= SD_1^2 \end{aligned} \quad (\text{E.3})$$

$$\overline{SD}_2^2 = SD_2^2 \quad (\text{E.4})$$

Bestimmung des Stichprobenmittelwertes und der Stichprobenvarianz unter Annahme differenzierter Fehler für alle Probanden

Die Gruppenmittelwerte entsprechen denen unter Annahme gleicher Fehler pro Proband (\nearrow Glg. E.1 und E.2) , wobei $E(\varepsilon_1) = \frac{1}{n_1} \sum_{r=1}^{n_1} (E(\varepsilon_r))$ und $E(\varepsilon_2) = \frac{1}{n_1} \sum_{i=1}^{n_2} (E(\varepsilon_i))$. Die Varianzen berechnen sich wie folgt:

$$\begin{aligned} \overline{SD}_1^2 &= \frac{1}{n_1 - 1} \sum_{r=1}^{n_1} (\overline{\mu}_r - \overline{MW}_1)^2 \\ &= \frac{1}{n_1 - 1} \sum_{r=1}^{n_1} ((\mu_r + E(\varepsilon_r)) - (MW_1 + E(\varepsilon_1)))^2 \\ &= \frac{1}{n_1 - 1} \sum_{r=1}^{n_1} ((\mu_r - MW_1) + (E(\varepsilon_r) - E(\varepsilon_1)))^2 \end{aligned} \tag{E.5}$$

$$\begin{aligned} &= \underbrace{\frac{1}{n_1 - 1} \sum_{r=1}^{n_1} (\mu_r - MW_1)^2}_{SD_1^2} + \underbrace{\frac{2}{n_1 - 1} \sum_{r=1}^{n_1} (\mu_r - MW_1) (E(\varepsilon_r) - E(\varepsilon_1))}_{\neq 0} \\ &\quad + \underbrace{\frac{1}{n_1 - 1} \sum_{r=1}^{n_1} (E(\varepsilon_r) - E(\varepsilon_1))^2}_{\text{Fehlervarianz} \neq 0} \end{aligned} \tag{E.6}$$

$$= SD_1^2 + SD(\varepsilon_1)^2 \tag{E.7}$$

$$\overline{SD}_2^2 = SD_2^2 + SD(\varepsilon_2)^2 \tag{E.8}$$

Gleichung E.5 entspricht der ersten binomischen Gleichung, diese ausmultipliziert ergibt Gleichung E.6. Die Varianzen der Fehler $SD(\varepsilon_1)^2$ und $SD(\varepsilon_2)^2$ können durchaus negativ werden.

Dies ist abhängig von der Multiplikation von $(\mu_r - MW_1)(E(\varepsilon_r) - E(\varepsilon_1))$, da der mittlere Fehler innerhalb der Gruppe negativ sein kann und die absoluten Fehler nicht kleiner sein können als der tatsächliche Mittelwert. In der Kombination mit einer fast vollständigen Unterschätzung und starken Streuung der tatsächlichen Mittelwerte kann sich ein negativer Wert herausbilden.

Mögliche Ausgänge eines statistischen Tests mit Angabe der entsprechenden Fehler

	H_0 wird angenommen	H_0 wird abgelehnt
H_0 ist wahr	richtig ($1 - \alpha$: Spezifität)	Fehler erster Art (α -Fehler)
H_0 ist unwahr	Fehler zweiter Art (β -Fehler)	richtig ($1 - \beta$: Power)

Tabelle E.1:Fehlerarten (α - und β -Fehler; H_0 - Nullhypothese)

Wissenschaftlicher Werdegang

	Seit 05/2008	Wissenschaftliche Mitarbeiterin am Institut für Sportwissenschaft der Technischen Universität Chemnitz
10/2000	– 04/2008	Diplomstudium der Wirtschaftsmathematik an der Technischen Universität Chemnitz mit Vertiefungsrichtung Stochastik und Investmentbanking/Commercial Banking Diplomarbeit zum Thema „Nichtlineare Regressionsmodelle und statistische Analysen zur Bestimmung biomechanischer Parameter der Rückfußbewegung“
01/2007	– 04/2008	Studentische Hilfskraft am sportwissenschaftlichen Institut
10/2001	– 12/2006	Bakkalaureusstudiengang Finanzmathematik an der Technischen Universität Chemnitz Bakkalaureusarbeit zum Thema „Statistische Methoden zur Analyse von Fragebögen mittels SPSS“
01/2004	– 07/2006	Leitung der Übung „Statistisches Methodenpraktikum (SPSS)“ (mit Unterbrechung)
10/2001	– 07/2005	Studentische Hilfskraft am Lehrstuhl Mathematik (mit Unterbrechung)

Publikationen

- Brauner, T., Milani, T. L., Sterzing, T. & Oriwol, D. (2008a). Determination of Pronation Parameters by Midsole Deformation is Independent of Running Velocity. In *4th North American Congress on Biomechanics*. Ann Arbor, MI. (URL: <http://www.asbweb.org/conferences/2008/abstracts/22.pdf>)
- Brauner, T., Milani, T. L., Sterzing, T. & Oriwol, D. (2008b). Usage of Discrete Pressure Sensors to Determine Rearfoot Motion During Heel-to-Toe-Running. In *Proceedings of the ESM* (S. 67). Dundee, Scotland. (URL: <http://novel.de/pdf/abstract/Proceedings-ESM2008.pdf>)
- Gras, N., Sterzing, T., Brauner, T., Oriwol, D., Heidenfelder, J. & Milani, T. L. (2009). Variability of stride frequency and pronation velocity during a 16 day relay-run around germany - a case study. In *27. International Society of Biomechanics in Sports Conference*. Limerick, Ireland. (ISBN: 978-1-907179-65-5)
- Milani, T. L., Brauner, T., Sterzing, T. & Oriwol, D. (2008). Does Midsole Deformation Reflect Rearfoot Motion During Running? - A Multiple Regression Approach to Evaluate Pronation by Hall Sensors. In *4th North American Congress on Biomechanics*. Ann Arbor, MI. (URL: <http://www.asbweb.org/conferences/2008/abstracts/167.pdf>)
- Oriwol, D. (2006). *Statistische Methoden zur Analyse von Fragebögen mittels SPSS*. Unveröffentlichte Bakkalaureusarbeit, Technische Universität Chemnitz.
- Oriwol, D. (2008). *Nichtlineare Regressionsmodelle und statistische Analysen zur Bestimmung biomechanischer Parameter der Rückfußbewegung*. Unveröffentlichte Diplomarbeit, Technische Universität Chemnitz.
- Oriwol, D. & Maiwald, C. (2011). Variability analysis of laboratory running. In *Proceedings of the Tenth Footwear Biomechanics Symposium (Tübingen, Germany)* (Bd. 3, S. S125–S127). Footwear Science.
- Oriwol, D., Sterzing, T., Maiwald, C., Brauner, T., Heidenfelder, J. & Milani, T. L. (2009). Pronation Velocity Values of Running Shoes are Dependent on the Mathematical Routines Applied During Data Post Processing. In *XXII Congress of the International Society of Biomechanics*. Cape Town, South Africa. (ISBN: 978-0-620-44037-0)
- Oriwol, D., Sterzing, T. & Milani, T. L. (2011). The position of medial dual density midsole elements in running shoes does not influence biomechanical variables. *Footwear Science*, 3(2), 107–116.
- Schlee, G., Milani, T. L., Sterzing, T. & Oriwol, D. (2009). Short-time lower leg ischemia reduces plantar foot sensitivity. *Neuroscience Letters*, 462(3), 286–288.
- Schubert, C., Oriwol, D. & Sterzing, T. (2011). Gender and age related requirements of running shoes: a questionnaire on 4501 runners. In *Proceedings of the Tenth Footwear Biomechanics Symposium (Tübingen, Germany)* (Bd. 3, S. S148–S150). Footwear Science.

Danksagung

Schlussendlich nach Fertigstellung dieser Arbeit ist es mir ein Bedürfnis die Personen zu würdigen, ohne die diese Arbeit nie entstanden wäre.

„Dankbarkeit ist das Eingeständnis unserer Abhängigkeit von anderen.“

- Erhard Blanck -

In erster Linie möchte ich mich bei Thomas und Chris für ihre Unterstützung und ihre Bereitschaft bedanken, mich in meinen Fragen und Ideen zu bestärken. Ihr wachsendes Interesse an methodischen Vorgehensweisen und Umsetzungen von Studien hat diese Untersuchung erst ermöglicht, wobei die Akzeptanz der Infragestellung bisher verwendeter Methoden und damit verbundener möglicher wissenschaftlicher Umstürze keine Selbstverständlichkeit darstellen.

Allen Kollegen, die durch fachliche Gespräche, technologische Umsetzung, Datenaufnahme und moralischen Beistand mich in meinem Vorhaben unterstützt und generell bestärkt haben zu promovieren, sei an dieser Stelle gesondert gedankt. Insbesondere gilt hierbei mein Dank Anna und Stefanie, welche für die Datenaufnahme der zweiten Studie verantwortlich waren und somit neben den Probanden maßgeblich zum Gelingen dieser Arbeit beigetragen haben.

Katha und Susa - eure psychologischen und mathematischen Kompetenzen haben zu vielen Fragen in dieser Arbeit Lösungen geliefert, mathematische Modelle verbessert und Anregungen gegeben - dafür vielen Dank! Liebe Tante Kerstin, dies ist die Gelegenheit dir für das Korrekturlesen meiner drei großen wissenschaftlichen Arbeiten zu danken. Danke für die stoische Gelassenheit, mit der du häufig in der letzten Minute alle meine falschen Kommas an die richtigen Stellen geschaufelt hast, mit der du all meine Vorhaben unterstützt hast und immer ein offenes Ohr für mich hattest.

Zu guter Letzt möchte ich mich bei meinen Eltern, meiner Familie und meinen Freunden bedanken, die über die Jahre der Entstehung dieser Arbeit und vor allem in den letzten Monaten immer mit Rat und Motivation zur Seite gestanden, all meine Launen ertragen haben und mich immer wieder meine Stärken erkennen ließen.