

Bewertung der Treffsicherheit von Schwingfestigkeitsschätzungen mittels Horizont- und Treppenstufenverfahren

Franz Ellmer

2019

Die folgende Dissertationsschrift wurde am 23. August 2018 an der Fakultät Maschinenwesen der Technischen Universität Dresden eingereicht und am 5. März 2019 erfolgreich verteidigt.

Gutachter: 1. Gutachter: Prof. Dr.-Ing. habil. Klaus-Georg Eulitz

2. Gutachter: Prof. Dr.-Ing. habil. Volker Ulbricht

© 2019 Franz Ellmer
Berggießhübler Str. 7
D-01277 Dresden

Alle Rechte vorbehalten. All rights reserved.

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter in der Arbeitsgruppe Betriebsfestigkeit am Institut für Festkörpermechanik der Technischen Universität Dresden. Mein Dank gilt Herrn Prof. Dr.-Ing. habil. Klaus-Georg Eulitz für die fachliche Betreuung, das entgegengebrachte Vertrauen und die aufgebrauchte Geduld, ohne die diese Arbeit nicht zustande gekommen wäre.

Herrn Prof. Dr.-Ing. habil. Volker Ulbricht danke ich für die Übernahme des Zweitgutachtens und die damit verbundenen Mühen.

Des weiteren gilt mein Dank allen Kollegen der Arbeitsgruppe für die angenehme und produktive Arbeitsatmosphäre und insbesondere Herrn Dr. Hantschke für sein ständig offenes Ohr in fachlichen Dingen. Frau Dr.-Ing. Katrin Fuhrmann und Herr Prof. Dr.-Ing. Sebastian Götz waren mir darüber hinaus immer auch moralische Unterstützung.

Franz Ellmer

Inhaltsverzeichnis

Verzeichnis der Formelzeichen und Symbole	11
1 Einleitung	15
1.1 Motivation	15
1.2 Inhalt und Aufbau der Arbeit	18
2 Grundlagen	19
2.1 Grundlagen der statistischen Versuchsauswertung	19
2.1.1 Grundlegende Begriffe	19
2.1.2 Beschreibung einer Stichprobe	20
2.1.3 Schätzen von Verteilungsparametern	22
2.1.3.1 Momentenmethode	23
2.1.3.2 Maximum-Likelihood-Methode	24
2.1.3.3 Grafische und quasigrafische Schätzung	24
2.1.4 Konfidenzintervalle für Verteilungsparameter	26
2.1.5 Anpassungs- und Ausreißertests	27
2.1.6 Einige Verteilungsmodelle und Parameterschätzungen	28
2.1.6.1 Binomial-Verteilung $\mathcal{Bi}(n; p)$	28
2.1.6.2 Hypergeometrische Verteilung	30
2.1.6.3 Normalverteilung $\mathcal{N}(\mu; \sigma^2)$	30
2.1.6.4 Logarithmische Normalverteilung $\mathcal{N}_L(\mu; \sigma^2)$	32
2.1.6.5 Weibullverteilung $\mathcal{W}(b; T)$	33
2.1.6.6 Logistische Verteilung $\mathcal{L}(\alpha; \beta)$	34
2.1.6.7 χ^2 -Verteilung	35
2.1.6.8 t-Verteilung	36
2.1.7 Bemerkung zur Wahl des Verteilungsmodells	37
2.2 Experimentelle Bestimmung der Bauteildauerfestigkeit	39
2.2.1 Durchführung und Auswertung von Horizontversuchen	40
2.2.1.1 Probitmethode	40
2.2.1.2 Maximum-Likelihood-Methode	41
2.2.1.3 Praktische Durchführung von Horizontversuchen	42
2.2.1.4 Auswertungsbeispiel	44
2.2.2 Durchführung und Auswertung von Treppenstufenversuchen	46
2.2.2.1 Nach Dixon/Mood	46
2.2.2.2 Nach Hück	48
2.2.2.3 Nach Deubelbeiss	50
2.2.2.4 Mit Probit- oder Maximum-Likelihood-Methode	51
2.2.2.5 Auswertungsbeispiel	51
2.2.3 Weitere Verfahren zur Bestimmung der Dauerfestigkeit	53

2.2.4	Bemerkung zur Umsortierung von Versuchsergebnissen	54
3	Simulation von Versuchen zur Bestimmung der Dauerfestigkeit	55
3.1	Bekannte Arbeiten	55
3.2	Ziele der eigenen Untersuchung	57
3.3	Beschreibung des Simulationsablaufes	57
3.3.1	Simulation der Versuchsreihen	59
3.3.2	Ein- und Ausgabeparameter	60
3.3.3	Statistische Auswertung der Simulationsergebnisse	62
3.3.4	Ermittlung von Korrelationskoeffizienten	62
3.4	Aufbereitung der Simulationsergebnisse	63
3.4.1	Verfahrensunabhängige Ergebnisse	63
3.4.1.1	Anzahl belegter Horizonte in Treppenstufenversuchen	63
3.4.1.2	Anzahl ungültiger Versuche	64
3.4.2	Erwartungswert bei bekannter Standardabweichung	64
3.4.2.1	Anteil auswertbarer Versuchsreihen	64
3.4.2.2	Systematischer und zufälliger Schätzfehler	65
3.4.2.3	Allgemeine Messunsicherheit	69
3.4.2.4	RMS-Fehler	70
3.4.2.5	Empirischer Vertrauensbereich	71
3.4.3	Schätzung der Standardabweichung	73
3.4.3.1	Anteil auswertbarer Versuchsreihen	73
3.4.3.2	Systematischer und zufälliger Schätzfehler	77
3.4.3.3	RMS-Fehler	78
3.4.3.4	Empirischer Vertrauensbereich	80
3.4.4	Erwartungswert bei unbekannter Standardabweichung	81
3.4.4.1	Systematischer und zufälliger Schätzfehler	81
3.4.4.2	Allgemeine Messunsicherheit	82
3.4.4.3	RMS-Fehler	83
3.4.4.4	Empirischer Vertrauensbereich	83
3.4.5	Quantil $x_{2,5\%}$ bei bekannter Standardabweichung	85
3.4.5.1	Systematischer und zufälliger Schätzfehler	85
3.4.5.2	RMS-Fehler	87
3.4.5.3	Empirischer Vertrauensbereich	88
3.4.6	Quantil $x_{2,5\%}$ bei unbekannter Standardabweichung	90
3.4.6.1	Systematischer und zufälliger Schätzfehler	90
3.4.6.2	RMS-Fehler	93
3.4.6.3	Empirischer Vertrauensbereich	93
4	Vergleich der Simulationsergebnisse	95
4.1	Vom Auswerteverfahren unabhängige Ergebnisse	95
4.2	Schätzung des Erwartungswertes bei bekannter Standardabweichung	99
4.2.1	Anteil auswertbarer Versuchsreihen	99
4.2.2	Mittlerer zufälliger und systematischer Schätzfehler	99
4.2.3	Verallgemeinerter Standardfehler	103
4.2.4	RMS-Fehler	108
4.2.5	Empirischer Vertrauensbereich	109

4.3	Schätzung der Standardabweichung	115
4.3.1	Anteil auswertbarer Versuchsreihen	115
4.3.2	Mittlerer zufälliger und systematischer Schätzfehler	115
4.3.3	RMS-Fehler	122
4.3.4	Empirischer Vertrauensbereich	126
4.4	Schätzung des 2,5 %-Quantils bei bekannter Standardabweichung	132
4.4.1	Mittlerer zufälliger und systematischer Schätzfehler	132
4.4.2	RMS-Fehler	138
4.4.3	Empirischer Vertrauensbereich	142
4.5	Schätzung des Erwartungswertes bei unbekannter Standardabweichung	146
4.5.1	Mittlerer zufälliger und systematischer Schätzfehler	146
4.5.2	Verallgemeinerter Standardfehler	150
4.5.3	Empirischer Vertrauensbereich	152
4.6	Schätzung des 2,5 %-Quantils bei unbekannter Standardabweichung	154
4.6.1	Mittlerer zufälliger und systematischer Schätzfehler	154
4.6.2	RMS-Fehler	157
4.6.3	Empirischer Vertrauensbereich	159
5	Zusammenfassung der Simulationsergebnisse	161
5.1	Bewertung der Auswertverfahren	161
5.1.1	Schätzung des Erwartungswertes	162
5.1.2	Schätzung der Standardabweichung	163
5.1.3	Schätzung des 2,5 %-Quantils	164
5.2	Vorgehen zur experimentellen Bestimmung der Verteilungsparameter	165
5.2.1	Bei bekannter Standardabweichung	165
5.2.2	Bei unbekannter Standardabweichung	166
5.2.3	Anwendungsbeispiel	169
5.2.4	Wahl der Fehlergröße	172
5.2.5	Anwendung bei nicht normalverteilten Grundgesamtheiten	172
5.3	Kritische Betrachtung zu Simulationsversuchen	173
	Literaturverzeichnis	175
A	Grafische Darstellung der Simulationsergebnisse	179
A.1	Verfahrensunabhängige Ergebnisse	179
A.1.1	Anzahl belegter Horizonte im Treppenstufenversuch	180
A.1.2	Anzahl ungültiger Versuche	186
A.2	Erwartungswertschätzung bei bekannter Standardabweichung	192
A.2.1	Mittlere Verzerrungen und Standardfehler Schätzfehler	192
	Treppenstufenversuche	192
	Horizontversuche	194
A.2.2	Verallgemeinerte Standardfehler	197
	Treppenstufenversuche	197
	Horizontversuche	198
A.2.3	RMS-Schätzfehler	200
	Treppenstufenversuche	200
	Horizontversuche	202

A.2.4	Mittlere verallgemeinerte 5%-Quantile	205
	Treppenstufenversuche	205
	Horizontversuche	206
A.3	Standardabweichung	208
A.3.1	Mittlere Verzerrungen und Standardfehler	208
	Treppenstufenversuche	208
	Horizontversuche	210
A.3.2	RMS-Fehler	213
	Treppenstufenversuche	213
	Horizontversuche	215
A.3.3	Mittlere 5%-Quantile	218
	Treppenstufenversuche	218
	Horizontversuche	219
A.4	Erwartungswert bei unbekannter Standardabweichung	221
A.4.1	Mittlere Verzerrungen und Standardfehler	221
	Treppenstufenversuche	221
	Horizontversuche	223
A.4.2	Verallgemeinerte Standardfehler	226
	Treppenstufenversuche	226
	Horizontversuche	227
A.4.3	RMS-Schätzfehler	229
	Treppenstufenversuche	229
	Horizontversuche	231
A.4.4	Mittlere verallgemeinerte 95 % Quantile	234
	Treppenstufenversuche	234
	Horizontversuche	235
A.5	Schätzung des 2,5%-Quantils bei bekannter Standardabweichung	237
A.5.1	Mittlere Verzerrungen und Standardfehler	237
	Treppenstufenversuche	237
	Horizontversuche	239
A.5.2	RMS-Fehler	242
	Treppenstufenversuche	242
	Horizontversuche	244
A.5.3	Mittlere 95%-Quantile	247
	Treppenstufenversuche	247
	Horizontversuche	248
A.6	Schätzung des 2,5%-Quantils bei unbekannter Standardabweichung	250
A.6.1	Mittlere Verzerrungen und Standardfehler	250
	Treppenstufenversuche	250
	Horizontversuche	252
A.6.2	RMS-Fehler	255
	Treppenstufenversuche	255
	Horizontversuche	257
A.6.3	Mittlere 95 %-Quantile	260
	Treppenstufenversuche	260
	Horizontversuche	261

B	Sonstiges	263
B.1	Treppenstufenauswertung	263
B.1.1	Treppenstufenauswertung nach HÜCK	263
B.1.1.1	Schätzung der Standardabweichung	263
B.1.1.2	Schätzung der Standardfehler	264
B.2	Zur Umrechnungen zwischen Verteilungsparametern	266
B.2.1	Näherung für die Parameter der WEIBULL-Verteilung	266
B.2.2	Näherung aus logarithmischer Normalverteilung	266
B.2.3	Näherung aus Normalverteilung	267
B.3	Zu erwartende Stützzahlen für ein ideal treffsicheres Stützzahlkonzept	268

Verzeichnis der Formelzeichen und Symbole

Bezeichnung verschiedener statistischer Verteilungen

Symbol	Bedeutung
$\text{Bi}(n; p)$...	Binomial-Verteilung für n Wiederholungen mit der Erfolgswahrscheinlichkeit p
$\text{Exp}(\lambda)$...	Exponentialverteilung mit der Ereignisrate λ
$\mathcal{L}(\alpha; \beta)$...	Logistische Verteilung mit dem Erwartungswert α und der Standardabweichung β
$\mathcal{N}(\mu; \sigma^2)$...	Normalverteilung mit dem Erwartungswert μ und der Varianz σ^2
$\mathcal{N}_L(\mu; \sigma^2)$...	Logarithmische Normalverteilung mit dem Erwartungswert μ und der Varianz σ^2
t_n ...	t-Verteilung mit dem Freiheitsgrad n
$\mathcal{W}(b; T)$...	Weibullverteilung mit Formparameter b und charakteristischer Lebensdauer T
χ_n^2 ...	χ^2 -Verteilung mit dem Freiheitsgrad n

Lateinische Symbole

Symbol	Bedeutung
d ...	Stufenabstand in Treppenstufenversuchen
d' ...	auf die Standardabweichung der Grundgesamtheit bezogener Stufenabstand in Treppenstufenversuchen
$C_{(\cdot)}$...	zufälliger Schätzfehler eines Schätzers (\cdot)
$C_{(\cdot),\max}$...	größter auftretender zufälliger Schätzfehler eines Schätzers (\cdot)
$C_{(\cdot),\min}$...	kleinster auftretender zufälliger Schätzfehler eines Schätzers (\cdot)
$\bar{C}_{(\cdot)}$...	mittlerer auftretender zufälliger Schätzfehler eines Schätzers (\cdot)
$\mathbf{E}[\cdot]$...	Erwartungswert einer Zufallsgröße oder Stichprobe
$G_{(\cdot)}$...	systematischer Schätzfehler eines Schätzers (\cdot)
$G_{(\cdot),\max}$...	größter auftretender systematischer Schätzfehler eines Schätzers (\cdot)
$G_{(\cdot),\min}$...	kleinster auftretender systematischer Schätzfehler eines Schätzers (\cdot)
$\bar{G}_{(\cdot)}$...	mittlerer auftretender systematischer Schätzfehler eines Schätzers (\cdot)

Inhaltsverzeichnis

h	...	Anzahl der Horizonte
$\mathbf{H}[\cdot]$...	Schiefe einer Zufallsgröße oder Stichprobe
m	...	Schätzer für den Erwartungswert einer normalverteilten Zufallsgröße
m'	...	auf die tatsächliche Standardabweichung bezogener Abstand zwischen dem Schätzer und dem wahren Wert des Erwartungswertes
$m'_{a\%}$...	Quantil des Schätzers m' zur Wahrscheinlichkeit $a\%$
$m'_{a\%,\max}$...	größtes auftretendes Quantil des Schätzers m'
$m'_{a\%,\min}$...	kleinstes auftretendes Quantil des Schätzers m'
$\overline{m'}_{a\%}$...	mittleres Quantil des Schätzers m'
$\overline{m'}_{ a\% }$...	verallgemeinertes mittleres Quantil des Schätzers m'
m^*	...	auf die geschätzte Standardabweichung bezogener Abstand zwischen dem Schätzer und dem wahren Wert des Erwartungswertes
$m^*_{a\%}$...	Quantil des Schätzers m^* zur Wahrscheinlichkeit $a\%$
$m^*_{a\%,\max}$...	größtes auftretendes Quantil des Schätzers m^*
$m^*_{a\%,\min}$...	kleinstes auftretendes Quantil des Schätzers m^*
$\overline{m^*}_{a\%}$...	mittleres Quantil des Schätzers m^*
$\overline{m^*}_{ a\% }$...	verallgemeinertes mittleres Quantil des Schätzers m^*
$\mathbf{M}[\cdot]$...	Median einer Zufallsgröße oder Stichprobe
n	...	Anzahl der (im Sinne der Auswertung gültigen) Einzelversuche
n_u	...	Anzahl der Versuche, die in der Auswertung nicht berücksichtigt werden, Anschnitt der Versuchsreihe
$n_{VR,\text{rel},m}$...	relative Anzahl auswertbarer Versuchsreihen bei Schätzung des Erwartungswertes
$n_{VR,\text{rel},s}$...	relative Anzahl auswertbarer Versuchsreihen bei Schätzung der Standardabweichung
$RMS[\cdot]$...	RMS-Fehler eines Schätzers
$RMS_{(\cdot),\max}$...	größter auftretender RMS-Fehler eines Schätzers
$RMS_{(\cdot),\min}$...	kleinster auftretender RMS-Fehler eines Schätzers
$\overline{RMS}_{(\cdot)}$...	mittlerer RMS-Fehler eines Schätzers
$\Delta RMS_{(\cdot)}$...	Spanne zwischen größtem und kleinstem RMS-Fehler eines Schätzers
s	...	Schätzer für die Standardabweichung einer normalverteilten Zufallsgröße
s'	...	auf die tatsächliche Standardabweichung bezogener Schätzer der Standardabweichung
s'_{\lg}	...	Logarithmus des auf die tatsächliche Standardabweichung bezogenen Schätzers der Standardabweichung
$s'_{\lg,a\%}$...	Quantile des Schätzers s'_{\lg} zur Wahrscheinlichkeit $a\%$
$s'_{\lg,a\%,\max}$...	größtes auftretendes Quantil des Schätzers s'_{\lg}
$s'_{\lg,a\%,\min}$...	kleinstes auftretendes Quantil des Schätzers s'_{\lg}
$\overline{s'}_{\lg,a\%}$...	mittleres Quantil des Schätzers s'_{\lg}
s^2	...	Schätzer für die Varianz einer normalverteilten Zufallsgröße

s'^2	...	auf die tatsächliche Varianz bezogener Schätzer der Varianz
$\mathbf{S}[\cdot]$...	Standardabweichung einer Zufallsgröße oder Stichprobe
T	...	Streuspanne
$U_{(\cdot)}$...	verallgemeinerter Standardfehler des Schätzers (\cdot)
$\mathbf{V}[\cdot]$...	Varianz einer Zufallsgröße oder Stichprobe
$\mathbf{W}[\cdot]$...	Wölbung einer Zufallsgröße oder Stichprobe
x	...	Realisation (Stichprobe) einer Zufallsgröße
$x_{2,5\%}$...	Schätzer für das 2,5 %-Quantil einer normalverteilten Zufallsgröße
x'	...	auf die tatsächliche Standardabweichung bezogener Abstand zwischen dem Schätzer und dem wahren Wert des 2,5 %-Quantils
$x'_{a\%}$...	Quantile des Schätzers x' zur Wahrscheinlichkeit $a\%$
$x'_{a\%,\max}$...	größtes auftretendes Quantil des Schätzers x'
$x'_{a\%,\min}$...	kleinstes auftretendes Quantil des Schätzers x'
$\overline{x'}_{a\%}$...	mittleres Quantil des Schätzers x'
x^*	...	auf die geschätzte Standardabweichung bezogener Abstand zwischen dem Schätzer und dem wahren Wert des 2,5 %-Quantils
$x^*_{a\%}$...	Quantile des Schätzers x^* zur Wahrscheinlichkeit $a\%$
$x^*_{a\%,\max}$...	größtes auftretendes Quantil des Schätzers x^*
$x^*_{a\%,\min}$...	kleinstes auftretendes Quantil des Schätzers x^*
$\overline{x^*}_{a\%}$...	mittleres Quantil des Schätzers x^*
x_0	...	Abstand zwischen Versuchsreihenstartwert und Erwartungswert der Grundgesamtheit in Treppenstufen- und Horizontversuchen
x'_0	...	auf die Standardabweichung der Grundgesamtheit bezogener Abstand zwischen Versuchsreihenstartwert und Erwartungswert der Grundgesamtheit in Treppenstufen- und Horizontversuchen
z'	...	auf die Standardabweichung der Grundgesamtheit bezogener Abstand zwischen dem höchstem und dem niedrigsten Lathorizont bei Horizontversuchen

Griechische Symbole

Symbol	Bedeutung
$\Gamma(\cdot)$... Gamma-Funktion
μ_G	... tatsächlicher Erwartungswert einer normalverteilten Zufallsgröße
μ_L	... Erwartungswert der logarithmierten logarithmisch normalverteilten Zufallsgröße
σ_G	... tatsächliche Standardabweichung einer normalverteilten Zufallsgröße
σ_L	... Standardabweichung der logarithmierten logarithmisch normalverteilten Zufallsgröße
σ_L^2	... Varianz der logarithmierten logarithmisch normalverteilten Zufallsgröße

1 Einleitung

1.1 Motivation

Die entscheidende Frage der Betriebsfestigkeit lautet: Hält das Bauteil den vorgegebenen zeitabhängigen Belastungen stand? Um diese Frage zu beantworten, gibt es zwei Möglichkeiten. Entweder der Gefragte führt den experimentellen Nachweis, indem er das Bauteil mit den realen, innerhalb der Bauteillebensdauer auftretenden Lasten beansprucht, oder er versucht bereits bekannte Versuchsergebnisse für andere Bauteile auf das aktuelle Bauteil zu übertragen. Während der erste Fall als ein kaum zu erreichendes und mit großen Kosten verbundenes Ideal betrachtet werden muss, ist das letztere Vorgehen mit geringerem Aufwand, aber wie von EULITZ [11] gezeigt, mit großen Unsicherheiten verbunden und kann in einer Vielzahl von Abstraktionsstufen durchgeführt werden. So können zum einen Annahmen über die tatsächlich im Betrieb auftretenden Beanspruchungen aus bisher bekannten Bauteillasten getroffen werden und zum anderen können bereits bekannte Bauteilbeanspruchbarkeiten auf andere Bauteile übertragen werden. Die Möglichkeiten reichen hier bis hin zur reinen Verwendung statischer Festigkeitskennwerte aus dem Werkstoffdatenblatt. Dabei steigt mit zunehmendem Grad der Abstraktion auch die Unsicherheit, mit der die anfangs gestellte Frage beantwortet werden kann.

Vor allem dem Thema der Übertragbarkeit von Bauteilbeanspruchbarkeiten widmen sich eine Vielzahl von wissenschaftlichen Arbeiten auf dem Gebiet der Betriebsfestigkeit. Im einfachsten Fall werden dazu Wechselfestigkeits einer ungekerbten Probe oder einer gekerbten Probe auf eine andere Probengeometrie gleichen Werkstoffs übertragen, z. B. in [3]. Der Beweis der aufgestellten Hypothesen erfolgt dann durch den Vergleich einzelner, meist anhand der Literatur entnommener Versuche. Durch die Angabe des Fehlers zwischen den experimentell ermittelten und den abgeschätzten Festigkeiten können auch verschiedene Hypothesen untereinander verglichen werden. Die experimentell ermittelten Festigkeiten sind Messwerte und als solche zwangsweise mit einer mehr oder weniger großen Unsicherheit behaftet. Ein Teil des angegebenen Fehlers beruht also auf anzunehmenden unzureichenden Modellvorstellungen. Ein weiterer Teil des Fehlers ist aber der zufälligen Abweichung der experimentellen Ergebnisse geschuldet. Diese Tatsache wird oft nicht berücksichtigt. Zwar kann der Anteil des zufälligen experimentellen Fehlers durch eine entsprechend hohe Anzahl an Vergleichsversuchen verringert werden, wie groß diese Verringerung aber tatsächlich ausfällt, kann nur angegeben werden, wenn die Größe des experimentellen Fehlers bekannt ist.

Ein bewährtes Konzept der Betriebsfestigkeit ist es, die Schwingfestigkeit σ_w einer (ungekerbten) Werkstoffprobe durch eine Stützzahl n auf die Schwingfestigkeit eines Bauteils

1 Einleitung

$\sigma_{wk} = n \cdot \sigma_w$ zu übertragen¹. In dieser Stützzahl werden dann alle die Schwingfestigkeit beeinflussenden Faktoren (Größeneinflüsse) zusammengefasst. Eine Bewertung der verschiedenen Ansätze zur Übertragung der Schwingfestigkeitskennwerte erfolgt durch den Vergleich von experimentell ermittelten Stützzahlen n_{exp} und theoretisch berechneten Stützzahlen n_{th} . Diese relativen Stützzahlen werden als Zufallsgrößen aufgefasst und als solche, wie z. B. in [13] gezeigt, statistisch ausgewertet. Als die treffsichersten und damit vorzuziehenden Übertragungskonzepte werden jene Konzepte herausgestellt, deren relative Stützzahlen über möglichst viele verschiedene Versuchsergebnisse hinweg die geringsten Streuungen aufweisen. Nicht vergessen werden darf dabei allerdings, dass jede relative Stützzahl mit der Unsicherheit behaftet ist, mit der die jeweiligen Beanspruchbarkeiten experimentell bestimmt wurden. Damit sind auch die ermittelten Streuungen der relativen Stützzahl von Unsicherheiten der experimentellen Ergebnisse abhängig.²

Ein weiteres Beispiel für die Notwendigkeit der Kenntnis der Unsicherheiten experimentell ermittelter Wechselfestigkeiten ist die Bestimmung der Mittelspannungsempfindlichkeit. Diese wird allgemein durch Wöhlerversuche mit einem Spannungsverhältnis von $R = -1$ und $R = 0$ ermittelt. In den seltensten Fällen wird dabei der Tatsache Rechnung getragen, dass beide Größen experimentell ermittelt und damit mit einer Unsicherheit behaftet sind.³ Das führt zum Beispiel dazu, dass experimentell ermittelte Mittelspannungsempfindlichkeiten mit Besonderheiten im Gefüge erklärt werden, siehe [18] S.21. Dabei wäre es viel plausibler, erst einmal die Aussagefähigkeit für aus Treppenstufenversuchen mit lediglich 8 bzw. 6 Versuchen gewonnenen Mittelspannungsempfindlichkeiten zu hinterfragen.⁴

Die Kenntnis der Unsicherheiten experimentell ermittelter Werkstoff- und Bauteilfestigkeiten ist also grundlegender Bestandteil für die Auslegung von Bauteilen oder die Ableitung weiterer Materialparameter. Werden sie ignoriert, müssen sie im besten Fall durch übermäßig große Sicherheitsfaktoren ausgeglichen werden, im schlechtesten Fall führen sie zum Bauteilversagen. Auch bei der Entwicklung neuer Ansätze zur Übertragbarkeit sind bei der Bewertung selbiger die Unsicherheiten experimenteller Ergebnisse nicht außer Acht zu lassen.

¹Hier bezeichnet σ_w und σ_{wk} jeweils die berechnete Maximalspannung bzw. Kerbspannung im Prüfquerschnitt oder der Kerbe.

²Wie groß diese Unsicherheiten ausfallen, wird im Anhang B.3 für den Fall eines ideal treffsicheren Stützzahlkonzeptes abgeschätzt.

³Wird die Mittelspannung M aus der Wechselfestigkeit $\sigma_{a,R=-1}$ und der Schwellfestigkeit $\sigma_{a,R=0}$ durch $M = \sigma_{a,R=-1}/\sigma_{a,R=0} - 1$ berechnet und sind die experimentell ermittelten Festigkeiten mit den Unsicherheiten $u_{\sigma_{a,R=-1}}$ und $u_{\sigma_{a,R=0}}$ behaftet, dann berechnet sich die Unsicherheit u_M der Mittelspannungsempfindlichkeit näherungsweise durch:

$$u_M = \sqrt{\left(\frac{\partial M}{\partial \sigma_{a,R=-1}}\right)^2 u_{\sigma_{a,R=-1}}^2 + \left(\frac{\partial M}{\partial \sigma_{a,R=0}}\right)^2 u_{\sigma_{a,R=0}}^2}$$

⁴ In diesem Fall wurde aus einer mittels Treppenstufenversuchen mit $n = 8$ Einzelversuchen bestimmten Wechselfestigkeit $\sigma_{a,R=-1} = 279$ MPa und einer aus Treppenstufenversuche mit $n = 6$ Einzelversuchen ermittelten Schwellfestigkeit $\sigma_{a,R=0} = 174$ MPa eine Mittelspannungsempfindlichkeit von $M_\sigma = 0,59$ berechnet. Die Streuspanne der Festigkeiten kann nach [14](S. 527) überschlägig mit $T_\sigma = 1,26$ angesetzt werden. Mit Vorgriff auf die Ergebnisse aus Abschnitt 4.2.3 der vorliegenden Arbeit folgen daraus Unsicherheiten von etwa $u_{\sigma_{a,R=-1}} = 24$ MPa bzw. $u_{\sigma_{a,R=0}} = 17$ MPa für die experimentell bestimmten mittleren Festigkeiten. Für die Mittelspannungsempfindlichkeit lässt sich damit eine Unsicherheit von etwa $u_M = 0,21$ abschätzen. Das Ergebnis sollte also besser $M = 0,59 \pm 0,21$ lauten.

Die experimentelle Bestimmung der Werkstoff- und Bauteilfestigkeiten erfolgt durch mehrere Wöhler- bzw. Einstufenversuche und deren entsprechende statistische Auswertung. Während im Bereich der Zeitfestigkeit die Versuche bis zum Versagen der Probe durchgeführt werden und sich die Beanspruchbarkeit aus der Auswertung der erreichten Lastwechsel bestimmt, können im Bereich der Dauerfestigkeit, also bei sehr kleinen Beanspruchungen und sehr großen zu erwartenden Lastwechselzahlen, nicht alle Proben bis zum Versagen geprüft werden. Die Bestimmung der Beanspruchbarkeiten ergibt sich im letzteren Fall immer aus dem Verhältnis von auf verschiedenen Lastniveaus ausgefallenen Proben (Brüchen) und nicht ausgefallenen Proben (Durchläufern). Dafür werden die Versuche in Form von Treppenstufenversuchen oder Horizontversuchen durchgeführt. Für die Auswertung der Versuchsreihen stehen dann verschiedene Methoden zur Verfügung, die alle unterschiedliche Schätzwerte für die Parameter der Festigkeitsverteilung liefern. Wie groß dabei die Unsicherheiten ausfallen, d. h. wie genau die ermittelten Festigkeitsparameter die wahren Werten der Festigkeitsverteilung wiedergeben, ist bereits von verschiedenen Autoren dargelegt worden. Für die Auswertung von Treppenstufenversuchen werden von DIXON und MOOD in [9] Standardfehler für die ermittelten Kennwerte der Festigkeit angegeben. Für ein verbessertes Treppenstufenverfahren liefert auch HÜCK in [21] die Standardfehler für die experimentell bestimmte mittlere Festigkeit und die Streuung der Festigkeit. Hierzu bedient sich der Autor der Möglichkeiten der Monte-Carlo-Simulationen und führt virtuelle Versuchsreihen durch, anhand deren Streuung die Standardfehler der Versuchsergebnisse ermittelt werden. Da dabei aber alle Versuchsreihen im Mittel von einem der mittleren Festigkeit entsprechenden Lastniveau beginnen, können systematische Schätzfehler nicht erkannt werden. Andere Autoren stellen dies fest, und führen Monte-Carlo-Simulationen mit verschiedenen mittleren Versuchsreihenstartwerten durch, z. B. LIU in [26]. Auch von MÜLLER werden in [32] und [31] durch Simulation die zu erwartenden Unsicherheiten der experimentell ermittelten Parameter der Festigkeitsverteilung für verschiedene Auswertungsmethoden von Treppenstufenversuche und Horizontversuche untersucht. Dabei werden die Angaben über zu erwartende Fehler bei der Parameterbestimmung immer in Abhängigkeit der zu Beginn der Versuchsreihe vorhandenen Proben durchgeführt. Da aber bei verschiedenen Auswertemethoden auch einzelne Versuchsergebnisse als für die Auswertung ungültige Ergebnisse betrachtet werden können, ist es schwierig, schon bekannte Literaturergebnisse, bei denen lediglich die Anzahl der für die Auswertung verwendeten Proben angegeben ist, hinsichtlich ihrer Unsicherheit zu bewerten. Alle bisherigen Untersuchungen konzentrieren sich außerdem darauf, die Unsicherheit der Festigkeitsparameter in Form der mittleren Festigkeit und einer Streuung der Festigkeit zu ermitteln. In der praktischen Anwendung dienen die Versuche aber dazu, die Festigkeit für eine möglichst geringe Ausfallwahrscheinlichkeit zu ermitteln, denn nur mit solchen Werten kann eine Bauteilauslegung nach den gültigen Normen und Regelwerken, z. B. nach der FKM-Richtlinie [35] erfolgen. Ein Vergleich verschiedener Auswertemethoden bezüglich der Treffsicherheit einer solchen geschätzten Mindestfestigkeit fehlt bisher ebenso, wie der einheitliche Vergleich der verschiedenen Auswerteverfahren bei der Schätzung der mittleren Festigkeit und der Festigkeitsstreuung.

1.2 Inhalt und Aufbau der Arbeit

Ziel der vorliegenden Arbeit ist es daher, die Unsicherheiten bei der Bestimmung der Parameter der Schwingfestigkeit bei sehr großen Schwingungszahlen für verschiedene Auswertemethoden von Horizontversuchen und Treppenstufenversuchen zu ermitteln und vergleichend gegenüberzustellen. Hierzu werden Monte-Carlo-Simulationen durchgeführt und virtuelle Versuchsreihen in Form von Treppenstufenversuchen und Horizontversuchen erstellt. Für die Auswertung solcher Versuche existieren verschiedene Auswertemethoden. Diese liefern jeweils unterschiedliche Schätzwerte für die Parameter der Festigkeitsverteilung. Die Abweichungen der geschätzten Parameter von den tatsächlichen Parametern sind selbst als Zufallsgrößen anzusehen, von denen durch die Simulationen jeweils eine Stichprobe gezogen wird. Anhand dieser Stichproben können die zu erwartenden Abweichungen der Parameterschätzungen ermittelt werden. Dabei werden die systematischen und zufälligen Abweichungen, die RMS-Fehler der Schätzungen sowie die Konfidenzintervalle betrachtet. Dies ermöglicht den Vergleich der verschiedenen Auswertemethoden. Als Parameter der Festigkeitsverteilung werden die mittlere Festigkeit, die Festigkeitsstreuung sowie die geschätzte Festigkeit für eine Ausfallwahrscheinlichkeit von 2,5 % betrachtet. Der Einfluss einer systematischen Startwertverschiebung wird ebenso untersucht, wie der Einfluss unterschiedlicher Definitionen des Versuchsreihenanschnitts, also jenen anfänglichen Versuchen, die in der Auswertung nicht berücksichtigt werden. Durch die Auswertung der Ergebnisse in Abhängigkeit der für die Auswertung tatsächlich betrachteten Versuche ist es möglich, die Unsicherheiten für eigene experimentell ermittelte Schwingfestigkeiten oder aber für in der Literatur angegebene Werte zu bestimmen.

Sowohl die verschiedenen Schätzverfahren als auch die Auswertung der durch die Simulation gezogenen Stichproben erfordern grundlegende Kenntnisse der beschreibenden und schließenden Statistik. Für eine korrekte Simulation ist es außerdem erforderlich, die zu untersuchenden Schätzverfahren zur experimentellen Bestimmung der Festigkeitsverteilung bei hohen Schwingungszahlen genauer zu beleuchten. Beides ist Gegenstand des zweiten Kapitels. Wie die Simulationsversuche im einzelnen aufgebaut sind, und welche Einschränkungen dabei getroffen werden, wird im Kapitel 3 erläutert. Hier wird auch beschrieben, wie sich aus den Simulationsergebnissen die Fehlergrößen zur Bewertung der Aussagegenauigkeiten der verschiedenen Schätzverfahren berechnen lassen. Dazu werden die Ergebnisse anhand von Versuchsreihen mit 14 bzw. 15 Einzelversuchen exemplarisch dargestellt. Für alle anderen Versuchsanzahlen sind die aufbereiteten Fehlergrößen als grafische Darstellungen im Anhang zu finden. Im Kapitel 4 werden die so gewonnenen Ergebnisse genauer betrachtet und die verschiedenen Arten der Versuchsführung und -auswertung miteinander verglichen. Für ein Auswertungsbeispiel werden im Kapitel 5 die anzusetzenden Fehler bestimmt, die Ergebnisse zusammengefasst und kritisch betrachtet.

2 Grundlagen

2.1 Grundlagen der statistischen Versuchsauswertung

Der folgende Abschnitt behandelt Grundlagen der Statistik. Er ist dabei auf die nach Meinung des Autors wesentlichen zur ingenieurmäßigen Bearbeitung der interessierenden Fragestellungen nötigen Teile beschränkt und kann auch bei Unklarheiten im Laufe des weiteren Textes zum Nachschlagen verwendet werden. Das in diesem Abschnitt Wiedergegebene ist allgemein bekannt und kann aus verschiedensten Lehrbüchern mit Abweichungen in der Notation nachvollzogen werden. Der Autor empfand folgende Lehrbücher als hilfreich: [40], [39], [41], [17], [33].

2.1.1 Grundlegende Begriffe

Die Statistik befasst sich im Allgemeinen mit der Gewinnung und Aufbereitung von Daten für ein oder mehrere Untersuchungsobjekte. Die interessierenden Eigenschaften des Untersuchungsobjektes werden Merkmale genannt. Werden die Ausprägungen eines Merkmals beobachtet, kann das Ergebnis dieser Beobachtungen auf einer Skala, die alle möglichen Ausprägungen beinhaltet, festgehalten werden. Je nach Art der verwendeten Skala kann zwischen verschiedenen Arten der Merkmalsausprägung unterschieden werden:

- nominale Merkmale Die Ausprägungen des Merkmals sind nicht vergleichbar und unterliegen keiner Rangfolge, z. B. *rot, grün, blau*.
- ordinale Merkmale Die Ausprägungen des Merkmals unterliegen einer Reihenfolge, die Abstände zwischen verschiedenen Ausprägungen besitzen keine Aussagekraft, z. B. *Schulnoten*.
- metrische Merkmale Die Ausprägungen des Merkmals unterliegen einer Reihenfolge, die Abstände zwischen den Ausprägungen sind interpretierbar. Besitzt die verwendete Skala einen absoluten Nullpunkt, ist auch das Verhältnis zweier Ausprägungen zueinander interpretierbar, z. B. *Temperaturen in Kelvin*.

Andere Arten der Klassifizierung sind z. B. die Einteilung in diskrete und stetige Merkmalsausprägungen. Die Menge aller der Untersuchung zugrunde liegenden Objekte wird als Grundgesamtheit bezeichnet. Da es meist nicht möglich oder sinnvoll ist, alle Objekte der Grundgesamtheit zu untersuchen, wird dieser eine Stichprobe entnommen. Anhand der aus

der Stichprobe ermittelten Kennwerte wird dann auf die Kennwerte der Grundgesamtheit geschlossen.

2.1.2 Beschreibung einer Stichprobe

Wird an n unabhängigen Untersuchungsobjekten ein Merkmal X beobachtet oder an einem Untersuchungsobjekt ein Merkmal unabhängig n mal gemessen, so wird eine Stichprobe $(x_1, \dots, x_i, \dots, x_n)$ mit n Merkmalswerten ermittelt. Diese bilden die Urliste und damit den Ausgang aller weiteren Betrachtungen. Werden die Einträge der Urliste in nicht fallender Reihenfolge sortiert, entsteht daraus einer Rangliste $(x_{(1)}, \dots, x_{(i)}, \dots, x_{(n)})$. Im Falle diskreter Merkmale kann sowohl die Urliste als auch die Rangliste nach Merkmalswerten gleicher Ausprägung durchsucht und die absolute Häufigkeit jeder auftretenden Ausprägung festgehalten werden. In einem Stabdiagramm lassen sich die so ermittelten absoluten Häufigkeiten oder die auf die Gesamtanzahl der Beobachtungen bezogenen relativen Häufigkeiten über den zugehörigen Merkmalsausprägungen übersichtlich darstellen. Im Falle stetiger oder sehr vieler verschiedener Merkmalsausprägungen werden verschiedene Merkmalswerte in Klassen zusammengefasst. Es ist dann möglich, die absoluten oder relativen Klassenhäufigkeiten in einem Histogramm, d. h. einem Stabdiagramm, in dem die Balkenbreite jeweils der zugehörigen Klassenbreite entspricht, darzustellen. Hierbei ist nach [16] darauf zu achten, dass die gewählten Klassen aneinander stoßen und bei ungleichmäßiger Klasseneinteilung die Fläche eines Balkens proportional zur ermittelten Häufigkeit ist. Die empirische Verteilungsfunktion wird durch Summation aller relativen Häufigkeiten für alle Merkmalsausprägungen, die kleiner oder gleich dem betrachteten Merkmalswert sind, ermittelt.

Für eine weiterführende Beschreibung einer Stichprobe werden Maßzahlen als Lagemaße zur Beschreibung des „Zentrums“ der beobachteten Werte und als Streuungsmaße zur Beschreibung der Abweichung der Beobachtungen von diesem „Zentrum“ ermittelt. Die wohl bedeutendsten Lagemaße sind der Mittelwert \bar{x} und der Median \tilde{x} mit:

$$\begin{aligned} \bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i, \\ \tilde{x} &= \begin{cases} x_{((n+1)/2)} & , \text{ wenn } n \text{ ungerade} \\ \frac{1}{2}(x_{(n/2)} + x_{((n/2)+1)}) & , \text{ wenn } n \text{ gerade.} \end{cases} \end{aligned}$$

Der Median kann im Gegensatz zum Mittelwert auch bei ordinal skalierten Merkmalen verwendet werden und ist unempfindlicher gegenüber Ausreißern innerhalb der Beobachtungsreihe. Weitere gegenüber Ausreißern unempfindlichere Lagemaße sind der gestutzte sowie der winsorisierte Mittelwert, auf die hier nicht näher eingegangen wird. Als Streumaße finden die Stichprobenvarianz s^2 und die sich daraus ergebende Stichprobenstandardabweichung s mit

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \tag{2.1}$$

Anwendung.¹ Um Stichproben mit unterschiedlichen Mittelwerten hinsichtlich ihrer Streuung zu vergleichen, wird der Variationskoeffizient

$$V = \frac{s}{\bar{x}}$$

verwendet. Eine allgemeinere Beschreibung der Maßzahlen kann über die Angabe der Momente m_k und zentralen Momente m'_k einer Stichprobe erreicht werden. Mit k als Ordnung der Momente sind diese definiert als:

$$m_k = \frac{1}{n} \sum_{i=1}^n x_i^k,$$

$$m'_k = \frac{1}{n} \sum_{i=1}^n (x_i - m'_1)^k.$$

Dadurch lassen sich leicht weitere Maßzahlen angeben, z. B. die empirische Schiefe v und die empirische Wölbung w einer Stichprobe, mit denen sich die Art und Weise der Verteilung der Stichprobenwerte genauer beschreiben lässt:

$$v = \frac{m_3}{s^3},$$

$$w = \frac{m_4}{s^4}.$$

Bei großen Stichprobenumfängen ist auch die Angabe von Stichprobenquantilen x_α eine Möglichkeit um die Stichprobe zu beschreiben. Hierzu wird die Rangliste der n Beobachtungen betrachtet und der Wert $k = n \cdot \alpha$ berechnet. Ist der Wert k gerade eine ganze Zahl, wird das zu einer Summenhäufigkeit α gehörende Quantil durch

$$x_\alpha = x_{(k)}$$

bestimmt, andernfalls wird der Wert k zur nächsten ganzen Zahl aufgerundet und es gilt

$$x_\alpha = \frac{1}{2} (x_{(k)} + x_{(k+1)}).$$

Ist die Stichprobe mit den Mitteln der deskriptiven Statistik beschrieben, wird im nächsten Schritt versucht ein mathematisches Modell zur Beschreibung der Grundgesamtheit zu finden. Die Wahl des Verteilungsmodells ist dabei meist durch die bisherigen Erfahrungen des Anwenders geprägt. Nach der Wahl des Modells müssen die Modellparameter bzw. Verteilungsparameter bestimmt werden. Dies erfolgt in der Form von Punktschätzern, also durch die Angabe konkreter Werte für jeden Verteilungsparameter. Dabei liegt es in der Natur der Sache, dass die geschätzten Modellparameter nicht mit den tatsächlichen Parametern übereinstimmen. Durch die Angabe eines Konfidenzintervalls bzw. Vertrauensbereichs zu jedem geschätzten Parameter kann deutlich gemacht werden, wie genau die Verteilungsparameter anhand der vorliegenden Stichprobe geschätzt werden können. Außerdem können in der vorhandenen Stichprobe auftretende extreme Beobachtungen mithilfe von Ausreißertests als Ausreißer erkannt und von der Auswertung ausgeschlossen werden. Mit einem Anpassungstest ist es des Weiteren möglich, zu zeigen, dass das gewählte Modell samt Parametern der gezogenen Stichprobe nicht widerspricht.

¹Verschiedentlich wird die Varianz einer Stichprobe mit dem Nenner $n - 1$ angegeben, da sie dazu benutzt wird, die Verteilungsparameter der Grundgesamtheit zu schätzen. Für die ausschließliche Beschreibung einer Stichprobe ist das nicht notwendig. Außerdem verschwindet der Unterschied für große Stichproben.

2.1.3 Schätzen von Verteilungsparametern

Für das angenommene Verteilungsmodell müssen nun die Verteilungsparameter möglichst zutreffend ermittelt werden. Ziel ist es, aus n Beobachtungen x_1, \dots, x_n der Zufallsgröße² X_1, \dots, X_n einen unbekanntem Parameter ϑ der angenommenen Wahrscheinlichkeitsverteilung abzuschätzen. Es wird also eine Funktion $\hat{\vartheta}(x_1, \dots, x_n)$ gesucht, die möglichst gleich dem wahren Parameter ϑ der Verteilung ist oder zumindest nah an ihn herankommt. An solch einen Schätzer werden verschiedene Anforderungen gestellt, nach [41], S.280 sollte ein Schätzer:

1. Möglichst keinen systematischen Fehler aufweisen, d.h. erwartungstreu³ sein mit

$$\mathbf{E}[\hat{\vartheta}] = \vartheta.$$

2. Auch für kleine Stichprobenumfänge eine möglichst geringe Varianz $\mathbf{V}[\hat{\vartheta}]$ aufweisen, d. h. effizient sein mit

$$\mathbf{V}[\hat{\vartheta}] = \text{minimal.}$$

3. Einen möglichst kleinen Fehler für große Stichprobenumfänge aufweisen, d. h. konsistent sein mit

$$\hat{\vartheta} \approx \vartheta \quad \text{für großes } n.$$

4. Alle Informationen in der Stichprobe über den Parameter nutzen, d. h. suffizient und robust sein gegenüber Abweichungen vom angenommenen Modell.

5. Normalverteilt (zumindest asymptotisch) sein, d. h.

$$\frac{\hat{\vartheta} - \mathbf{E}[\hat{\vartheta}]}{\sqrt{\mathbf{V}[\hat{\vartheta}]}} \approx \mathcal{N}(0; 1).$$

So ist zum Beispiel der Mittelwert aller Beobachtungen ein erwartungstreuer Schätzer $\hat{\mu}$ für den Erwartungswert μ einer $\mathcal{N}(\mu; \sigma^2)$ -Verteilung, da gilt:

$$\mathbf{E}[\hat{\mu}] = \mathbf{E}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n \mathbf{E}[X_i] = \mu.$$

Darüber hinaus ist $\hat{\mu}$, wie in [41] auf S. 281 nachzulesen, auch ein effizienter, konsistenter und suffizienter Schätzer für μ einer $\mathcal{N}(\mu; \sigma^2)$ -Verteilung.

Die Güte eines Schätzers kann über den mittleren quadratischen Fehler bzw. mean squared error (MSE) nach Gleichung 2.2 angegeben werden. Dieser beschreibt die zu erwartende

²Die Zufallsgröße X_i ist dann z. B. die i -te unabhängige Messung der Größe X . Jede Messung wird für sich als Zufallsgröße aufgefasst.

³Erwartungstreu bedeutet, dass der Erwartungswert $\mathbf{E}[\hat{\vartheta}]$ des Schätzers $\hat{\vartheta}$ dem Wert des wahren Parameters ϑ der Verteilung entspricht.

quadratische Abweichung des Schätzers vom wahren Parameter und kann mit dem Verschiebungssatz bzw. dem Satz von Steiner aus der Varianz nach Gleichung 2.3 ermittelt werden.

$$MSE[\hat{\vartheta}, \vartheta] = \mathbf{E}[(\hat{\vartheta} - \vartheta)^2] \quad (2.2)$$

$$= \mathbf{V}[\hat{\vartheta}] + (\mathbf{E}[\hat{\vartheta}] - \vartheta)^2. \quad (2.3)$$

Der Ausdruck $(\mathbf{E}[\hat{\vartheta}] - \vartheta)$ wird auch als Bias oder Verzerrung bezeichnet und verschwindet für einen erwartungstreuen Schätzer. Der MSE eignet sich also, um verschiedene Schätzer hinsichtlich der ersten zwei Punkte in obiger Aufzählung zu vergleichen. Hierbei ist es möglich, einen mit einem systematischen Schätzfehler behafteten Schätzer als günstiger zu bewerten, wenn dessen Varianz nur klein genug ausfällt. Vor allem dann, wenn in der praktischen Anwendung keine Wiederholungen des Experimentes durchgeführt werden. Um die Ergebnisse eines solchen Vergleichs besser interpretieren zu können, wird meist die Quadratwurzel des MSE verwendet. Der RMS-Fehler nach Gleichung 2.4 und 2.5 besitzt dann die gleichen physikalischen Einheiten wie die zu schätzende Größe.

$$RMS[\hat{\vartheta}, \vartheta] = \sqrt{\mathbf{E}[(\hat{\vartheta} - \vartheta)^2]} \quad (2.4)$$

$$= \sqrt{\mathbf{V}[\hat{\vartheta}] + (\mathbf{E}[\hat{\vartheta}] - \vartheta)^2}. \quad (2.5)$$

Um für ein angenommenes Verteilungsmodell einen Schätzer zu bestimmen, existieren mehrere Methoden. Die wichtigsten sind die Momenten- und die Maximum-Likelihood-Methode. Anschaulich können die Schätzer auch durch eine Auswertung in einem verteilungsspezifischen Wahrscheinlichkeitsnetz grafisch bzw. quasigrafisch bestimmt werden.

2.1.3.1 Momentenmethode

Zur Bestimmung der k Parameter des angenommenen Verteilungsmodells nach der Momentenmethode werden die Momente der angenommenen Verteilung bis zur k -ten Ordnung mit den jeweiligen Momenten der Stichprobe gleichgesetzt, wobei das k -te Moment der Stichprobe aus n Beobachtungen als

$$m_k = \frac{1}{n} \sum_{i=1}^n x_i^k$$

definiert ist.⁴ Besitzt das angenommene Verteilungsmodell mehrere Parameter, entsteht ein Gleichungssystem, welches gelöst werden muss. Die Momentenschätzer sind in einigen Fällen, z. B. für die Varianz der Normalverteilung, nicht erwartungstreu.

⁴So ist zum Beispiel für die Exponentialverteilung mit dem Verteilungsparameter λ das erste Moment, also der Erwartungswert, gerade

$$m_1 = \mu = \mathbf{E}[X^1] = \frac{1}{\lambda}.$$

Durch Gleichsetzen des empirischen Momentes mit dem ersten Moment der Verteilung folgt der Schätzer für λ zu

$$\hat{\lambda} = \frac{1}{m_1} = \frac{n}{\sum_{i=1}^n x_i}.$$

2.1.3.2 Maximum-Likelihood-Methode

Der Maximum-Likelihood-Methode liegt die simple Annahme zugrunde, dass das eingetretene Ergebnis eines Zufallsexperiments mehr oder weniger das wahrscheinlichste Ergebnis ist. Der zu schätzende Parameter sollte also so gewählt werden, dass die Wahrscheinlichkeitsfunktion bzw. Wahrscheinlichkeitsdichte der angenommenen Verteilung für den konkreten Wert der Beobachtung gerade ihr Maximum besitzt. Bei mehreren Beobachtungen sollte dann auch das Produkt aller Wahrscheinlichkeitsfunktionen bzw. -dichten für alle Beobachtungen ein Maximum annehmen. Dieses Produkt wird als Likelihoodfunktion $L(\vartheta)$ bezeichnet

$$L = L(\vartheta) = f(x_1, \vartheta) \cdot f(x_2, \vartheta) \cdot \dots \cdot f(x_n, \vartheta). \quad (2.6)$$

Sie ist, wenn die bekannten Werte der Beobachtungen eingesetzt werden, nur von dem zu schätzenden Parameter abhängig. Der wahrscheinlichste Parameter ist nun gerade der, für den diese Funktion ihr Maximum annimmt, d. h. der Schätzwert des unbekanntes Parameters wird über die notwendige Bedingung

$$\left. \frac{\partial L(\vartheta)}{\partial \vartheta} \right|_{\vartheta=\hat{\vartheta}} = 0$$

bestimmt. Durch eine Logarithmierung der Likelihoodfunktion kann die Rechnung vereinfacht werden, da dann das Produkt der Wahrscheinlichkeitsdichten aus Gleichung 2.6 in eine Summe überführt wird, ohne dass sich die Position des Maximums verschiebt. Der Schätzer $\hat{\vartheta}$ ist eine Funktion der Beobachtungen in der Stichprobe; die entsprechende zugehörige Schätzfunktion Θ wird als Maximum-Likelihood-Schätzfunktion bezeichnet

$$\begin{aligned} \hat{\vartheta} &= g(x_1, x_2, \dots, x_n), \\ \Theta &= g(X_1, X_2, \dots, X_n). \end{aligned}$$

2.1.3.3 Grafische und quasigrafische Schätzung

Eine andere Möglichkeit der Parameterbestimmung besteht in der Regression der Verteilungsfunktion. Dazu werden aus den Beobachtungen x_i eines Zufallsexperiments Werte der empirischen Verteilungsfunktion $F(x_i)$ ermittelt. Bei großen Stichprobenumfängen geschieht dies zweckmäßig durch Aufsummierung der relativen Häufigkeiten der klassierten Merkmale, siehe [17]. Werden diese über den Merkmalswerten als Punktwolke grafisch dargestellt, folgen sie im Allgemeinen einem S-förmigen (sigmodalen) Kurvenverlauf. Durch Verzerrung der Ordinate⁵ sowie der Abszissenachse kann dieser Verlauf linearisiert werden. Dann ist es möglich, mit einer Regressionsrechnung oder mit der Methode des „scharfen Hinsehens“ eine Gerade durch die Datenpunkte zu legen. Schätzwerte für den Median und weitere Quantile der Verteilung können dann abgelesen werden.

Für kleine Stichprobenumfänge ist diese Methode mit sehr großen Fehlern behaftet. Hier bietet es sich an, die Stichprobe nach den Merkmalswerten zu sortieren und für jede Rangposition

⁵Den als Wahrscheinlichkeiten interpretierten relativen Häufigkeiten werden die entsprechenden Quantile der angenommenen Verteilung zugeordnet. Diese werden dann linear aufgetragen.

das zu erwartende Quantil der angenommenen Verteilung zu ermitteln.⁶ Diese als „Normal Ranks“ bezeichneten Werte sind für normalverteilte Zufallsgrößen in [12], in „Tabelle XX“ zu finden. Zur Darstellung im Wahrscheinlichkeitsnetz werden anstelle der berechneten Quantile die zugeordneten Wahrscheinlichkeiten verwendet. Ohne Rechentechnik ist sowohl die Bestimmung der Quantile als auch der zugeordneten Wahrscheinlichkeiten sehr aufwendig. Nach [20] können die zugeordneten Wahrscheinlichkeiten $P_{i,n}$ des i -kleinsten Wertes von n Werten aber mit der Näherungsformel

$$P_{i,n} = \frac{i - \alpha}{n - 2\alpha + 1}$$

bestimmt werden. Der Wert α ist hierbei selbst von i und n abhängig. Für ein geeignetes α gibt es für normalverteilte Zufallsgröße eine ganze Reihe von Vorschlägen:

$$\alpha = 0 : \quad P_{i,n} = \frac{i}{n + 1} \quad \text{nach WEIBULL oder GUMBEL} \quad (2.7)$$

$$\alpha = \frac{1}{3} : \quad P_{i,n} = \frac{3i - 1}{3n + 1} \quad \text{nach ROSSOW oder TUKEY} \quad (2.8)$$

$$\alpha = 0,375 : \quad P_{i,n} = \frac{i - 0,375}{n + 0,25} \quad \text{nach BLOM} \quad (2.9)$$

$$\alpha = 0,5 : \quad P_{i,n} = \frac{i - 0,5}{n} \quad \text{nach SCHMIDT oder LOHSE} \quad (2.10)$$

$$\alpha = 0,535206 : \quad P_{i,n} = \frac{i - 0,535206}{n - 0,070413} \quad \text{nach HÜCK} \quad (2.11)$$

Aus einer durch die im Wahrscheinlichkeitsnetz aufgetragenen Punkte führenden Regressionsgeraden können dann die den Quantilen zugehörigen Merkmalswerte ermittelt werden. So ist beispielsweise der geschätzte Median der Verteilung jener Merkmalswert, für den die Regressionsgerade den Funktionswert von $P = 0,5$ bzw. $u = 0$ annimmt. Aus der Differenz zwischen den Merkmalswerten für $P = 0,5$ bzw. $u = 0$ und $P = 0,8413$ bzw. $u = 1$ folgt die Standardabweichung.

Auf die Frage der Qualität der so ermittelten Schätzer für Varianz und Erwartungswert der Grundgesamtheit wird in [20] eingegangen. Es wird festgestellt, dass der geschätzte Median und damit im Falle der Normalverteilung auch der Erwartungswert ein erwartungstreuer Schätzer bei allen verwendeten Verfahren ist. Schätzt man aus dem Wahrscheinlichkeitsnetz als Verteilungsparameter die Standardabweichung, so ist diese nach [20] für die „Normal Ranks“ erwartungstreu. Mit der Näherungsformel nach ROSSOW (2.8) und der Schätzformel nach BLOM (2.9) ist diese Schätzung systematisch, aber vernachlässigbar, fehlerbehaftet. Die Schätzformeln nach WEIBULL oder GUMBEL, SCHMIDT oder LOHSE sowie nach HÜCK sind lediglich asymptotisch erwartungstreu. Während mit der erstgenannten Schätzformel (2.7) die Standardabweichungen systematisch überschätzt werden, liefern die beiden letzteren Formeln (2.10) und (2.11) im Mittel zu kleine Schätzungen der Standardabweichung.

Ebenso wie für die Normalverteilung können auch für die logistische Verteilung oder die Weibullverteilung Wahrscheinlichkeitsnetze konstruiert werden. Für die logistische Verteilung

⁶Wird das Ziehen der Stichprobe nur oft genug wiederholt, so wird der jeweils i -te Wert der sortierten Stichprobe im Mittel jeweils einen Wert annehmen, der einem Quantil der Grundgesamtheit zugeordnet werden kann.

werden dazu für vorgegebene Wahrscheinlichkeiten zwischen 0 und 1 die Quantile nach Gleichung 2.22 (mit $\alpha = 0$ und $\beta = 0$) berechnet und linear über der Ordinate aufgetragen. Die Skalierung der Abszissenachse entsteht aus den linear aufgetragenen möglichen Merkmalswerten. Natürlich ist es möglich, auch andere Skalierungen zu wählen, z. B. eine logarithmische Skalierung, um eine logarithmische logistische Verteilung abzubilden. Ist der Stichprobenumfang ausreichend groß, können die ermittelten Punkte der empirischen Verteilungsfunktion eingetragen werden. Bei kleinem Stichprobenumfang können die einzutragenden Prozentpunkte nach Gleichung 2.12, vgl. [22] berechnet werden.

$$P_i = \frac{i - 0,61}{n + 0,22} \tag{2.12}$$

Für die Weibullverteilung ist sowohl die Konstruktion des Wahrscheinlichkeitsnetzes als auch das Ablesen der Parameter etwas komplizierter, aber in der Literatur erklärt, z. B. in [43], Blatt 5, S. 16 ff. Für kleine Stichprobenumfänge können die Prozentpunkte nach folgender Gleichung bestimmt und im Wahrscheinlichkeitsnetz der Weibullverteilung dargestellt werden:

$$P_i = \frac{i - 0,3}{n + 0,4} \tag{2.13}$$

2.1.4 Konfidenzintervalle für Verteilungsparameter

Eine andere Möglichkeit, um eine statistische Aussage über den wahren aber unbekanntem Parameter einer Verteilung zu treffen, ist die Angabe eines Intervalls, das diesen Parameter mit einer Vertrauenswahrscheinlichkeit $1 - \alpha$ überdeckt. Die Vertrauenswahrscheinlichkeit gibt dabei den relativen Anteil jener in gleicher Art und Weise bestimmten Intervalle an, die bei erneuter Durchführung des Experiments den wahren Parameter enthalten werden.⁷ Die Länge eines solchen Konfidenzintervalls nimmt mit wachsender Anzahl an Beobachtungen ab. Ebenso kann durch Verringerung der Vertrauenswahrscheinlichkeit bzw. Erhöhung der Irrtumswahrscheinlichkeit eine Verkleinerung des Intervalls erreicht werden, allerdings zu Lasten der statistischen Sicherheit der Aussage. Wird die Größe des Konfidenzintervalls mit einer bestimmten Vertrauenswahrscheinlichkeit vorgegeben, kann umgekehrt auch die Anzahl der Beobachtungen bestimmt werden, die nötig sind, um ein Ergebnis einer bestimmten Güte zu erhalten.⁸ Erläuterungen zur Konstruktion von Konfidenzintervallen sind in [17] oder [33] zu finden.

⁷Daher wird dieses Intervall allgemein als Konfidenzintervall zum Niveau $1 - \alpha$ oder als Vertrauensbereich zur Irrtumswahrscheinlichkeit α bezeichnet.

⁸Soll etwa die Varianz einer als normalverteilt angenommenen metrischen Größe bestimmt werden, so ist bekannt, dass gilt:

$$\frac{(n - 1)s^2}{\sigma^2} = z \sim \chi_{n-1}^2$$

Wird nun gefordert, dass die geschätzte Varianz mit einer Vertrauenswahrscheinlichkeit von $1 - \alpha = 0,95$ um nicht mehr als 10% zu niedrig ermittelt wird, lässt sich die Anzahl der dafür nötigen Beobachtungen iterativ aus der Gleichung

$$P\left(\chi_{n-1}^2 < \frac{0,9 \cdot (n - 1)\sigma^2}{\sigma^2}\right) = 0,05$$

berechnen. Man benötigt etwa $n = 520$ Beobachtungen, um den Parameter σ^2 mit der geforderten Genauigkeit zu schätzen.

2.1.5 Anpassungs- und Ausreißertests

Sind die Parameter für eine angenommene Verteilung geschätzt, kann mithilfe eines Anpassungstests geprüft werden, ob sich die Stichprobe damit plausibel beschreiben lässt. Hierzu wird eine Prüfgröße definiert und die Nullhypothese aufgestellt, dass die Stichprobe einer entsprechenden Verteilung entnommen wurde. Die Prüfgröße selbst ist eine Zufallsgröße und nimmt für jede Stichprobe einen anderen Wert an. Anhand der Verteilungsfunktion der Prüfgröße kann festgelegt werden, welche Werte der Prüfgröße als unwahrscheinlich beim Vorliegen der entsprechenden Verteilung gelten. Diese Werte werden als Ablehnungsbereich bezeichnet und müssen vor Anwendung des Tests festgelegt werden. Wenn die Prüfgröße einen Wert aus diesem Bereich annimmt, wird die Nullhypothese verworfen. Es ist dabei zu bedenken, dass lediglich die Größe des Ablehnungsbereiches bekannt ist und daher nur abgeschätzt werden kann, wie groß die Sicherheit gegen das Verwerfen der Nullhypothese bei zutreffender Verteilung ist (Fehler 1. Art). Nicht bekannt ist der Fehler, mit dem die Nullhypothese nicht abgelehnt wird, obwohl die Stichprobe einer anderen Verteilung entnommen ist (Fehler 2. Art). Das Vorliegen einer bestimmten Verteilung mit einer festgelegten Sicherheit nachzuweisen, ist also nicht möglich.

Es existieren eine Vielzahl von Anpassungstests, die wohl wichtigsten sind der KOLMOGOROV-SMIRNOV-Test und der χ^2 -Anpassungstest, die ihre Popularität ihrer einfachen Handhabbarkeit verdanken und in nahezu allen Lehrbüchern erklärt werden. Der SHAPIRO-WILK-Test wird gerne bei Normalverteilungsanpassungen mit geringem Stichprobenumfang verwendet. Für ihn gibt es eine Reihe von Modifikationen, die es auch erlauben, Tests für sehr große Stichprobenumfänge durchzuführen. Von ROYSTON wird in [38] recht ausführlich die Berechnung der Teststatistik beschrieben, sodass diese auch problemlos nachprogrammiert werden kann.

Sowohl von HARTUNG [16], S. 190 als auch in der Norm DIN 53804 [7], Teil 1, S. 7 wird die Auswertung einer Stichprobe im Wahrscheinlichkeitsnetz als Alternative aufgeführt, um die Gültigkeit eines angenommenen Verteilungsmodells zu überprüfen. Nach HÜCK kann die Frage nach der Gültigkeit des angenommenen Verteilungsmodells bei einer Auswertung im Wahrscheinlichkeitsnetz nicht beantwortet werden. Dies wird in [22] S. 43 gezeigt, indem 300 Dauerfestigkeitsversuche, verteilt auf 10 Belastungsstufen, unter der Annahme einer Normal-, Logit- und Arc sin \sqrt{P} -Verteilung, im Wahrscheinlichkeitsnetz ausgewertet werden, ohne dass eine der Verteilungen bevorzugt oder abgelehnt werden kann. Wie auch bei den Anpassungstests wird bei einer Auswertung im Wahrscheinlichkeitsnetz die Nullhypothese aufgestellt, dass die Stichprobe einer bestimmten Verteilung entstammt und sich gerade deshalb im Wahrscheinlichkeitsnetz durch eine Gerade abbilden lässt. Während der Ablehnungsbereich bei einem Anpassungstest fest vorgegeben wird, liegt es bei einer Auswertung im Wahrscheinlichkeitsnetz allein an der subjektiven Einschätzung des Betrachters, ob die Nullhypothese abgelehnt wird.

Neben der Frage, ob sich eine Stichprobe in ihrer Gesamtheit durch eine bestimmte Verteilung darstellen lässt, ist auch die Frage, ob sich die in einer Stichprobe vorhandenen größten und kleinsten Werte noch durch die angenommene Verteilung beschreiben lassen in der praktischen Arbeit bedeutsam. Prinzipiell wird dabei nur die Frage gestellt, ob die vorkommenden kleinsten bzw. größten Werte einer Stichprobe bei Betrachtung der restlichen Stichprobenwerte plausibel erscheinen. Dazu wird wieder eine Prüfgröße definiert und die Hypothese aufgestellt, dass

es sich bei den extremen Stichprobenwerten nicht um Ausreißer handelt. Liegt der Wert der ermittelten Prüfgröße im Ablehnungsbereich, wird diese Hypothese verworfen und der Extremalwert als Ausreißer betrachtet.

Wird ein Wert als Ausreißer identifiziert, sollte versucht werden, einen Grund für dessen Außergewöhnlichkeit zu finden. Kann kein solcher Grund gefunden werden, ist entweder die Verteilungsannahme zu überprüfen, oder es ist davon auszugehen, dass der Ausreißer irrtümlich als solcher identifiziert wurde. Wenn mit Sicherheit bekannt ist, dass alle Stichprobenwerte der gleichen Grundgesamtheit entstammen, führt ein Ausreißertest immer zu Erkennungsraten von Ausreißern, die mit der Irrtumswahrscheinlichkeit übereinstimmen. Gerade bei Simulationsversuchen führt das zu Fehleinschätzungen, wie von HÜCK in [20], S. 27 gezeigt wird. Verschiedene Testverfahren sind in [16], [29] beschrieben oder in Normen, DIN 53804-1, [7], dokumentiert.

2.1.6 Einige Verteilungsmodelle und Parameterschätzungen

Die Modelle zur Beschreibung von Zufallsgrößen sind zahlreich. Prinzipiell ist es möglich, jede Funktion mit einem sigmodialen Kurvenverlauf zur Darstellung einer Verteilungsfunktion zu verwenden, solange die empirische Verteilungsfunktion damit wiedergegeben werden kann. Einen Überblick über die wichtigsten Verteilungen ist in jedem Statistiklehrbuch zu finden. Eine grundsätzlich sehr umfassende Zusammenstellung bietet auch ein im weltweiten Rechnernetzwerk zu findendes, von einer gemeinnützigen Organisation und Freiwilligen betriebenes, kostenfreies Nachschlagewerk.

2.1.6.1 Binomial-Verteilung $\mathcal{Bi}(n; p)$

Ein Zufallsexperiment mit nur zwei möglichen, sich gegenseitig ausschließenden Ergebnissen, nämlich Erfolg oder Misserfolg mit konstanter Erfolgswahrscheinlichkeit p bzw. Misserfolgswahrscheinlichkeit $q = 1 - p$, wird als Bernoulli-Experiment bezeichnet. Wird es mehrmalig durchgeführt, kann die Anzahl der Erfolge X bei n -maliger Wiederholung des Experiments als diskrete Zufallsgröße aufgefasst und mit der Binomialverteilung beschrieben werden. Die Wahrscheinlichkeit, dass eine binomialverteilte Zufallsgröße $X \sim \mathcal{Bi}(n; p)$ genau einen Wert x annimmt ist

$$P(X = x) = f(x) = \binom{n}{x} p^x (1 - p)^{n-x}.$$

Die Verteilungsfunktion der Binomialverteilung gibt die Wahrscheinlichkeit an, mit der die Zufallsgröße einen Wert kleiner oder gleich dem Wert x annimmt und lautet entsprechend

$$P(X \leq x) = F(x) = \sum_{k=0}^x \binom{n}{k} p^k (1 - p)^{n-k}. \quad (2.14)$$

Eine binomialverteilte Zufallsgröße $X \sim \mathcal{Bi}(n; p)$ besitzt sowohl einen Erwartungswert $\mathbf{E}[X]$ als auch eine Varianz $\mathbf{V}[X]$. Beide Größen lassen sich leicht aus den Verteilungsparametern

$$\mathbf{E}[X] = n \cdot p \quad \text{und} \quad \mathbf{V}[X] = n \cdot p \cdot q$$

berechnen. Nach dem zentralen Grenzwertsatz von de Moivre-Laplace konvergiert die Binomialverteilung für $n \rightarrow \infty$ gegen die Normalverteilung

$$\lim_{n \rightarrow \infty} P(X \leq x) = \Phi \left(\frac{x - np}{\sqrt{npq}} \right).$$

In der praktischen Anwendung wird diese Annäherung in ausreichendem Maß erreicht, wenn die Bedingung $npq \geq 9$ erfüllt wird. Um dabei Fehler aus dem Übergang von einer diskreten zu einer stetigen Zufallsgröße zu kompensieren, wird eine Stetigkeitskorrektur eingeführt, vgl. [16], und es gilt dann

$$P(X \leq x) = F(x) \approx \Phi \left(\frac{x - np + 0,5}{\sqrt{npq}} \right) - \Phi \left(\frac{x - np - 0,5}{\sqrt{npq}} \right).$$

Ein Schätzer⁹ \hat{p} für die unbekannte Ausfallwahrscheinlichkeit p ist die relative Häufigkeit

$$\hat{p} = \frac{x}{n}. \quad (2.15)$$

In [16] wird die Berechnung des Konfidenzintervalls $[p_u; p_o]$ zum Niveau $1 - \alpha$ für den wahren Parameter p vorgeführt. Demnach kann das Intervall, wenn die Binomialverteilung durch die Normalverteilung approximiert werden kann ($npq \geq 9$), durch

$$p_u = \frac{2x + u_{1-\alpha/2}^2 - u_{1-\alpha/2} \sqrt{u_{1-\alpha/2}^2 + 4x \left(1 - \frac{x}{n}\right)}}{2 \left(n + u_{1-\alpha/2}^2\right)},$$

$$p_o = \frac{2x + u_{1-\alpha/2}^2 + u_{1-\alpha/2} \sqrt{u_{1-\alpha/2}^2 + 4x \left(1 - \frac{x}{n}\right)}}{2 \left(n + u_{1-\alpha/2}^2\right)}$$

ermittelt werden.¹⁰ Ist eine Approximation durch eine Normalverteilung nicht möglich, lässt sich das Konfidenzintervall durch

$$p_u = \frac{x F_{2x; 2(n-x+1); \alpha/2}}{n - x + 1 + x F_{2x; 2(n-x+1); \alpha/2}},$$

$$p_o = \frac{(x + 1) F_{2(x+1); 2(n-x); 1-\alpha/2}}{n - x + (x + 1) F_{2(x+1); 2(n-x); 1-\alpha/2}}$$

exakt angeben.¹¹ Die Werte p_u und p_o werden auch als Person-Clopper-Werte bezeichnet. Die obere Intervallgrenze p_o ist hierbei anschaulich der Wert der Erfolgswahrscheinlichkeit, mit dem die Wahrscheinlichkeit in n Versuchen x oder weniger Erfolge zu erzielen gleich der halben Irrtumswahrscheinlichkeit ist; also der Wert, für den die Verteilungsfunktion (Gleichung 2.14) den Wert $\alpha/2$ annimmt. Die untere Intervallgrenze ist hingegen gerade der Wert, für den die Wahrscheinlichkeit in n Versuchen $n - x$ oder weniger Misserfolge zu erzielen gleich der halben Irrtumswahrscheinlichkeit ist.

⁹Der Schätzer kann z. B. nach der Maximum-Likelihood-Methode ermittelt werden. Die Likelihood-Funktion ist: $L(p) = \binom{n}{x} p^x (1-p)^{(n-x)}$. Die notwendige Bedingung für die partielle Ableitung der logarithmierten Likelihood-Funktion nach dem Parameter p lautet: $\frac{\partial \ln(L(p))}{\partial p} = \frac{x}{p} - \frac{n-x}{1-p} \stackrel{!}{=} 0$. Die Auflösung nach p ergibt den Schätzer \hat{p} .

¹⁰Mit $u_{1-\alpha/2}$ wird das Quantil der Standardnormalverteilung $\Phi(u) = 1 - \alpha/2$ beschrieben.

¹¹Hierbei sind $F_{m_1; m_2; \gamma}$ die Quantile zur Wahrscheinlichkeit γ einer mit den Freiheitsgraden m_1 (im Zähler) und m_2 (im Nenner) F-verteilter Zufallsgröße.

2.1.6.2 Hypergeometrische Verteilung

Werden bei dem Experiment des Ziehens von Kugeln aus einer Urne die gezogenen Kugeln nicht wieder zurückgelegt, verändert sich mit jedem Zug die Erfolgswahrscheinlichkeit. Die Wahrscheinlichkeit bei n Versuchen genau x mal erfolgreich zu sein, wenn sich in der Urne genau M weiße von insgesamt N Kugeln befinden, wird durch

$$P(X = x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

berechnet. Entsprechend lautet die Verteilungsfunktion einer hypergeometrisch verteilten Zufallsgröße X

$$P(X \leq x) = \sum_{m=0}^x \frac{\binom{M}{m} \binom{N-M}{n-m}}{\binom{N}{n}}.$$

Der Erwartungswert $\mathbf{E}[X]$ und die Varianz $\mathbf{V}[X]$ werden aus den Parametern nach

$$\mathbf{E}[X] = n \frac{M}{N} \quad \text{und} \quad \mathbf{V}[X] = n \frac{N}{M} \left(1 - \frac{N}{M}\right) \frac{N-n}{N-1}$$

bestimmt. Wird hier der Anteilswert N/M durch den Parameter p dargestellt, ergibt sich im Vergleich zur Binomialverteilung ein identischer Erwartungswert. Die Varianzen der Verteilungen unterscheiden sich um einen als Endlichkeitskorrektur bezeichneten Faktor $(N-n)/(N-1)$. Ist der Umfang der gezogenen Stichprobe deutlich kleiner als der Umfang N der Grundgesamtheit, d. h. $n/N \leq 0,05$, kann eine hypergeometrisch verteilte Zufallsgröße durch eine Binomialverteilung approximiert werden. Der Schätzer für den unbekanntem Anteilswert $\frac{M}{N} = p$ ist direkt aus den Beobachtungsergebnis mit $\hat{p} = \frac{x}{n}$ bestimmbar.

2.1.6.3 Normalverteilung $\mathcal{N}(\mu; \sigma^2)$

Das bekannteste Verteilungsmodell zur Beschreibung stetiger, metrischer Zufallsgrößen ist die Normalverteilung. Sie ist geeignet, die Verteilung einer Zufallsgröße zu beschreiben, die sich aus der Summe vieler weiterer, voneinander unabhängigen, zufälligen Größen zusammensetzt.¹² So kann zum Beispiel der Mittelwert aus mehreren ($n > 30$) unabhängigen Messungen einer physikalischen Größe meist als eine normalverteilte Zufallsgröße aufgefasst werden, für die dann Erwartungswert und Streuung angegeben bzw. bestimmt werden können. Die Wahrscheinlichkeitsdichte der Normalverteilung wird durch

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

¹²Diese Erkenntnis ist Gegenstand des zentralen Grenzwertsatzes, siehe [16] oder überall sonst. Der Mittelwert von n voneinander unabhängigen, gleichverteilten Zufallsgrößen mit endlichem Erwartungswert und endlicher Varianz ist selbst eine Zufallsgröße, deren Verteilungsfunktion sich für $n \rightarrow \infty$ durch eine Normalverteilung beschreiben lässt.

beschrieben. Die Verteilungsfunktion lautet

$$F(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\tilde{x}-\mu)^2}{2\sigma^2}} d\tilde{x}.$$

Die in beiden Gleichungen verwendeten Parameter beschreiben den Erwartungswert $\mathbf{E}[X] = \mu$ und die Varianz $\mathbf{V}[X] = \sigma^2$ der Zufallsgröße X . Während der Erwartungswert jenen Wert beschreibt, den die Zufallsgröße im Mittel annimmt, wird durch die Varianz die mittlere quadratische Abweichung der Zufallsgröße vom Erwartungswert beschrieben. Wird die Zufallsgröße mithilfe dieser beiden Parameter normiert, vgl. Gleichung 2.16, dann besitzt die daraus entstehende neue Zufallsgröße einen Erwartungswert von $\mathbf{E}[Z] = 0$ und eine Varianz von gerade $\mathbf{V}[Z] = 1$.

$$Z = \frac{X - \mu}{\sqrt{\sigma^2}} \quad (2.16)$$

Die Verteilung dieser Zufallsgröße Z wird als Standardnormalverteilung bezeichnet, deren tabellierte Verteilungsfunktion

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{z}^2}{2}} d\tilde{z}$$

in nahezu allen analogen und digitalen Nachschlagewerken verfügbar ist. Mit ihrer Hilfe ist es möglich, unter Nutzung der Beziehung 2.16 für jede beliebige normalverteilte Zufallsgröße den Wert der nicht elementar berechenbaren Verteilungsfunktion bzw. deren Umkehrung einfach zu bestimmen.

Die Parameter μ und σ^2 der Normalverteilung werden durch die erwartungstreuen Schätzer m und s^2 mit

$$m = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (2.17)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.18)$$

ermittelt. Während der Schätzer des Erwartungswertes mit dem Stichprobenmittelwert gleichgesetzt werden kann, muss die Stichprobenvarianz immer dann, wenn der Erwartungswert unbekannt ist, mit dem Faktor $n/(n-1)$ korrigiert werden. Hierbei gilt es zu beachten, dass der Schätzer s^2 zwar die Varianz, nicht aber deren Wurzel, die Standardabweichung im Mittel erwartungstreu schätzt. Sollen die Ergebnisse mehrerer Versuchsreihen zusammengefasst werden, so ist die mittlere Standardabweichung aus $\bar{s} = \frac{1}{m} \sqrt{\sum_{i=1}^m s_i^2}$ zu berechnen. Eine erwartungstreue Schätzung der Standardabweichung kann durch den Schätzer

$$\tilde{s} = \sqrt{\frac{n-1}{2} \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})}} \cdot s$$

erreicht werden. Auch bei der Auswertung im Wahrscheinlichkeitsnetz ist prinzipiell die Möglichkeit gegeben, erwartungstreue Schätzer der Standardabweichung zu erhalten, siehe Abschnitt 2.1.3.3.

Für die beiden nach Gleichung 2.17 und 2.18 geschätzten Parameter der Normalverteilung können Konfidenzintervalle angegeben werden. Die Intervalle des jeweiligen Parameters fallen dabei grundsätzlich kleiner aus, wenn der andere Parameter bereits bekannt ist. Der Vertrauensbereich des Erwartungswertes bei bekannter Varianz σ zu einer Irrtumswahrscheinlichkeit von α ist

$$\left[m - u_{1-\alpha/2} \cdot \sqrt{\frac{\sigma^2}{n}}; \quad m - u_{\alpha/2} \cdot \sqrt{\frac{\sigma^2}{n}} \right].$$

Für den Vertrauensbereich des Erwartungswertes bei unbekannter Varianz wird anstelle des Quantils u der Standardnormalverteilung das Quantil $t_{k,\alpha}$ der t-Verteilung mit dem Freiheitsgrad $k = n - 1$ verwendet, vgl. Abschnitt 2.1.6.8. Das Konfidenzintervall lautet dann

$$\left[m - t_{n-1,1-\alpha/2} \cdot \sqrt{\frac{s^2}{n}}; \quad m - t_{n-1;\alpha/2} \cdot \sqrt{\frac{s^2}{n}} \right].$$

Für die Varianz kann das Konfidenzintervall zum Niveau $1 - \alpha$ bei bekanntem Erwartungswert mithilfe der χ^2 -Verteilung mit dem Freiheitsgrad $k = n$ durch

$$\left[\frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi_{n;1-\alpha/2}^2}; \quad \frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi_{n;\alpha/2}^2} \right]$$

angegeben werden. Ist der Erwartungswert hingegen unbekannt, lautet das Konfidenzintervall

$$\left[\frac{n-1}{\chi_{n-1;1-\alpha/2}^2} \cdot s^2; \quad \frac{n-1}{\chi_{n-1;\alpha/2}^2} \cdot s^2 \right].$$

2.1.6.4 Logarithmische Normalverteilung $\mathcal{N}_L(\mu; \sigma^2)$

Setzt sich eine Zufallsgröße Y aus dem Produkt vieler einzelner, identisch verteilter Zufallsgrößen zusammen, kann diese oft durch eine logarithmische Normalverteilung approximiert werden. Durch Logarithmieren der Zufallsgröße Y wird eine neue Zufallsgröße X gebildet

$$Y \sim \mathcal{N}_L(\mu; \sigma^2) \qquad X = \log_a Y \qquad X \sim \mathcal{N}(\mu_L; \sigma_L^2)$$

und das Produkt wird in eine Summe gewandelt, welche sich nach dem zentralen Grenzwertsatz bei ausreichend großen n durch eine Normalverteilung darstellen lässt. Es sind dann die oben beschriebenen Verteilungsparameter der logarithmierten Zufallsgröße zu bestimmen.

Während die Zufallsgröße Y den Erwartungswert $\mathbf{E}[Y] = \mu$ und die Varianz $\mathbf{V}[Y] = \sigma^2$ besitzt, werden die Parameter der logarithmierten Zufallsgröße $X = \log_a Y$ durch den Erwartungswert $\mathbf{E}[X] = \mu_L$ und die Varianz $\mathbf{V}[X] = \sigma_L^2$ angegeben. Hier ist stets darauf zu achten, welcher Parametersatz tatsächlich angegeben wird und zu welcher Basis a der Logarithmus zu bilden ist. Die ermittelten Parameter lassen sich nach den folgenden Gleichungen ineinander umrechnen:

$$\mu = e^{(\mu_L + \frac{1}{2}\sigma_L^2 \cdot (\ln a)^2)}, \tag{2.19}$$

$$\sigma^2 = e^{(2\mu_L + \sigma_L^2 \cdot (\ln a)^2) \cdot (e^{\sigma_L^2 \cdot (\ln a)^2} - 1)}. \tag{2.20}$$

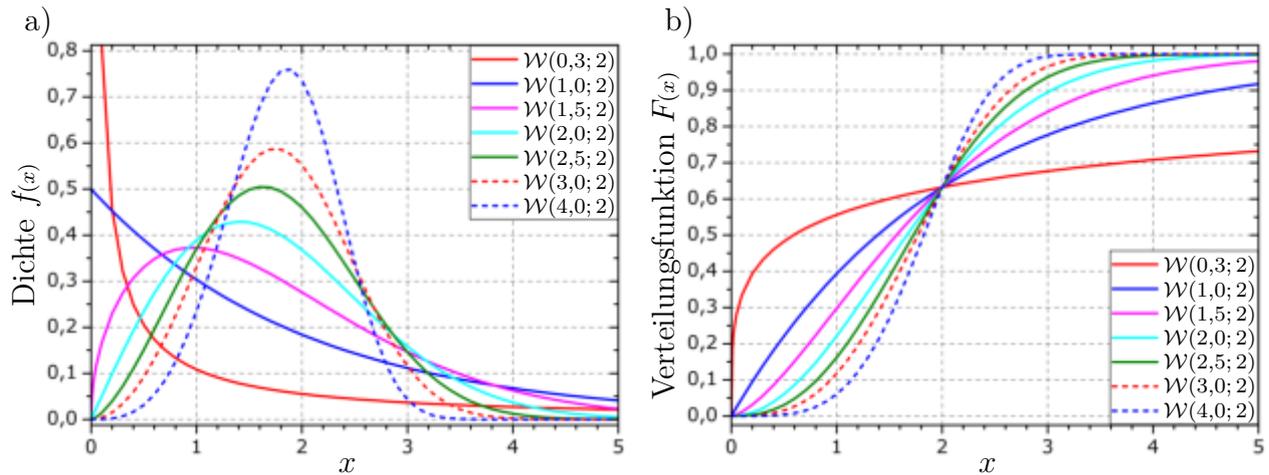


Abb. 2.1: Beispielhafte Darstellung verschiedener Wahrscheinlichkeitsdichten und Verteilungsfunktionen der Weibullverteilung $\mathcal{W}(b;T)$ mit verschiedenen Parametern.

Wird der natürliche Logarithmus verwendet, vereinfacht sich die Umrechnung zu:

$$\begin{aligned} \mu &= e\left(\mu_L + \frac{1}{2}\sigma_L^2\right), \\ \sigma^2 &= e\left(2\mu_L + \sigma_L^2\right) \cdot (e^{\sigma_L^2} - 1). \end{aligned}$$

Eine Potenzierung der Basis a mit dem Erwartungswert μ_L der logarithmierten Zufallsgröße liefert immer den Median der Verteilung.

In der Betriebsfestigkeit wird zur Beschreibung der Streuung gerne die „Streuspanne“ T verwendet. Sie ist als das Verhältnis des 90 %-Merkmalswertes zum 10 %-Merkmalswert definiert und beschreibt damit einem Bereich, in dem 80 % aller Merkmalsausprägungen der Zufallsgröße Y zu erwarten sind. Sie lässt sich aus den geschätzten Verteilungsparametern m_L und s_L^2 der logarithmierten Zufallsgröße recht einfach durch

$$T = \frac{a^{m_L+1,2816s_L}}{a^{m_L+1,2816s_L}} = a^{2,5632s_L} \quad (2.21)$$

angeben. Wenn dieses Streumaß verwendet wird, ist zu beachten, dass sich durch die Mittelung mehrerer Streuspannen nicht zwangsweise ein genauerer Schätzwert für die Streuspanne der Grundgesamtheit ergibt. Je nach Auswertemethode ist nur die Varianz s_L^2 oder nur die Standardabweichung s_L der logarithmierten Zufallsgröße ein erwartungstreuer Schätzer. Es ist also immer erst dessen Mittelwert zu bestimmen, bevor mit Hilfe der Gleichung 2.21 eine mittlere Streuspanne angegeben werden kann.

2.1.6.5 Weibullverteilung $\mathcal{W}(b;T)$

Für die Beschreibung von Lebensdauerproblemen ist die Weibullverteilung ein oft verwendetes Modell. Ihre Popularität verdankt sie vor allem der Tatsache, dass sie je nach Wahl der Verteilungsparameter verschiedenen anderen Verteilungen ähnelt, siehe Abbildung 2.1. Die Verteilungsdichte der Weibullverteilung ist

$$f(x) = \frac{b}{T} \left(\frac{x}{T}\right)^{b-1} e^{-\left(\frac{x}{T}\right)^b}.$$

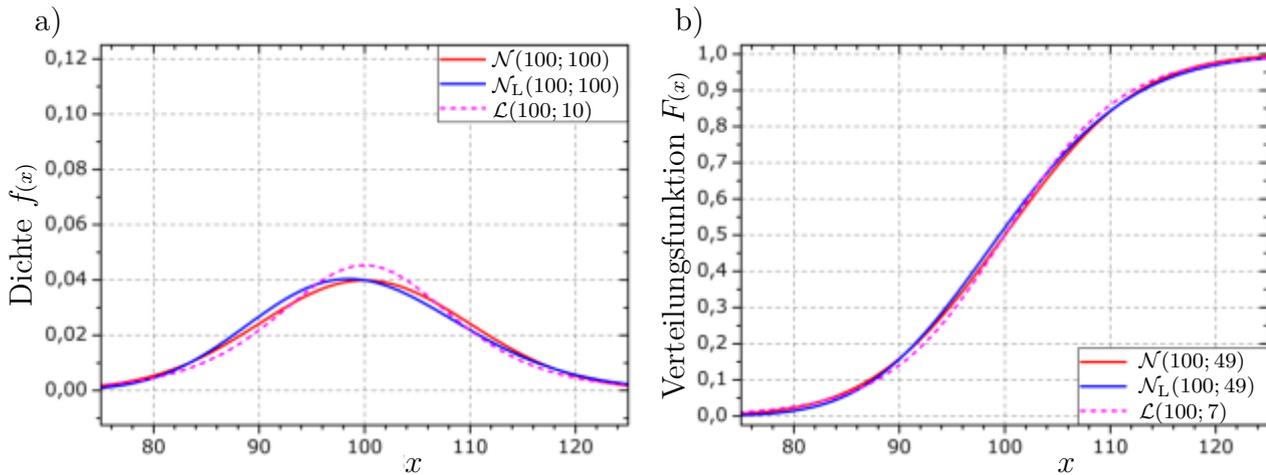


Abb. 2.2: Vergleich der Wahrscheinlichkeitsdichten und Verteilungsfunktionen der Normalverteilung $\mathcal{N}(100; 100)$, logarithmischen Normalverteilung $\mathcal{N}_L(100; 100)$ und der logistischen Verteilung $\mathcal{L}(100; 10)$, jeweils für den Erwartungswert $\mathbf{E}[X] = 100$ und die Varianz $\mathbf{V}[X] = 100$.

Mit der Verteilungsfunktion

$$F(x) = P(X \leq y) = 1 - e^{-\left(\frac{x}{T}\right)^b}$$

kann dann die Wahrscheinlichkeit des Bauteilausfalls bis zum Erreichen einer Lebensdauer x berechnet werden. Wie aus der Abbildung 2.1 zu entnehmen ist, lassen sich durch Variation des Formparameters b verschiedene, auch der Normalverteilung ähnliche, Verteilungen erzeugen. Die Verteilungsfunktion nimmt unabhängig davon stets an der Stelle $x = T$ den Wert $F(x = T) = 0,632$ an. Daher wird der Parameter T als charakteristische Lebensdauer interpretiert. Der Erwartungswert und die Varianz einer weibullverteilten Zufallsgröße X berechnen sich mit der Gammafunktion $\Gamma(\cdot)$ zu

$$\mathbf{E}[X] = T \cdot \Gamma\left(1 + \frac{1}{b}\right) \quad \text{und} \quad \mathbf{V}[X] = T^2 \cdot \left(\Gamma\left(1 + \frac{2}{b}\right) - \Gamma^2\left(1 + \frac{1}{b}\right)\right).$$

Neben der zweiparametrischen Weibullverteilung existiert auch ein Modell mit drei Verteilungsparametern, mit der es möglich ist, eine ausfallfreie Anfangszeit bzw. untere Beanspruchungsgrenze zu berücksichtigen. Im Allgemeinen ist es aber recht aufwendig und unsicher, eine solche Grenze zu bestimmen. Nach MOGWITZ [30], S. 24, ist die dreiparametrische Weibullverteilung besser als eine logarithmische Normalverteilung geeignet, die Lebensdauer von im Zeitfestigkeitsgebiet mittels Wöhlerversuch geprüften Proben zu beschreiben.

Wie die Parameter der Weibullverteilung geschätzt werden können, wird von HARTUNG in [16] S. 791 ff. beschrieben. Wie die Konfidenzintervalle der so geschätzten Parameter berechnet werden, ist dem Autor¹³ nicht bekannt.

2.1.6.6 Logistische Verteilung $\mathcal{L}(\alpha; \beta)$

Obwohl sie nach LIU [26] zur Beschreibung der Dauerfestigkeitsstreuung geeignet ist, wird die logistische Verteilung nur selten verwendet. Die Form der Verteilungsfunktion lässt sich

¹³Also mir!

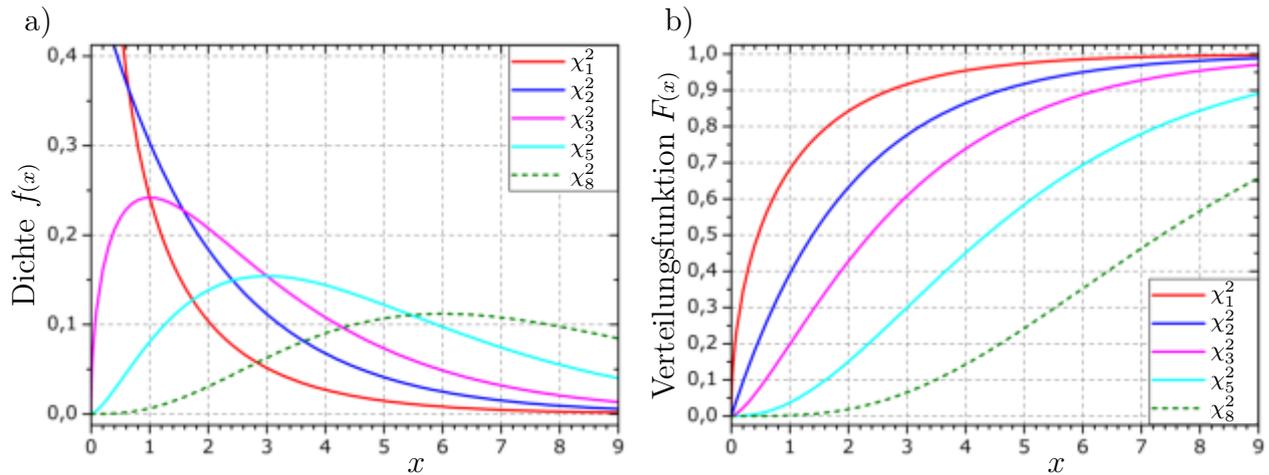


Abb. 2.3: Vergleich der Wahrscheinlichkeitsdichte und Verteilungsfunktion χ^2 -Verteilungen für verschiedene Freiheitsgrade.

direkt aus den Überlegungen zur Ausfallwahrscheinlichkeit einer n -gliedrigen Kette herleiten, siehe [22]. Ihre Wahrscheinlichkeitsdichte wird durch

$$f(x) = \frac{\pi}{\beta\sqrt{3}} \frac{e^{-\frac{\pi}{\sqrt{3}} \frac{x-\alpha}{\beta}}}{\left(1 + e^{-\frac{\pi}{\sqrt{3}} \frac{x-\alpha}{\beta}}\right)^2}$$

angegeben. Die Verteilungsfunktion berechnet sich daraus mit:

$$F(x) = \frac{1}{1 + e^{-\frac{\pi}{\sqrt{3}} \frac{x-\alpha}{\beta}}}.$$

In der hier verwendeten Notation lassen sich Erwartungswert und Varianz der Verteilung leicht angeben. So ist der Erwartungswert der Verteilung gerade $\mathbf{E}[X] = \alpha$ und die Varianz wird durch $\mathbf{V}[X] = \beta^2$ beschrieben. Ein Vorteil der logistischen Verteilung liegt in der einfachen Ermittlung des zu einem Wert der Wahrscheinlichkeitsfunktion $F(x) = P(X \leq x)$ gehörenden Quantils. Dieses ist direkt nach folgender Gleichung 2.22 berechenbar:

$$x = \frac{\sqrt{3}}{\pi} \beta \ln\left(\frac{P}{1-P}\right) + \alpha. \tag{2.22}$$

Die Verteilungsparameter werden aus dem Stichprobenmittelwert und der mit dem Faktor $n/(n-1)$ korrigierten Stichprobenvarianz ermittelt.

2.1.6.7 χ^2 -Verteilung

Die χ^2 -Verteilung mit dem Freiheitsgrad n beschreibt die Summe aus n unabhängigen standardnormalverteilten Zufallsgrößen. Sie wird benötigt, um statistische Tests zu konstruieren, wenn Fehler in Form von quadratischen Abständen, wie es zum Beispiel bei Stichprobenvarianzen der Fall ist, betrachtet werden. So ist z. B. die aus einer n -elementigen Stichprobe einer Normalverteilung $X \sim \mathcal{N}(\mu; \sigma^2)$ ermittelte Größe $\sum(x_i - \mu)^2/\sigma^2$ gerade χ_n^2 -verteilt. Die

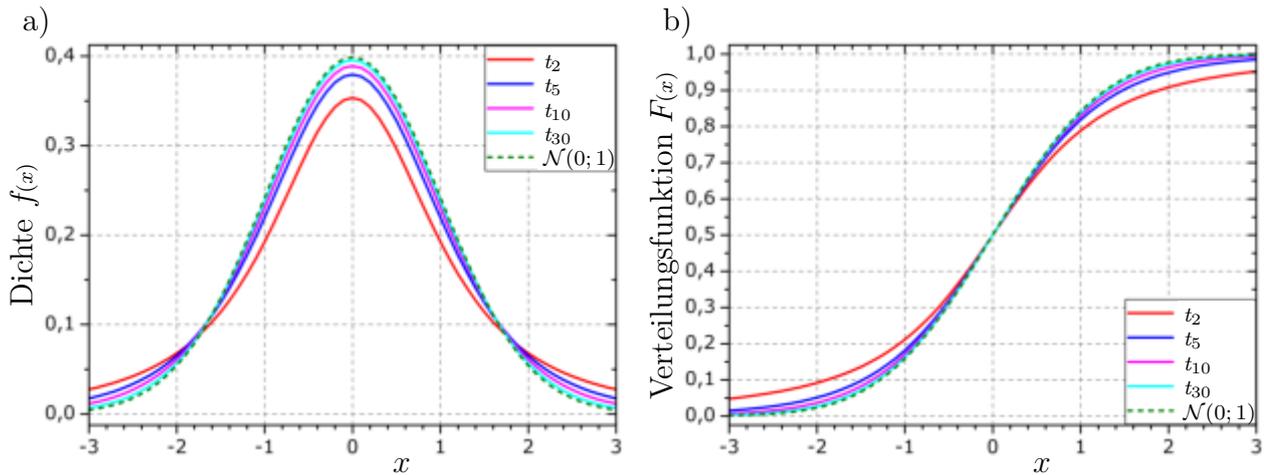


Abb. 2.4: Vergleich der Wahrscheinlichkeitsdichte und Verteilungsfunktion von Standardnormalverteilung und t-Verteilung für verschiedene Freiheitsgrade.

Verteilungsdichte der χ_n^2 -Verteilung wird nach Gleichung 2.23 mithilfe der Gammafunktion $\Gamma(\cdot)$ berechnet.

$$f(x) = \frac{x^{\frac{n-2}{2}} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \quad (2.23)$$

Die Verteilungsfunktion

$$F(x) = \int_{-\infty}^x \frac{\tilde{x}^{\frac{n-2}{2}} e^{-\frac{\tilde{x}}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} d\tilde{x} \quad (2.24)$$

kann nicht elementar dargestellt werden.¹⁴ Die für statistische Tests interessanten Quantile sind aber in nahezu allen Statistiklehrbüchern und Nachschlagewerken tabelliert. Ein Beispiel für die Dichte- und die Verteilungsfunktion einer χ^2 -verteilten Zufallsgröße ist in Abbildung 2.3 dargestellt. Bei einem ausreichend großen Freiheitsgrad ($n \geq 100$) kann die Verteilungsfunktion einer χ^2 -verteilten Zufallsgröße durch eine Normalverteilung approximiert werden.

2.1.6.8 t-Verteilung

Mit Hilfe der t-Verteilung ist es möglich, die Variabilität eines Mittelwertes zu bewerten. Wird der Mittelwert einer Stichprobe vom Umfang n einer $\mathcal{N}(\mu; \sigma)$ verteilten Zufallsgröße selbst als eine Zufallsgröße \bar{X} betrachtet und auf die Parameter der Grundgesamtheit normiert und mit \sqrt{n} multipliziert, so ist diese Größe Z standardnormalverteilt, d.h.:

$$\bar{X} = \sum_{i=1}^n X_i, \quad Z = \sqrt{n} \cdot \frac{\bar{X} - \mu}{\sigma}, \quad Z \sim \mathcal{N}(0; 1). \quad (2.25)$$

Damit ist es möglich, die Verteilung der Mittelwerte zu modellieren. Insbesondere kann so das Konfidenzintervall für einen geschätzten Erwartungswert ermittelt werden, siehe Abschnitt 2.1.4. Wenn allerdings die Mittelwerte nicht auf die Standardabweichung der Grundgesamtheit,

¹⁴Hierzu ist die regularisierte unvollständige Gammafunktion nötig.

sondern auf die jeweilige korrigierte Stichprobenstandardabweichung normiert werden, ist dies nicht möglich. Die Stichprobenstandardabweichung ist hier selbst eine Zufallsgröße S und die normierten Mittelwerte Y sind mit dem Freiheitsgrad $(n - 1)$ t -verteilt.

$$S = \sum_{i=1}^n (X_i - \bar{X})^2, \quad Y = \sqrt{n} \cdot \frac{\bar{X} - \mu}{S}, \quad Y \sim t_{n-1}. \quad (2.26)$$

Hierdurch ist es möglich, die Variabilität eines Mittelwertes zu bewerten, wenn die Streuung der Grundgesamtheit aus der selben Stichprobe abgeschätzt wird, vgl. Abschnitt 2.1.4. Ein Vergleich der Verteilungsdichten und Verteilungsfunktionen in Abbildung 2.4 zeigt, wie sich die Verteilungsfunktionen für kleine Stichprobenumfänge vor allem bei sehr kleinen oder sehr großen Quantilen unterscheiden. Auch zu erkennen ist, dass ab einem Freiheitsgrad von $n > 30$ die t -Verteilung durch eine Normalverteilung dargestellt werden kann. Für die ingenieurtechnische Anwendung sind vor allem die Quantile der t -Verteilung interessant. Diese sind für verschiedene Freiheitsgrade in nahezu allen sich mit Statistik befassenden Büchern vertafelt.

2.1.7 Bemerkung zur Wahl des Verteilungsmodells

Wird aufgrund von getroffenen Beobachtungen ein Modell zur Beschreibung einer Zufallsgröße gewählt, stellt sich vor allem dann, wenn zur Ableitung des Modells nur ein geringer Stichprobenumfang zur Verfügung stand, die Frage nach der Gültigkeit des Modells für die Beschreibung sehr großer oder kleiner Merkmalswerte. Meist ist bei geringem Stichprobenumfang die richtige Wahl des Verteilungsmodells von geringer Bedeutung, denn der Bereich der Verteilungsfunktion, der mit diesen wenigen Versuchen experimentell belegt wird, besitzt bei vielen Verteilungsfunktionen einen ähnlichen Verlauf. Die Auswirkungen unterschiedlicher Verteilungsannahmen zeigen sich meist erst für Quantile u_α sehr kleiner bzw. großer Auftretenswahrscheinlichkeit $\alpha < 0,05$ bzw. $\alpha > 0,95$. In Abbildung 2.5 sind hierzu exemplarisch die Verteilungsfunktionen, der oben beschriebenen Verteilungen, dargestellt. Alle Verteilungen besitzen hierbei einen Erwartungswert von $\mathbf{E}[X] = 100$ und eine Varianz von $\mathbf{V}[X] = 49$. Zusätzlich sind noch die im Mittel zu erwartenden Ergebnisse einer Stichprobe (Normalverteilung) von $n = 10$ Beobachtungen eingezeichnet. Gut zu erkennen ist, dass alle betrachteten Verteilungen sich im Bereich der erwarteten Beobachtungen kaum unterscheiden. Die Unterschiede ergeben sich vielmehr erst, wenn das Auftreten von sehr kleinen oder großen Merkmalswerten modelliert werden soll. Es ist einleuchtend, dass dann aber auch eine viel größere Anzahl an Beobachtungen nötig ist, um den Verlauf der Kurven dort experimentell belegen zu können.

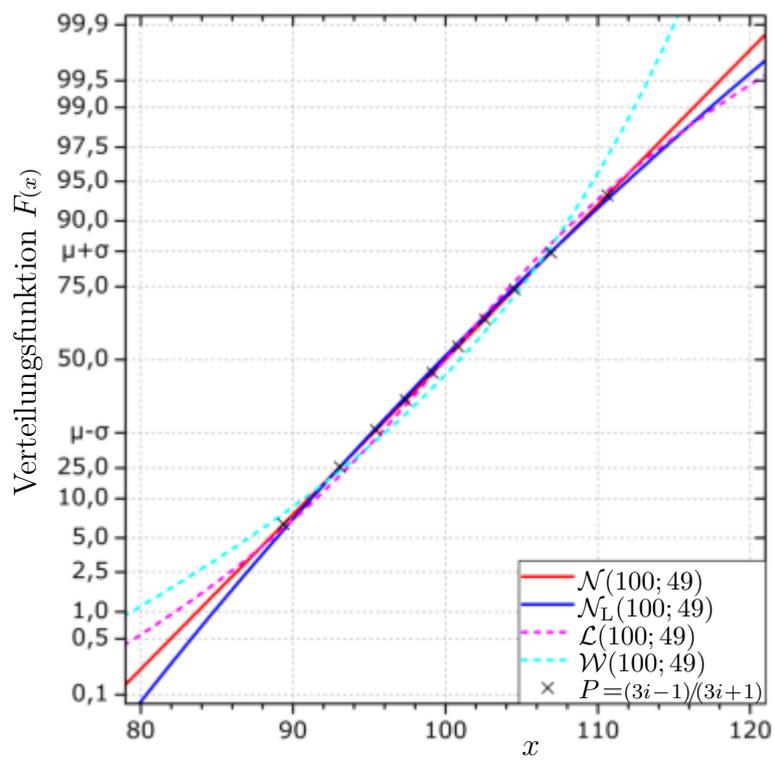


Abb. 2.5: Vergleich verschiedener Verteilungsfunktionen im Wahrscheinlichkeitsnetz der Normalverteilung. Alle Verteilungen besitzen einen Erwartungswert $\mathbf{E}[X] = 100$ und eine Varianz von $\mathbf{V}[X] = 49$.

2.2 Experimentelle Bestimmung der Bauteildauerfestigkeit

Im Bereich sehr großer Lastwechselzahlen und entsprechend kleinen Beanspruchungsamplituden ist die Bestimmung der statistischen Verteilung der ertragbaren Lastwechsel eines Bauteils mit sehr hohem zeitlichen Aufwand verbunden bzw. nicht möglich. Anstelle der statistischen Verteilung der Lastwechsel wird daher die Verteilung der ertragbaren Belastung bestimmt. Da aber experimentell das Versagen bei einer festgelegten Lastwechselzahl nicht bestimmt werden kann, werden Bauteile bis zu einer Grenzlasterwechselzahl N_G geprüft. Je nachdem, ob die Probe vor dem Erreichen der Grenzlasterwechselzahl versagt oder nicht, wird das Versuchsergebnis als Ausfall (Bruch) oder Durchläufer bezeichnet¹⁵. Unter der Annahme unendlich vieler verfügbarer Proben ist dieses Experiment das „Ziehen mit Zurücklegen aus einer Urne“, also ein n -faches BERNOULLI-Experiment. Die Wahrscheinlichkeit auf einem Lasthorizont¹⁶ mit einer Ausfallwahrscheinlichkeit p in n Versuchen genau r Brüche zu erhalten, wird mithilfe der Binomialverteilung beschrieben:

$$P(r, n, p) = \binom{n}{r} \cdot p^r \cdot (1 - p)^{(n-r)}. \quad (2.27)$$

Der Schätzer für die tatsächliche Ausfallwahrscheinlichkeit p eines solchen Horizontes ist die relative Bruchhäufigkeit $\hat{p} = r/n$. Liegen für mehrere Horizonte die jeweiligen Schätzer vor, können durch Regression die Parameter der als bekannt vorausgesetzten Verteilungsfunktion ermittelt werden, praktischerweise durch Auswertung im Wahrscheinlichkeitsnetz [10]. Dieses Vorgehen ist als PROBIT-METHODE bekannt und wird im folgenden Abschnitt 2.2.1.1 genauer betrachtet. Werden alle Versuche auf nur zwei Lasthorizonte verteilt, wird das Vorgehen oft als Abgrenzungsverfahren, siehe [27], bezeichnet. Eine solche Unterscheidung ist unnötig und wird daher für eine einheitliche Systematik ignoriert.

Liegen Versuchsergebnisse auf mehreren verschiedenen Lasthorizonten mit bekannter Ausfallwahrscheinlichkeit vor, dann kann die Eintrittswahrscheinlichkeit für genau diese vorliegende Kombination aus Brüchen und Durchläufern als Produkt der Einzelwahrscheinlichkeiten mit

$$P = \prod_{i=1}^k \binom{n_i}{r_i} \cdot p_i^{r_i} \cdot (1 - p_i)^{(n_i - r_i)} \quad (2.28)$$

angegeben werden. Umgekehrt können unter der Annahme, dass das eingetretene Versuchsergebnis das wahrscheinlichste ist, mit Hilfe der als Likelihood-Funktion bezeichneten Gleichung 2.28 die Ausfallwahrscheinlichkeiten bzw. die Verteilungsparameter geschätzt werden. Dieses Vorgehen wird als MAXIMUM-LIKELIHOOD-METHODE bezeichnet und ist im Abschnitt 2.2.1.2 genauer beschrieben. Der rechnerische Aufwand zur Bestimmung der Schätzwerte ist hierbei deutlich aufwendiger als bei der PROBIT-METHODE. Beide Methoden gestatten prinzipiell

¹⁵Welches Ereignis genau als Ausfall der Probe bezeichnet wird, ist verschieden. Das Versagenskriterium kann Bruch der Probe, Anriss der Probe, ein vorher festgelegter Steifigkeitsabfall oder ähnliches sein. Allgemeiner kann auch von Erfolg, wenn eine Reaktion auf die Belastung eintritt und Misserfolg beim gegenteiligen Fall die Rede sein, wenn z. B. die tödliche Dosis von Substanzen an Organismen getestet werden soll. Dann ist natürlich auch die Grenzlasterwechselzahl durch einen zeitlichen Bezug zu ersetzen. Im folgenden wird der Begriff „Bruch“ für den Ausfall einer Probe verwendet.

¹⁶Mehrere Versuche unter identischer Belastungshöhe bzw. Beanspruchungshöhe oder allgemein unter identischen Bedingungen.

die Auswertung von Versuchen auf beliebig vielen, in unterschiedlichen Abständen zueinander liegenden, unterschiedlich stark mit Versuchen belegten Lasthorizonten.

Wird die Versuchsdurchführung derart angelegt, dass beim Eintreten des Versuchsergebnisses Bruch die Last für den nächsten Prüfling um einen konstanten Wert vermindert und im entgegengesetzten Fall um einen konstanten Wert gesteigert wird, entsteht ein getreppter, sich um den Median der Verteilung zentrierender Versuchsablauf. Für die Auswertung dieser als Treppenstufenversuche bezeichneten Versuchsmethode existieren eine Reihe von Vorschlägen, von denen die gebräuchlichsten Auswertemethoden nach DIXON/MOOD, HÜCK und DEUBELBEISS in Abschnitt 2.2.2 beschrieben werden. Außerdem können Treppenstufenversuche formal auch mit der PROBIT-METHODE sowie der MAX.-LIKELIHOOD-METHODE ausgewertet werden, da die verschiedenen Niveaus der Treppenstufen auch als Horizonte aufgefasst werden können.

2.2.1 Durchführung und Auswertung von Horizontversuchen

2.2.1.1 Probitmethode

Mit Hilfe der PROBIT-METHODE ist es möglich, binär ausgeprägte Merkmale auszuwerten. Einige umfassende Erläuterungen der Methode sind von DORFF in [10] und HÜCK in [22] oder [19] dargelegt. Grundgedanke der Auswertung ist die Vorstellung, dass auf jedem Lasthorizont, entsprechend der zugrunde liegenden Verteilung der Festigkeit, eine konstante Ausfallwahrscheinlichkeit $P_A = p$ vorliegt. Die Anzahl der Ausfälle wird dann mit der Binomialverteilung (Gleichung 2.27) beschrieben.

Der Schätzwert \hat{p} für die unbekannte Ausfallwahrscheinlichkeit p auf einem Horizont ist die relative Bruchhäufigkeit

$$\hat{p} = \frac{r}{n}.$$

Bei endlichem Stichprobenumfang treten mitunter auf einem Horizont nur Brüche ($r = n$) oder nur Durchläufer ($r = 0$) auf. Die Ausfallwahrscheinlichkeit wird dann mit $\hat{p} = 1$ bzw. $\hat{p} = 0$ über- bzw. unterschätzt¹⁷ und lässt sich auch nicht im Wahrscheinlichkeitspapier eintragen. Von HÜCK wird daher in [19] eine Abschätzung in Form einer Ungleichung vorgeschlagen. Dazu wird in Gleichung 2.27 für die Wahrscheinlichkeit P der Wert 0,5 festgelegt und es folgt:

$$\hat{p} \geq 0,5^{\frac{1}{n}} \quad \text{für „nur Brüche“ und} \quad (2.29)$$

$$\hat{p} \leq 1 - 0,5^{\frac{1}{n}} \quad \text{für „nur Durchläufer“}. \quad (2.30)$$

¹⁷Die Wahrscheinlichkeit auf einem Lasthorizont mit einer Ausfallwahrscheinlichkeit von $P_A = p = 90\%$ genau 10 Brüche in 10 Versuchen zu erhalten beträgt 34,9%. Es ist also davon auszugehen, dass in einer von drei Versuchsreihen auf einem Lasthorizont mit $P_A = 90\%$ diesem eine Bruchwahrscheinlichkeit von $P_A = 100\%$ zugeordnet werden würde.

Diese Ungleichungen beschreiben den Bereich, in dem die Ausfallwahrscheinlichkeit so groß bzw. klein ist, dass es wahrscheinlicher ist in n Versuchen genau n Brüche bzw. Durchläufer zu erhalten, als alle anderen möglichen Kombinationen von Brüchen und Durchläufern¹⁸.

Die ermittelten Schätzwerte für die Ausfallwahrscheinlichkeiten der einzelnen Horizonte werden im Wahrscheinlichkeitsnetz ausgewertet, wie es in Abschnitt 2.1.3.3 beschrieben ist. Da die jeweiligen geschätzten Ausfallwahrscheinlichkeiten je nach tatsächlicher Ausfallwahrscheinlichkeit und Versuchsanzahl mit unterschiedlichen statistischen Sicherheiten bestimmt werden, müssen sie bei der Regression unterschiedlich stark berücksichtigt werden. Nach HÜCK in [22], S.42, können die Wichtungsfaktoren g_i der jeweiligen geschätzten Ausfallwahrscheinlichkeit nach Gleichung 2.31 berechnet werden. Für reine Bruch- bzw. Durchläuferhorizonte werden die Gewichte nach Gleichung 2.32 bestimmt.

$$g_i = \sqrt{\frac{n_i^2}{n_i - 1} \cdot \hat{p}_i \cdot (1 - \hat{p}_i)} \quad (2.31)$$

$$g_i = \sqrt{\frac{n_i^2}{n_i - 1} \cdot 0,5^{1/n_i} \cdot (1 - 0,5^{1/n_i})} \quad \text{für } n_i = r_i \text{ bzw. } r_i = 0 \quad (2.32)$$

Bei der Regression ist der Abstand zur veränderlichen Größe, in diesem Fall also zur geschätzten Ausfallwahrscheinlichkeit, zu minimieren. Horizonte mit ausschließlich Brüchen oder Durchläufern werden nur dann in der Regression berücksichtigt, wenn die Regressionsgerade außerhalb des durch die Ungleichung 2.29 und 2.30 beschriebenen Bereiches verläuft. Die Lösung des durch diese Forderung entstehenden nichtlinearen Optimierungsproblems erfolgt praktischerweise, indem zunächst mehrere Regressionsrechnungen ungeachtet der letzten Forderung durchgeführt werden, wobei die ausschließlich mit Brüchen oder Durchläufern belegten Horizonte gezielt¹⁹ von der Regression ausgeschlossen werden. Aus den so für alle Kombinationsmöglichkeiten ermittelten Regressionsgeraden wird jene gewählt, deren Reststreuung unter Berücksichtigung der letzten Forderung und aller Horizonte den kleinsten Wert besitzt.

2.2.1.2 Maximum-Likelihood-Methode

Grundlage der MAX.-LIKELIHOOD-METHODE ist die Annahme, dass das im Experiment eingetretene Ergebnis das wahrscheinlichste ist. Um die Verteilungsparameter der Festigkeit zu bestimmen, wird die als bekannt vorausgesetzte Verteilungsfunktion in die auch als Likelihood-Funktion bezeichnete Gleichung der Eintrittswahrscheinlichkeit 2.28 eingesetzt. Im Falle einer angenommenen Normalverteilung $\mathcal{N}(\mu; \sigma^2)$ also:

$$p_i = F(x_i, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x_i} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dt.$$

¹⁸Der durch die Ungleichungen 2.29 und 2.30 beschriebene Bereich für \hat{p} ist identisch mit den Intervallgrenzen des einseitigen Vertrauensbereichs für den Schätzer $\hat{p} = \frac{x}{n}$ einer Binomialverteilung nach PEARSON und CLOPPER, vergl. Abschnitt 2.1.6.1, S. 28, wenn das Konfidenzniveau $1 - \alpha = 0,5$ angesetzt wird.

¹⁹Ist nur ein Bruch- oder Durchläuferhorizont vorhanden, erfolgt die Regression einmal ohne und einmal mit diesem Horizont. Sind zwei Bruch- und/oder Durchläuferhorizonte vorhanden, erfolgt die Regression ohne diese Horizonte, mit diesen Horizonten und jeweils mit einem der beiden Horizonte, usw.

Die Schätzwerte m und s für die Verteilungsparameter μ und σ sind diejenigen Werte, für die die Likelihood-Funktion ihren maximalen Wert annimmt. Durch Logarithmieren vereinfacht sich die Likelihood-Funktion, ohne dass das Maximum der Funktion verschoben wird. Die zu Null gesetzten partiellen Ableitungen der logarithmierten Likelihood-Funktion nach den zu schätzenden Größen bilden die notwendige Bedingung für das Vorliegen eines Maximums. Es entsteht ein Gleichungssystem, im Falle einer Normalverteilung

$$\begin{aligned} 0 &= \sum_i \left[\frac{r_i}{F(x_i, m, s)} - \frac{n_i}{1 - F(x_i, m, s)} \right] f(x_i, m, s) \\ 0 &= \sum_i \left[\frac{r_i}{F(x_i, m, s)} - \frac{n_i}{1 - F(x_i, m, s)} \right] (x_i - m) f(x_i, m, s) \end{aligned}$$

mit $F(x, m, s)$ als Verteilungs- und $f(x, m, s)$ als Dichtefunktion der Normalverteilung. Die Lösung dieses Gleichungssystems erfolgt numerisch. Alternativ ist es natürlich möglich, das vermutete Gebiet der Lösung ausreichend fein zu unterteilen und das Maximum der ursprünglichen Likelihood-Funktion direkt zu suchen.

2.2.1.3 Praktische Durchführung von Horizontversuchen

Den augenscheinlich größten Einfluss auf das Ergebnis der Versuche hat die Lage der Lasthorizonte. Wie deren Lage festgelegt wird, ist in der Literatur nicht eindeutig ausgeführt. Von MAENNIG wird in [27] sowie [28] das Vorgehen beim **Abgrenzungsverfahren**, also dem Probitverfahren auf zwei Horizonten, beschrieben. Demnach wird zunächst eine Probe im Bereich der zu erwartenden Dauerfestigkeit geprüft. Je nach Ergebnis dieses Versuches, also Durchläufer oder Bruch, wird in den folgenden Versuchen die Last erhöht oder vermindert, und zwar solange bis das entsprechende Gegenereignis erstmalig auftritt. Auf diesem Lastniveau L_1 werden dann mehrere Versuche durchgeführt. Die Position des zweiten Lastniveaus ergibt sich aus der Anzahl r der registrierten Brüche unter den n Versuchen des ersten Lasthorizontes aus

$$L_2 = L_1 + d_{AV} \tag{2.33}$$

mit

$$d_{AV} = \begin{cases} (1 - \frac{r}{n}) \cdot C \cdot L_1 & r \leq 0,5 \cdot n \\ -\frac{r}{n} \cdot C \cdot L_1 & r \geq 0,5 \cdot n. \end{cases} \tag{2.34}$$

Der Faktor C dient dabei als Maß für die zu erwartende Streubreite der zu ermittelnden Dauerfestigkeit. Er wird in [27] pauschal mit $C=0,1$ festgelegt und später in [28] differenzierter nach Tabelle 2.1 angegeben. Von BÖHM wird in [2], S. 58 das Abgrenzungsverfahren für die Annahme einer Extremwertverteilung modifiziert und für die Bestimmung des Horizontabstandes die Bedingung aufgestellt, dass der zweite Lasthorizont im Bereich zwischen den Lasten mit 15% und 85 % Ausfallwahrscheinlichkeiten liegen sollte, d. h.

$$C = \frac{L_{85\%} - L_{15\%}}{L_{50\%}}.$$

Da die Auswertung für die Extremwertverteilung durchgeführt wird, gibt BÖHM den Faktor C in Abhängigkeit des zu erwartenden Weibullexponenten k , wie in Tabelle 2.2, an.

2.2 Experimentelle Bestimmung der Bauteildauerfestigkeit

Tabelle 2.1: Zu erwartendes Streumaß C der Dauerfestigkeit nach MAENNIG aus [28] S. 287.

Proben-/ Bauteilart	Werkstoffprobe		einfache Bau- teile (Schraube)	komplizierte Bauteile	
	gekerbt	ungekerbt		gut definierter Form	schlecht definierter Form
Faktor C	0,05...0,15	0,1...0,2	0,2...0,3 (0,4)	0,3...0,4	0,4...0,6

Tabelle 2.2: Zu erwartendes Streumaß C der Dauerfestigkeit nach BÖHM aus [2] S. 58 für verschiedene Exponenten k der Weibullverteilung.

Exponent k	3	4	6	8	10	12	14	16	18	20	22	24
Faktor C	0,78	0,59	0,40	0,30	0,24	0,2	0,17	0,15	0,13	0,12	0,11	0,10

Von MUELLER wird in [31] bei der Untersuchung der Treffsicherheit der Probitmethode der Verschiebungsfaktor x_{Probit} eingeführt und die Höhe des kleinsten und größten Lasthorizontes mithilfe der vorab zu schätzenden Verteilungsparameter der Grundgesamtheit bestimmt:

$$L_{\min, \max} = m_{\text{schätz}} \mp x_{\text{Probit}} \cdot s_{\text{schätz}}$$

Innerhalb dieser Schranken wird die vorgegebene Anzahl an Lasthorizonten gleichmäßig verteilt. Je nach Anzahl der Lasthorizonte werden von MÜLLER in [31], S. 110, optimale Verschiebungsfaktoren ermittelt, wie sie in Tabelle 2.3 wiedergegeben werden. Es wird also in jedem Fall vorausgesetzt, dass die statistische Verteilung der Dauerfestigkeit bzw. deren Parameter vor Beginn der Versuchsreihe ungefähr bekannt sind. Dabei muss vor allem für das Streumaß ein vertrauenswürdiger Wert vorliegen.

Beide oben beschriebenen Herangehensweisen haben zum Ziel, die Lasthorizonte so zu positionieren, dass mit der vorhandenen Anzahl an Proben möglichst gemischte Horizonte, also Horizonte mit Brüchen und Durchläufern, ermittelt werden. Dazu versuchen beide Verfahren den obersten und untersten Lasthorizont möglichst gleichmäßig um einen zentralen Wert der Verteilung zu positionieren. Der Abstand zwischen dem obersten und untersten Horizont kann in Vielfachen der vorab geschätzten Standardabweichung angegeben werden und wird im Folgenden als Spreizung $z = z' \cdot s_{\text{schätz}}$ der Horizonte bezeichnet.

Während nach dem Vorgehen, wie es von MÜLLER beschrieben wird, die Position der Lasthorizonte schon vor dem Beginn der Versuchsreihe festzulegen ist, wird nach MAENNIG mit einer Art kurzem Treppenstufenversuch die Lage des ersten Horizontes in einem möglichst zentralen Bereich der Verteilung sichergestellt. Außerdem wird das Ergebnis des ersten Horizontes für die Lagebestimmung des zweiten Horizontes mit berücksichtigt. Dieses empfehlenswerte Vorgehen sollte daher deutlich unempfindlicher gegen zu starke Verschätzung der Parameter sein. Allerdings ist es dann nicht mehr uneingeschränkt möglich, mehrere Proben gleichzeitig

Tabelle 2.3: Verschiebungsfaktoren x_{Probit} bzw. Spreizung z' für die Lage der Lasthorizonte L_{\min} und L_{\max} , nach MÜLLER aus [31], S. 110.

Horizontanzahl	x_{Probit}	z'
3	1,0	2,0
4	1,0	2,0
6	1,5	3,0
8	2,0	4,0

2 Grundlagen

Tabelle 2.4: Beispielhafter Ablauf des Horizontverfahrens, Annahme normalverteilter Festigkeit.

Stufe i	Hori- zont	$\frac{x}{\text{MPa}}$	Versuchsnummer																		\hat{p}			
			1	2	3	4	5	6	7	8	(3)	9	10	11	12	(1)	(2)	13	14	15		16	17	18
-	L_3	162,4														x	x	x	x	x	x	x	x	> 0,92
3	-	162	x																					
2	-	156		x																				
-	L_2	153,2								x	x	o	o	o										0,4
1	-	150			x																			
0	L_1	144				o	o	o	o	x														0,2

zu prüfen. Diese Einschränkung bringt, nach Einschätzung des Autors, heutzutage allerdings die wenigsten Prüflabore in Bedrängnis.

2.2.1.4 Beispielhafte Auswertung eines Horizontversuchs

Wie die Prüfung und Auswertung von Versuchen zur Bestimmung der Festigkeitsparameter in Horizontversuchen aussehen kann, wird im Folgenden anhand eines Beispiels²⁰ dargestellt.

Vor Prüfbeginn ist die Anzahl der Horizonte h , die Anzahl der Proben n , eine erste Schätzung der mittleren Festigkeit $m_{\text{schätz}}$ und der Streuung $s_{\text{schätz}}$, ein Startwert x_0 (zweckmäßig der Schätzwert der mittleren Festigkeit), ein Stufenabstand d sowie eine Spreizung z der Horizonte festzulegen. In dieser Beispielrechnung sind diese Werte:

$$\begin{aligned}
 h &= 3 & d &= 6 \text{ MPa} & z &= 2 s_{\text{schätz}} \\
 n &= 18 & s_{\text{schätz}} &= 10 \text{ MPa} & x_0 &= m_{\text{schätz}} = 162 \text{ MPa}
 \end{aligned}$$

Auf dem Lastniveau $x_0 = 162 \text{ MPa}$ wird der erste Versuch geprüft und ein Bruch erreicht. Das Lastniveau wird daraufhin um $d = 6 \text{ MPa}$ gesenkt. Auch der zweite und dritte Versuch sind jeweils Brüche, die zu einer Absenkung des Lastniveaus führen. Der vierte Versuch wird ohne den Ausfall der Probe beendet und legt die Lasthöhe des ersten Prüfhorizontes L_1 fest. Im Beispiel stehen noch 14 Proben zur Verfügung. Daher werden für jeden der 3 Horizonte 5 Proben eingeplant und auf dem Horizont L_1 noch weitere 4 Proben geprüft. Die Ausfallwahrscheinlichkeit des Horizontes L_1 wird anschließend zu $\hat{p} = 0,2$ geschätzt. Das zu dieser Wahrscheinlichkeit korrespondierende Quantil der Standardnormalverteilung ist $u_{L1} = -0,8416$. Damit liegt der Horizont L_1 vermutlich deutlich unterhalb des 50%-Wertes der Festigkeit. Der Horizont L_3 wird aus der vorab festgelegten Spreizung ermittelt. Er soll möglichst bei $z/2$ Standardabweichungen oberhalb des Mittelwertes und damit etwa $(z'/2 - u_{L1})$ Standardabweichungen über dem Horizont L_1 liegen, also bei $L_3 = L_{\text{max}} = L_1 + z/2 - u_{L1} s_{\text{schätz}} = 162,4 \text{ MPa}$. Der letzte Horizont L_2 wird gleichmäßig zwischen den beiden vorhandenen Horizonten, bei $x = 153,2 \text{ MPa}$ platziert. Anschließend wird der Versuch Nr. 3 dem mittleren Horizont, die Versuche Nr. 1 und Nr. 2 dem oberen Horizont zugeordnet. Auf beiden Horizonten werden

²⁰Die dem Beispiel zugrunde liegenden Festigkeitswerte der verwendeten 18 Proben sind fiktiv und eine Stichprobe aus einer Reihe von simulierten, normalverteilten Zufallszahlen mit einem tatsächlichen Erwartungswert von $\mu_G = 150 \text{ MPa}$ und einer tatsächlichen Standardabweichung von $\sigma_G = 10 \text{ MPa}$. Die Stichprobe selbst hat einen Mittelwert $\bar{x} = 149,3 \text{ MPa}$ und eine korrigierte Stichprobenvarianz von $s^2 = 91,07 \text{ MPa}$.

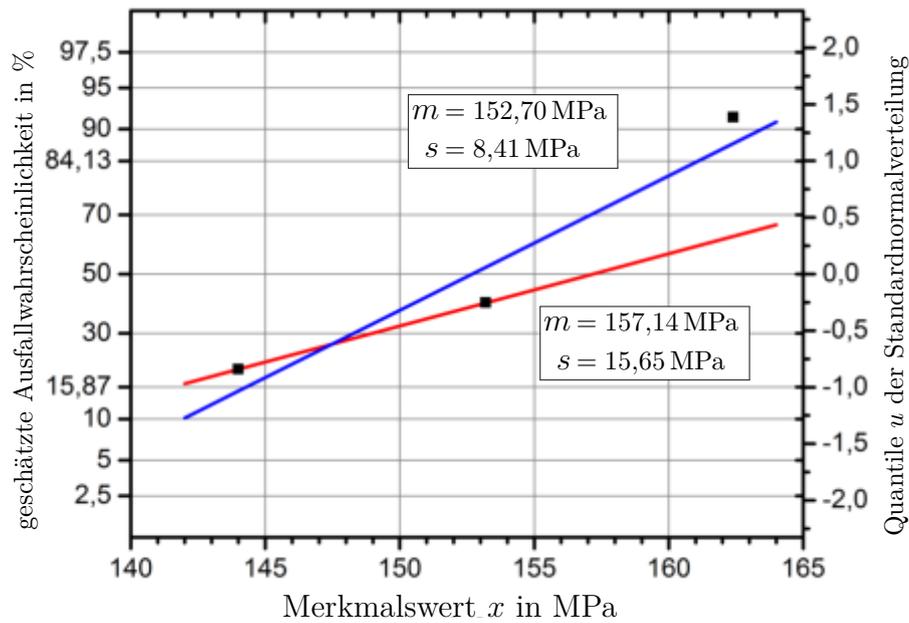


Abb. 2.6: Beispiel für die Regression im Wahrscheinlichkeitsnetz für Auswertung nach der PROBIT-METHODE.

nun weitere Proben geprüft bis jeder Horizont mit 5 Versuchen belegt ist. Die 3 verbleibenden Proben werden auf dem Horizont mit dem geringsten Regressionsgewicht geprüft, in diesem Beispiel, dem Horizont L_3 . Bei einer Auswertung nach der PROBIT-METHODE erfolgt die Regressionsrechnung einmal für alle drei und einmal nur für die mit Brüchen und Durchläufern belegten Horizonte, siehe Abbildung 2.6. Da die ermittelte Regressionsgerade im letzteren Fall (rote Gerade) bei $x = 162,4$ MPa die Ungleichung 2.29 nicht erfüllt, trägt der Horizont L_3 zum Regressionsfehler bei. Dieser fällt damit größer aus als im ersten Fall. Die Schätzwerte für den 50%-Wert und die Standardabweichung sind

$$m_{\text{Probit}} = 152,7 \text{ MPa} \quad \text{und} \quad s_{\text{Probit}} = 8,41 \text{ MPa}.$$

Für eine Auswertung nach der MAX.-LIKELIHOOD-METHODE muss die entsprechende Maximum-Likelihood-Funktion nach Gleichung 2.28 aufgestellt werden. Aus einer grafischen Darstellung der Funktion, siehe Abbildung 2.7, kann das Ergebnis direkt abgelesen werden und beträgt

$$m_{\text{ML}} = 151,6 \text{ MPa} \quad \text{und} \quad s_{\text{ML}} = 6,4 \text{ MPa}.$$

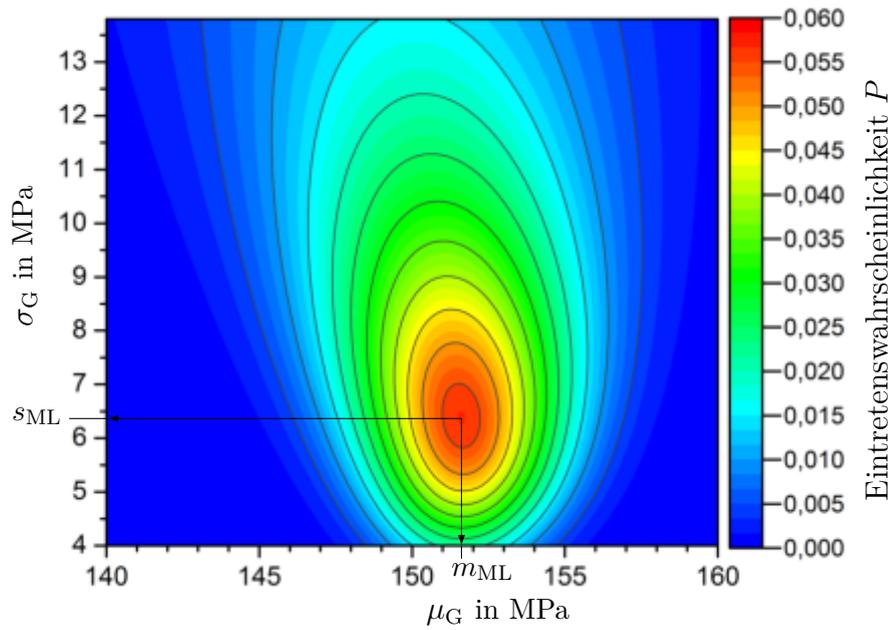


Abb. 2.7: Beispiel einer Likelihood-Funktion für die Ergebnisse eines Horizontversuches in Abhängigkeit der Parameter der Grundgesamtheit μ_G und σ_G . Die Schätzer m_{ML} und s_{ML} für die unbekannt Parameter sind die Koordinaten des Maximums.

2.2.2 Durchführung und Auswertung von Treppenstufenversuchen

2.2.2.1 Nach Dixon/Mood

Von DIXON und MOOD stammt der Vorschlag [9] für die Durchführung und Auswertung von Treppenstufenversuchen. Ausgangspunkt ist die Überlegung, dass bei Treppenstufenversuchen mit ausreichend großer Probenanzahl die Anzahl der Brüche auf einem Lastniveau x immer gleich oder um eins verschieden von der Anzahl der Durchläufer des nächst niedrigeren Niveaus ist. Damit ist die Information über die Versuchsfolge entweder in der Beschreibung der Brüche oder der Durchläufer nahezu vollständig enthalten. In [9] wird gezeigt, wie sich aufgrund dieser Überlegung Gleichung 2.28 zu

$$P = \prod_i \binom{n_i}{r_i} \cdot [p_i \cdot (1 - p_{i-1})]^{r_i} \quad (2.35)$$

umformen lässt. Weiter zeigen DIXON und MOOD, dass sich der Schätzer m für den Erwartungswert μ_G unter der Annahme einer Normalverteilung der ertragbaren Beanspruchung und einem konstanten Stufenabstand für $d/\sigma_G < 2$ proportional zu dem mit der jeweiligen Versuchsanzahl gewichteten Mittelwert der belegten Lasthorizonten verhält. Für den Schätzer s der Standardabweichung σ_G wird bei einem Stufenabstand von $d/\sigma_G = 1$ eine Proportionalität zu einer Varianzkennzahl k_{DM} des seltener eintretenden Ereignisses nachgewiesen. Für den Schätzer m des Erwartungswertes und den Schätzer s der Standardabweichung werden die einfach zu handhabenden Bestimmungsgleichungen

$$m = x_0 + d \cdot \frac{A}{F} \pm d \cdot \frac{1}{2} \quad (2.36)$$

$$\text{und} \quad s = 1,62 \cdot d (k_{DM} + 0,029) \quad (2.37)$$

abgeleitet. Die hierbei verwendeten Hilfsgrößen N , A und B bestimmen sich nach den Gleichungen 2.38 bis 2.40. Die Varianzkennzahl k_{DM} wird nach Gleichung 2.41 ermittelt, n_i ist die Anzahl der gezählten Ereignisse auf dem i -Lastniveau x_i und mit x_0 wird die Lasthöhe des niedrigsten Lastniveaus angegeben. Das variable Vorzeichen ist bei Auswertung von Brüchen mit „-“ und bei Auswertung von Durchläufern mit „+“ zu setzen.

$$N = \sum_i n_i \quad (2.38) \quad A = \sum_i i \cdot n_i \quad (2.39) \quad B = \sum_i i^2 \cdot n_i \quad (2.40)$$

$$k_{\text{DM}} = \frac{N \cdot B - A^2}{N^2} \quad (2.41)$$

Die Gleichungen 2.36 und 2.37 gelten für $k_{\text{DM}} > 0,3$. Andernfalls sind die Schätzer, analog zur MAX.-LIKELIHOOD-METHODE, diejenigen Werte, für die die Auftretenswahrscheinlichkeit nach Gleichung 2.35 maximal wird.

Beginnt die Versuchsreihe weit entfernt vom zu schätzenden Erwartungswert der Verteilung, treten oft am Anfang der Versuchsreihe nur Versuche mit dem Ergebnis Bruch bzw. Durchläufer auf. Von DIXON/MOOD werden solche Versuche von der Auswertung ausgeschlossen. Alle Versuche nach dem erstmaligem Wechsel des Versuchsergebnisses sowie der direkt davor liegende Versuch werden als die für die Auswertung zu berücksichtigenden gültigen Versuche betrachtet. Da aber in der Auswertung stets nur das seltener auftretende Ereignis berücksichtigt wird, ist zu erwarten, dass diese Behandlung des Anschnitts nur für kleine Stichprobenumfänge und bei kleinen Stufenabständen Einfluss auf das Schätzergebnis besitzt.

Für die Bewertung der Schätzung unerlässlich ist die Kenntnis über deren Unsicherheit. Von DIXON/MOOD wird diese in Form der Standardfehler s_m und s_s für die Schätzer des Erwartungswertes und der Standardabweichung mit den Gleichungen

$$s_m = \frac{C_m}{\sqrt{N}} \cdot s \quad (2.42)$$

$$s_s = \frac{C_s}{\sqrt{N}} \cdot s \quad (2.43)$$

angegeben. Die dabei verwendeten Faktoren C_m und C_s sind aus dem in Abbildung 2.8 wiedergegebenen Diagramm abzulesen.²¹

Die Gleichungen zur Schätzung der Verteilungsparameter können für ein besseres Verständnis der Schätzmethode auch anders angegeben werden: So beschreibt die Gleichung 2.36 nichts anderes als den um einen halben Stufenabstand verschobenen Mittelwert des Lastniveaus aller gezählten Brüche bzw. Durchläufer, da gilt:

$$\begin{aligned} m &= x_0 + d \cdot \frac{A}{N} \pm \frac{d}{2} = x_0 + d \cdot \frac{\sum_i i \cdot n_i}{\sum_i n_i} \pm \frac{d}{2} \\ &= x_0 + \frac{\sum_i n_i \cdot (x_i - x_0)}{N} \pm \frac{d}{2} \\ &= x_0 + \frac{\sum_i n_i \cdot x_i}{N} - x_0 \frac{N}{N} \pm \frac{d}{2} = \bar{x} \pm \frac{d}{2}. \end{aligned}$$

²¹Im Original werden für die Faktoren die Bezeichnungen G und H verwendet. Die Unsicherheiten sind bei größeren Stufenabständen von der Lage des Erwartungswertes zu den Lastniveaus abhängig. Die gestrichelt dargestellten Kurven gelten, wenn der Erwartungswert genau mittig zwischen zwei Lastniveaus liegt. Die durchgezogene Kurve gilt für den Fall eines auf einem Lastniveau liegenden Erwartungswertes.

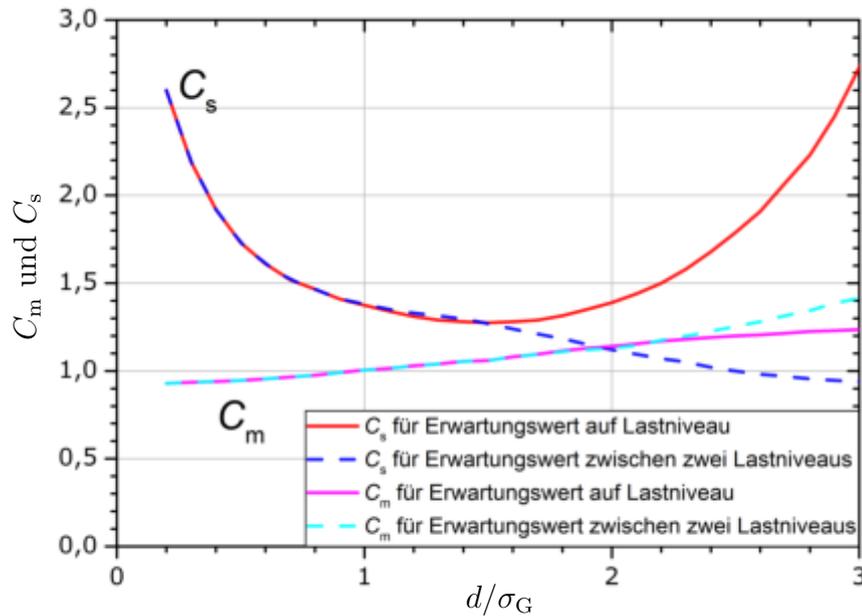


Abb. 2.8: Faktoren²¹ für die Standardfehler der Schätzungen der Verteilungsparameter im Treppenstufenversuch nach DIXON & MOOD [9].

Die nach Gleichung 2.41 zu bestimmende Varianzkennzahl ist die auf den Stufenabstand normierte Varianz des Lastniveaus aller geprüften Brüche bzw. Durchläufer, denn es gilt:

$$\begin{aligned}
 k_{\text{DM}} &= \frac{N \cdot B - A^2}{N^2} = \frac{B}{N} - \left(\frac{A}{N}\right)^2 = \frac{d^2}{d^2} \left(\frac{\sum i^2 \cdot n_i}{N} - \frac{(\bar{x} - x_0)^2}{d^2} \right) \\
 &= \frac{1}{d^2} \left(\frac{\sum n_i \cdot (x_i - x_0)^2}{N} - (\bar{x} - x_0)^2 \right) \\
 &= \frac{1}{d^2} \left(\frac{\sum n_i \cdot x_i^2}{N} - \bar{x}^2 \right) = \frac{1}{d^2} \cdot \frac{\sum n_i (x_i - \bar{x})^2}{N}.
 \end{aligned}$$

Die Bestimmungsgleichung 2.37 kann dann als eine Korrekturfunktion interpretiert werden, mit der die auf den Stufenabstand normierte Varianz der Brüche bzw. Durchläufer in Abhängigkeit vom gewählten Stufenabstand berichtigt werden muss, um die Standardabweichung der Grundgesamtheit zu schätzen. Gleichung 2.37 wird dann zu

$$s = 1,62 \cdot \frac{1}{d} \cdot \frac{\sum n_i (x_i - \bar{x})^2}{N} + 0,04698 \cdot d.$$

2.2.2.2 Nach Hück

Von HÜCK wird in [21] eine, an die Auswertemethode nach DIXON/MOOD angelehnte, ähnliche Auswertemethode vorgestellt. Hierbei werden alle Versuche, unabhängig von ihrem Ergebnis gezählt, wie es auch von anderen Autoren, z. B. [4], vorgeschlagen wird. Außerdem wird ein fiktiver Versuch am Ende der Versuchsreihe hinzugefügt, der wengleich sein Ergebnis unbekannt ist, durch die gemeinsame Auswertung von Brüchen und Durchläufern in die

Auswertung mit einbezogen werden muss. Aus der Gesamtanzahl der geprüften Proben jedes Horizontes werden die folgenden Hilfsgrößen bestimmt:

$$N = \sum_i n_i \quad (2.44) \quad A = \sum_i i \cdot n_i \quad (2.45) \quad B = \sum_i i^2 \cdot n_i \quad (2.46)$$

$$k_H = \frac{N \cdot B - A^2}{N^2} \quad (2.47)$$

Der Schätzer für den Erwartungswert folgt aus der Gleichung 2.48. Die Standardabweichung der Grundgesamtheit wird mithilfe des in Abbildung 2.9 a) dargestellten Diagramms aus der Varianzkennzahl k_H geschätzt.

$$m = x_0 + d \cdot \frac{A}{N} \quad (2.48)$$

Auch die Definition des Anschnitts einer Versuchsreihe wird von HÜCK verändert. In der Auswertung nicht zu berücksichtigen sind alle Versuche am Anfang der Versuchsreihe, deren Lasthöhe mit den folgenden Versuchen nicht wieder erreicht wird.

Für beide Schätzer werden von HÜCK in [21] auch die Standardfehler s_m und s_s der Schätzer angegeben, wie sie in Gleichung 2.49 und 2.50 wiedergegeben sind. Die darin enthaltenen Faktoren C_m und C_s sind aus Diagrammen zu entnehmen, siehe Abbildung 2.10. Sie lassen sich aber, wie im Anhang B.1.1, S. 263 dargestellt, auch mathematisch beschreiben.

$$s_m = C_m \cdot s \quad (2.49)$$

$$s_s = C_s \cdot d \quad (2.50)$$

Analog zur Auswertung nach DIXON/MOOD können die Gleichungen zur Parameterschätzung auch anders dargestellt werden. Gleichung 2.48 ist nichts anderes als der Mittelwert und die Varianzkennzahl nach Gleichung 2.47 ist die auf den quadrierten Stufenabstand normierte Varianz aller gültigen Versuche. Aus letzterer wird mit der entsprechenden Korrekturfunktion die Standardabweichung ermittelt.

$$m = x_0 + d \cdot \frac{A}{N} = \frac{\sum n_i \cdot x_i}{N} = \bar{x} \quad (2.51)$$

$$k_H = \frac{N \cdot B - A^2}{N^2} = \frac{1}{d^2} \cdot \frac{\sum n_i (x_i - \bar{x})^2}{N} \quad (2.52)$$

Die eigentlich feste Höhe der Lastniveaus der Einzelversuche wird also als das zufällige Ereignis interpretiert. Die Schätzwerte der Verteilungsparameter werden dann aus den Momentenschätzern abgeleitet. Das ist für die Schätzung des Erwartungswertes möglich, da es gerade die Eigenschaft der Treppenstufenfolge ist, sich um den Median und damit bei symmetrischen Verteilungen auch um den Erwartungswert zu zentrieren. Für die Schätzung der Standardabweichung wird zusätzlich eine Korrekturfunktion benötigt. Diese wurde von HÜCK durch Monte-Carlo-Simulationen ermittelt. Hierzu wurden für konstante Verhältnisse von vorgegebenen Standardabweichungen und Stufenabständen die Mittelwerte der dabei berechneten Varianzkennzahlen grafisch dargestellt. Die auf diese Weise ermittelten Korrekturfunktionen werden Abbildung 2.9 a) wiedergegeben. Für die mathematische Beschreibung der Kurvenzüge sind verschiedene Ansätze bekannt und im Anhang, Abschnitt B.1.1 ab S. 263, zu finden. Vor allem für kleine Stichprobenumfänge ist dabei zu bedenken, dass die ermittelten Kurvenzüge zum einen nur mittlere Korrekturfaktoren darstellen und zum anderen

2 Grundlagen

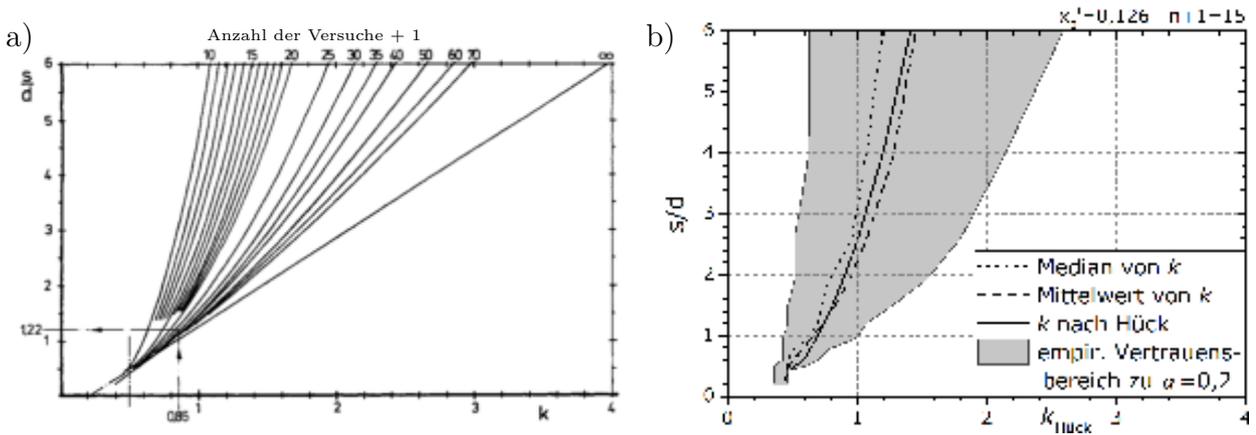


Abb. 2.9: Korrekturfunktion zur Schätzung der Standardabweichung σ_G aus der Varianzkennzahl k_H , nach HÜCK [21] in a) und aus eigenen Monte-Carlo-Simulationen für $n + 1 = 15$ mit grau hinterlegtem Bereich für 80% aller ermittelten Werte für k in b).

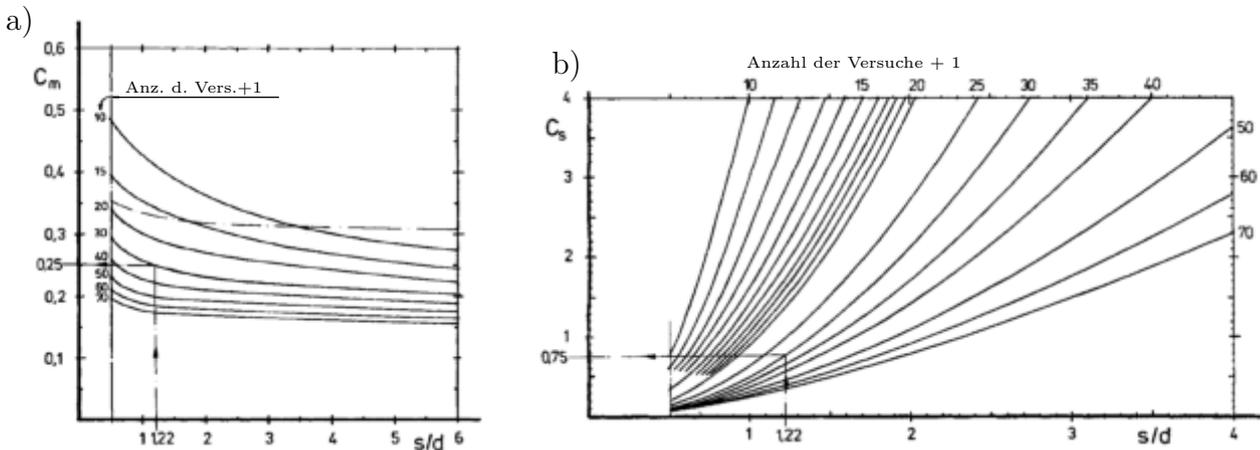


Abb. 2.10: Kennlinien aus [21] zur Bestimmung der Standardfehler der Parameterschätzungen. Standardfehler der Erwartungswertschätzung in a) und Standardfehler der Standardabweichungsschätzung in b).

bereits kleine Abweichungen der ermittelten Varianzkennzahlen zu sehr stark veränderten Korrekturfaktoren führen. Wie stark die ermittelte Varianzkennzahl bei einem fest vorgegebenen Verhältnis aus tatsächlicher Standardabweichung und vorgegebenem Stufenabstand streuen kann, ist in Abbildung 2.9 b) exemplarisch dargestellt. Demnach ist es aufgrund dieser Streuung möglich, für eine Versuchsreihe mit $n = 14$ gültigen Einzelversuchen nahezu beliebige Korrekturfaktoren zu ermitteln.

2.2.2.3 Nach Deubelbeiss

Die Auswertung von Treppenstufenversuchen nach DEUBELBEISS erfolgt als Auswertung im Wahrscheinlichkeitsnetz. Hierfür wird nach [5] für große Stichproben ($n > 25$) die Anzahl der Versuche einer jeden Laststufe auf die Gesamtversuchsanzahl normiert. Die so gebildeten relativen Häufigkeiten werden anschließend von der kleinsten bis zur aktuellen Laststufe summiert. Diese so ermittelten Ausfallwahrscheinlichkeiten der verschiedenen Laststufen sind dann in das Wahrscheinlichkeitsnetz einzutragen und auszuwerten. Für kleine Stichproben mit $n < 25$ werden die Einzelversuche entsprechend ihrer Laststufe und ohne Rücksicht auf das

tatsächliche Versuchsergebnis in aufsteigender Reihenfolge sortiert. Anschließend wird jedem der n Einzelversuche eine Ordnungszahl j , $1 \leq j \leq n$ zugeordnet. Nach Gleichung 2.53 kann dann jedem Versuch eine Ausfallwahrscheinlichkeit zugeordnet und in das Wahrscheinlichkeitsnetz eingetragen werden. Aus der Regressionsgeraden durch alle eingetragenen Punkte erfolgt anschließend die Schätzung der Parameter der Grundgesamtheit, vgl. Abschnitt 2.1.3.3.

$$P_A = \frac{100}{n} \left((j - 0,5) + \frac{n + 1 - 2j}{8(n - 1)} \right) \quad (2.53)$$

Problematisch an dieser Schätzmethode ist allerdings, dass sowohl bei kleinen als auch bei größeren Stichprobenumfängen die in einem Treppenstufenversuch geprüften Einzelversuche als Stichproben einer metrischen Zufallsgröße interpretiert werden. So ist Gleichung 2.53 gerade jene Formel, die von ROSSOW in [36] durch $(3j - 1) / (3n + 1)$ approximiert wird. Genauso gut könnte dann aber auch die Momentenmethode für die Schätzung der Verteilungsparameter verwendet werden. Wenn n_i die Anzahl der Versuche auf einer Lasthöhe x_i und $N = \sum n_i$ die Gesamtanzahl der Versuche bezeichnet, dann wären nach der Momentenmethode die Schätzer des Erwartungswertes und der Standardabweichung mit $m = \sum (n_i \cdot x_i) / N$ und $s^2 = \sum (n_i (x_i - \bar{x})^2) / N$ zu berechnen. Wie schon in Abschnitt 2.1.3 beschrieben, sind die Unterschiede der Parameterschätzungen zwischen der Momentenmethode und der Auswertung im Wahrscheinlichkeitsnetz gering. Allenfalls für die Schätzer der Standardabweichung ergeben sich hier Unterschiede. Von größerer Bedeutung ist hingegen, dass die nach der Momentenmethode ermittelte Standardabweichung genau so berechnet wird, wie die nach Gleichung 2.52 ermittelte Varianzkennzahl k_H bei einer Auswertung nach HÜCK. Während die Letztere aber mit einer Korrekturfunktion in eine Schätzung der Standardabweichung überführt wird, dient bei der Auswertung nach DEUBELBEISS dieser Schätzwert direkt als Schätzung der Standardabweichung. Ziel der verwendete Korrekturfunktion ist es, in Abhängigkeit des Stichprobenumfangs und Stufenabstandes möglichst erwartungstreue Schätzwerte für die Standardabweichung zu erhalten. Daher ist davon auszugehen, dass die Auswertung nach DEUBELBEISS keine erwartungstreuen und auch keine asymptotisch erwartungstreuen Schätzungen der Standardabweichung liefern kann.

2.2.2.4 Auswertung mit Probit- oder Max.-Likelihood-Methode

Prinzipbedingt können Treppenstufenversuche auch mit der PROBIT-METHODE oder der MAX.-LIKELIHOOD-METHODE ausgewertet werden. Während für die MAX.-LIKELIHOOD-METHODE keine Einschränkungen zu erwarten sind, ist bei der PROBIT-METHODE zu bedenken, dass die Güte der Schätzungen von der Anzahl der Versuche je Horizont abhängig ist und für gute Schätzungen möglichst sowohl mit Brüchen als auch mit Durchläufern belegte Horizonte benötigt werden. Bei Treppenstufenversuchen ist der höchste und niedrigste Horizont meist aber gerade ein mit wenigen Versuchen belegter, reiner Bruch- bzw. Durchläuferhorizont.

2.2.2.5 Beispielhafte Auswertung eines Treppenstufenversuches

Für eine beispielhafte Auswertung einer Treppenstufenversuchsreihe dient die gleiche Stichprobe simulierter Zufallszahlen, wie sie für die Beispielrechnung der Horizontversuche im

2 Grundlagen

Tabelle 2.5: Beispiel des Ablaufs des Treppenstufenverfahrens bei Auswertung nach DIXON/MOOD.

Stufe i	$\frac{x}{\text{MPa}}$	Versuchsnummer																n_i	$i \cdot n_i$	$i^2 \cdot n_i$		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16				17	18
4	162	A																		0	0	0
3	156		A				X					X								0	0	0
2	150			X		O		X			O		X					X		2	4	8
1	144				O				X		O			X			O		X	3	3	3
0	138									O					X		O			2	0	0
																Summe:	7	7	11			
																	<i>F</i>	<i>A</i>	<i>B</i>			

Tabelle 2.6: Beispiel des Ablaufs des Treppenstufenverfahrens bei Auswertung nach HÜCK.

Stufe i	$\frac{x}{\text{MPa}}$	Versuchsnummer																n_i	$i \cdot n_i$	$i^2 \cdot n_i$		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16				17	18
4	162	A																		0	0	0
3	156		X				X					X								3	9	27
2	150			X		O		X			O		X				X			6	12	24
1	144				O				X		O			X			O		X	6	6	6
0	138									O					O				f	2 + 1	0	0
																Summe:	18	27	57			
																	<i>F</i>	<i>A</i>	<i>B</i>			

Abschnitt 2.2.1.4 verwendet wird. Vor Versuchsbeginn sind die Parameter der Festigkeitsverteilung abzuschätzen, um einen Startwert x_0 und einen Stufenabstand d festlegen zu können. In diesem Beispiel besitzen diese Größen folgende Werte:

$$s_{\text{schätz}} = 6 \text{ MPa} \qquad x_0 = m_{\text{schätz}} = 162 \text{ MPa}$$

$$d = s_{\text{schätz}} \qquad n = 18$$

Beginnend von x_0 werden die Proben geprüft und es entsteht der Treppenstufenverlauf, wie er in Tabelle 2.5 dargestellt ist. Für die Auswertung nach DIXON/MOOD fallen die ersten zwei Versuche in den Bereich des Versuchsreihenanschnitts. Es wird nur das seltener eingetretene Ereignis, hier das Auftreten von Durchläufern, zur Berechnung der Hilfsgrößen N , A und B betrachtet. Die Varianzkennzahl der Durchläufer wird zu $k_{\text{DM}} = 0,571$ berechnet, die Schätzwerte für die Verteilung der Festigkeit sind

$$m_{\text{DM}} = 147,0 \text{ MPa} \qquad \text{und} \qquad s_{\text{DM}} = 5,84 \text{ MPa.}$$

Für die Auswertung nach HÜCK sowie DEUBELBEISS werden Brüche und Durchläufer gemeinsam ausgewertet. Die Lasthöhe des ersten Versuchs wird nicht wieder erreicht, sodass dieser in beiden Auswertungen dem Anschnitt zugeordnet wird. Außerdem wird bei beiden Verfahren noch ein fiktiver, am Ende der Versuchsreihe hinzugefügter Versuch in die Auswertung mit einbezogen. Die Hilfsgrößen für die Auswertung nach HÜCK sind in Tabelle 2.6 dargestellt. Mit $k_{\text{H}} = 0,9167$ werden die Schätzer zu

$$m_{\text{H}} = 147,0 \text{ MPa} \qquad \text{und} \qquad s_{\text{H}} = 10,36 \text{ MPa}$$

berechnet. Die für die Auswertung nach DEUBELBEISS berechneten Überschreitungs- bzw. Positionswahrscheinlichkeiten sind in Abbildung 2.11 im Wahrscheinlichkeitsnetz aufgetragen.

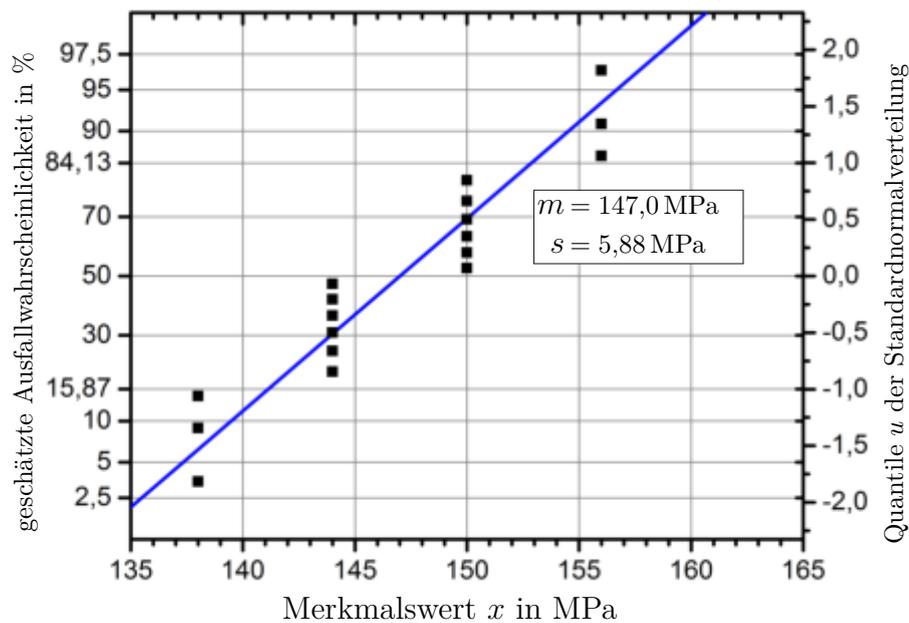


Abb. 2.11: Beispiel für die Regression im Wahrscheinlichkeitsnetz bei einer Auswertung nach DEUBELBEISS.

Es werden die Schätzer

$$m_D = 147,0 \text{ MPa} \quad \text{und} \quad s_D = 5,88 \text{ MPa}$$

ermittelt. Für die Auswertung des Treppenstufenverlaufes mit der PROBIT-METHODE oder der MAX.-LIKELIHOOD-METHODE lauten die Schätzwerte

$$\begin{aligned} m_{PR} &= 144,4 \text{ MPa} & \text{und} & & s_{PR} &= 12,39 \text{ MPa bzw.} \\ m_{ML} &= 145,7 \text{ MPa} & \text{und} & & s_{ML} &= 6,03 \text{ MPa.} \end{aligned}$$

2.2.3 Weitere Verfahren zur Bestimmung der Dauerfestigkeit

Neben den oben beschriebenen Auswerteverfahren zur Schätzung der Verteilungsparameter der Dauerfestigkeit gibt es eine Reihe von weiteren Ansätzen und Auswerteverfahren mit unterschiedlichen Zielstellungen. Ist das Ziel die Verminderung des Versuchsaufwandes, so sind die Laststeigerungsverfahren nach PROT, LOCATI (siehe hierzu [34]) oder ROTEM [37] zu erwähnen. Hierbei werden weitere Annahmen, z. B. über den Anstieg der Wöhlerlinie im Zeitfestigkeitsbereich und die Gültigkeit der linearen Schadensakkumulationshypothese nach PALMGREN und MINER vorausgesetzt. Diese zusätzlichen Voraussetzungen schieben den experimentellen Aufwand auf andere durchzuführende Versuche. Die Güte der Schätzungen zu bewerten, wird dadurch nahezu unmöglich. In der Literatur sind auch Vorschläge für die Auswertung sehr kurzer Treppenstufenversuchsreihen zu finden, z. B. von BROWNLEE et. al. [4] oder DIXON [8], mit denen die Lage des Medians der Verteilung geschätzt werden kann. Das Verfahren nach KLUBBERG [25] hat die möglichst exakte Schätzung des Streuparameters zum Ziel und mischt hierfür Treppenstufen- und Horizontverfahren.

2.2.4 Bemerkung zur Umsortierung von Versuchsergebnissen

Bedingt durch die meist begrenzte Anzahl an Proben, besteht oft der Wunsch des Versuchsführenden, die vorhandenen Proben effektiv zu nutzen. So ist es beispielsweise denkbar, nicht gebrochene Proben erneut auf einem höheren Lasthorizont zu prüfen oder einem niedrigeren Lasthorizont, auf dem jene Proben ebenfalls nicht versagt hätten, zuzuordnen. Ebenso könnten bereits gebrochene Proben einem höheren Lasthorizont zugeordnet werden, mit der Begründung, dass sie auch auf diesem Versagen würden. Bei Treppenstufenversuchen ist es ebenfalls vorstellbar, die Reihenfolge der Versuche zu verändern um „Lücken“ in der Treppenstufenfolge zu füllen oder Versuchsergebnisse so zu sortieren, dass für die Auswertung gültige Treppenstufenfolgen entstehen. Zu beachten ist dabei jedoch, dass die Grundannahme der Versuchsführung, dass nämlich jede Prüfung ein Bernoulliexperiment (horizontweiser) gleicher Erfolgswahrscheinlichkeit ist, nicht verletzt wird. Dies ist gerade dann der Fall, wenn Proben abhängig von ihrem bisherigen Ergebnis auf andere Lasthorizonte verschoben werden, denn durch das bereits bekannte Ergebnis werden diese Proben nach Proben hoher und niedriger Festigkeit sortiert.^{22,23}

²²Nach dem Satz von der totalen Wahrscheinlichkeit ([16], S. 100) für die nicht bedingte Wahrscheinlichkeit des Ereignisses A , bei zwei sich gegenseitig ausschließenden Ereignissen B_1 und B_2 gilt:

$$P(A) = P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2)$$

Ist $P(A)$ die Ausfallwahrscheinlichkeit auf einem Lasthorizont A und sind $P(B_1)$ bzw. $P(B_2)$ die Ausfall- bzw. Überlebenswahrscheinlichkeiten auf einem Horizont B , so ist es möglich, die bedingte Wahrscheinlichkeit $P(A|B_2)$ für den Ausfall eines auf dem Horizont B erhaltenen Durchläufers auf dem Horizont A zu ermitteln. Werden zum Beispiel alle Durchläufer des Horizontes B mit $P(B_1) = 0,5$ auf einem Horizont A mit $P(A) = 0,8$ noch einmal geprüft, dann wird mit $P(A|B_1) = 1$ eine Ausfallwahrscheinlichkeit von

$$P(A|B_2) = \frac{P(A) - P(B_1) \cdot P(A|B_1)}{P(B_2)} = 0,6$$

erwartet und damit die Ausfallwahrscheinlichkeit unterschätzt, da nur die „besseren“ Proben erneut geprüft wurden. Wird ein solcher Fehler systematisch begangen, lassen sich nicht vorhandene „Trainiereffekte“ erzeugen.

²³In Abschnitt 2.2.1.4 werden beim Horizontverfahren die ursprünglich zum Auffinden des Dauerfestigkeitsgebietes benötigten Versuche anderen Lasthorizonten zugeordnet. Damit wird prinzipiell derselbe, oben dargestellte, Fehler begangen. Die entstehenden Schätzfehler werden jedoch durch die Zuordnung zum nächstmöglichen Horizont minimiert.

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

3.1 Bekannte Arbeiten

Zur Simulation von Treppenstufenversuchen durch Monte-Carlo-Simulation¹ sind verschiedene Arbeiten erschienen. Die früheste dem Autor bekannte Veröffentlichung hierzu stammt von HÜCK [21] und ist die Grundlage der im Abschnitt 2.2.2.2 beschriebenen Treppenstufenauswertung nach HÜCK (IABG-Methode). Darin wird gezeigt, welche Verbesserungen bei der Auswertung von Treppenstufenversuchen erreicht werden können, wenn eine gemeinsame Auswertung der Brüche und Durchläufer vorgenommen wird. Dazu werden für die Verhältnisse aus Standardabweichung der Grundgesamtheit zu Stufenabstand $\sigma_G/d = \{0,5; 1; 2; 4; 6\}$ je eine aus jeweils 150000 Einzelversuchen bestehende Versuchsreihe simuliert und abschließend für jede untersuchte Versuchsanzahl $n = \{10; \dots; 70\}$ in kleinere Versuchsreihen unterteilt und ausgewertet. Dadurch ist der Startwert der ausgewerteten Versuchsreihe immer durch das Ergebnis des letzten Versuches der vorherigen Versuchsreihe vorbestimmt und damit selbst eine Zufallsgröße, deren Erwartungswert mit dem Erwartungswert der simulierten Dauerfestigkeit zusammenfällt und dessen Streuung durch das Verhältnis σ/d bestimmt ist. Es wird gezeigt, dass sich die Vertrauensbereiche der Schätzungen für Erwartungswert und Standardabweichung unter Annahme der Normalverteilung der Schätzer aus der ermittelten Standardabweichung berechnen lassen und diese im Vergleich zu einer Auswertung nach DIXON und MOOD [9] zu einer schärferen Schätzung führen, d. h. ein kleineres Konfidenzintervall für die geschätzten Parameter angegeben werden kann.

Von LIU werden in [26] ebenfalls Treppenstufenversuche simuliert. Dabei wird neben der Frage nach der statistisch besseren Auswertemethode auch der Frage nach dem Einfluss des Startwertes nachgegangen. Hierzu wird eine geschätzte mittlere Dauerfestigkeit als eine normalverteilte Zufallsgröße, deren Erwartungswert dem Erwartungswert der Grundgesamtheit entspricht und deren Streuung variiert, eingeführt. In der Simulation wird dann die Treffsicherheit der Parameterschätzung untersucht, wenn als Startwert dieser Schätzwert direkt verwendet wird oder aber bewusst ein Startwert sehr kleiner oder sehr großer vermuteter Ausfallwahrscheinlichkeit gewählt wird. Im Ergebnis dieser Untersuchungen zeigt sich, dass nur bei einem Startwert, der dem Erwartungswert der Dauerfestigkeit der Grundgesamtheit

¹Der Begriff der Monte-Carlo-Simulation wird allgemein verwendet, wenn für die Eingangsgrößen eines Modells zufällige Werte aus dem Bereich der möglichen Werte gewählt werden und mit diesen dann die Ausgangsgrößen des Modells berechnet werden. Nach dem Gesetz der großen Zahlen können so die Eigenschaften der Ausgangsgrößen des Modells, insbesondere die Parameter der statistischen Verteilung der Ausgangsgrößen ermittelt werden.

entspricht, die überprüften Verfahren erwartungstreue Schätzwerte liefern. Eine Betrachtung zum Schätzer für die Standardabweichungen der Dauerfestigkeit der Grundgesamtheit erfolgt nur für die Auswertung nach der MAX.-LIKELIHOOD-METHODE.

Die Idee, den Startwert als normalverteilte Zufallsgröße zu simulieren, wird auch von MÜLLER ET. AL. [32] verwendet. Hier erfolgt die Monte-Carlo-Simulation allerdings unter der Annahme logarithmisch normalverteilter Festigkeiten. Untersucht wird die Güte der Schätzungen für die verschiedenen Auswerteverfahren nach HÜCK, DEUBELBEISS, und der MAX.-LIKELIHOOD-METHODE. Es wird dabei herausgestellt, dass alle Verfahren den Mittelwert vergleichbar gut und erwartungstreu abschätzen, sich die Standardabweichung allerdings nur nach HÜCK ab einer größeren Versuchsanzahl erwartungstreu schätzen lässt. Weiterhin stellen die Autoren fest, dass der systematische Fehler der Schätzung für die Standardabweichung bei der Auswertung nach DEUBELBEISS vor allem von der Wahl des Stufenabstandes im Treppenstufenversuch abhängig ist. Alle Ergebnisse werden in Abhängigkeit der zu Beginn der Prüfung vorhandenen Proben dargestellt. Dadurch ist es nicht möglich, in der Literatur veröffentlichte Ergebnisse, bei denen (wenn überhaupt) meist nur die Anzahl der zur Auswertung verwendeten Versuche angegeben ist, einzuschätzen. Auf Fragen einer systematischen Startwertverschiebung wird ebenso wenig eingegangen, wie auf den Einfluss unterschiedlicher Anschnittsdefinitionen.

Für die Simulation von Horizontversuchen ist dem Autor nur die Arbeit von MÜLLER [31] bekannt, die neben einer vertieften Betrachtung zur Treffsicherheit von Treppenstufenversuchen auch die Treffsicherheit von Horizontversuchen bei einer Auswertung nach der PROBIT-METHODE untersucht.² Unter der Annahme logarithmisch normalverteilter Festigkeiten werden Versuche auf Horizonten mit definierten Abständen simuliert, um die Treffsicherheit der PROBIT-METHODE abzuschätzen. Es wird dabei wieder von einem zufälligen, im Mittel aber die mittlere Festigkeit treffenden, Versuchsreihenstartwert ausgegangen. Die Schätzung der Ausfallwahrscheinlichkeiten der verschiedenen Horizonte erfolgt bemerkenswerter Weise nicht durch die relative Ausfallhäufigkeit, sondern mithilfe der Schätzformel nach ROSSOW, siehe Gleichung 2.8 in Abschnitt 2.1.3.3. Wie von MÜLLER selbst bemerkt, führt das zu einem systematischen Verschätzen der mittleren Festigkeit.

Die Idee, die Treffsicherheit der Schätzungen von mittleren Festigkeiten und Festigkeitsstreuungen aus Treppenstufenversuchen und Horizontversuchen anhand von Monte-Carlo-Simulationen zu überprüfen, ist also bereits mehrfach umgesetzt. Für die Verteilung der Festigkeiten werden sowohl Normalverteilung als auch logarithmische Normalverteilung angenommen. Da sich die logarithmische Normalverteilung ohne Schwierigkeiten in eine Normalverteilung transformieren lässt, erschwert die Unterscheidung lediglich die Interpretation der verschiedenen Ergebnisse und trägt nur wenig zum Verständnis bei. Die Festlegung des Versuchsreihenstartwertes erfolgt zufällig. Ist der mittlere Startwert mit der mittleren Festigkeit identisch, führt das zu erwartungstreuen Schätzern der mittleren Festigkeit. Für praktische Versuche scheint diese Annahme wenig glaubhaft. Allein schon um die Versuchszeit zu verkürzen und damit Kosten zu sparen, wird der Versuchsführende eher mit einem zu großen als mit einem zu kleinen Versuchsreihenstartwert beginnen. Die Ergebnisse sind immer von der Anzahl der tatsächlich geprüften Proben abhängig. Dabei ist für die Einschätzung

²Von HÜCK wird in [19] außerdem skizziert, wie man die Unsicherheiten bei der Bestimmung der mittleren Festigkeit bei Horizontversuchen und der Auswertung mittels PROBIT-METHODE mit Methoden der Kombinatorik lösen könnte.

der Ergebnisse weniger interessant, wie viele Proben am Anfang des Versuches zur Verfügung stehen, als vielmehr, wie viele Proben im Sinne der Auswertung gültige Versuche liefern. Weiterhin unbekannt sind die Auswirkungen unterschiedlicher Anschnittsdefinition nach DIXON oder HÜCK, siehe Abschnitt 2.2.2, S. 46 ff. auf die Treffsicherheit der Schätzungen der Verteilungsparameter der Grundgesamtheit. In allen bisherigen Untersuchungen wird lediglich die Treffsicherheit der Verteilungsparameter untersucht. Die für die praktische Anwendung wichtige Treffsicherheit bzgl. der Schätzung kleiner Quantile der Festigkeit, d. h. die Festigkeit bei geringer Ausfallwahrscheinlichkeit, wird nicht untersucht.

3.2 Ziele der eigenen Untersuchung

Ziel der eigenen Untersuchung ist es daher, durch Simulationsversuche die Treffsicherheit der oben beschriebenen Auswerteverfahren für Treppenstufen- und Horizontversuche in einheitlicher Art und Weise zu bestimmen. Dabei soll untersucht werden, wie sich eine systematische Startwertverschiebung auf die Treffsicherheit auswirkt und welchen Einfluss die Wahl der Anschnittsdefinition auf die Ergebnisse besitzt. Um anhand der ermittelten Ergebnisse die Angaben aus der Literatur besser einschätzen zu können, werden als Maß für die Probenanzahl der Versuchsreihe die Anzahl der im Sinne der Auswertung gültigen Versuche herangezogen. Untersucht werden sollen zum einen die Treffsicherheiten bei der Schätzung des Erwartungswertes sowie der Standardabweichung und zum anderen die Treffsicherheit der Schätzung des 2,5%-Quantils der Verteilung. Letzteres ist der Schätzwert für die Festigkeit bei einer hohen Überlebenswahrscheinlichkeit von $P_U = 97,5\%$ und für die Bauteildimensionierung von entscheidender Bedeutung.

3.3 Beschreibung des Simulationsablaufes

Um die Qualität der gewonnenen Ergebnisse aus den Simulationsversuchen einschätzen zu können, ist es wichtig, den Simulationsablauf detailliert darzustellen. Das betrifft nicht nur die Wahl der Simulationsparameter, sondern ebenso die Art und Weise der virtuellen Versuchsführung. Der Ablauf der Untersuchung gliedert sich grob in drei Teilschritte. Zunächst wird eine große Zahl an Treppenstufenversuchen für jeweils verschiedenen Eingangsparemkombinationen simuliert und ausgewertet. Anschließend werden diese Ergebnisse für jeden Simulationspunkt getrennt statistisch ausgewertet. Abschließend erfolgt die Aufbereitung dieser statistischen Auswertung. Eine detailliertere Darstellung des Vorgehens ist in Abbildung 3.1 skizziert. Auf die darin dargestellten Einzelheiten wird im Folgenden genauer eingegangen.

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

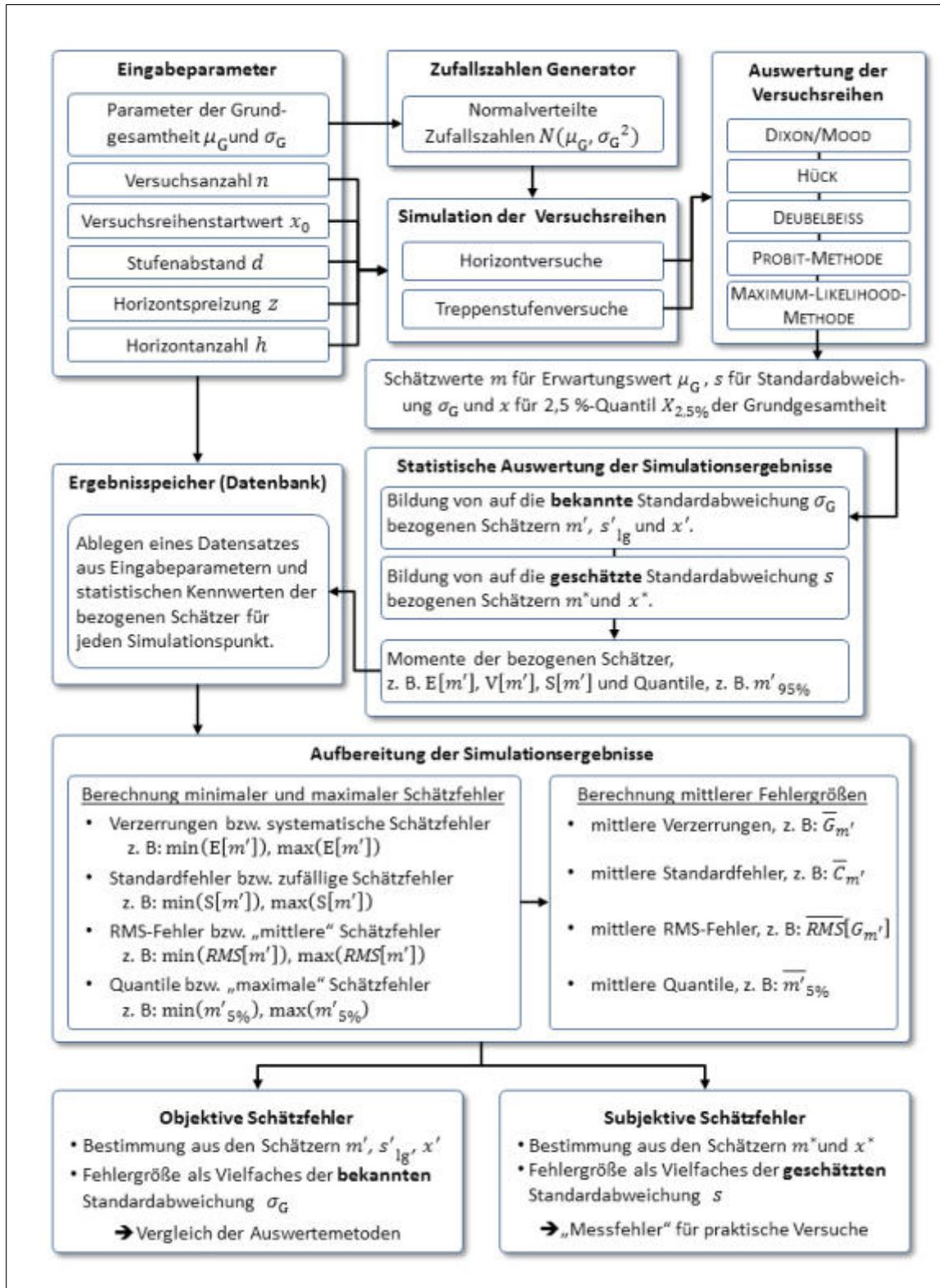


Abb. 3.1: Schematischer Ablauf der Durchführung und Auswertung der Simulationsversuche.

3.3.1 Simulation der Versuchsreihen

Wie bei der praktischen Durchführung sind auch bei der Simulation von Treppenstufenversuchen ein Startwert der Beanspruchung x_0 , der Stufenabstand d und die Anzahl der zu erhaltenden gültigen Versuche n bei der vorgegebenen Anschnittsdefinition festzulegen³. Außerdem ist die statistische Verteilung der zu untersuchenden Größe, in diesem Fall die Schwingfestigkeit, durch den Verteilungstyp sowie die nötigen Parameter zu charakterisieren. Alle Untersuchungen werden mit einer Normalverteilung und deren Verteilungsparametern, dem Erwartungswert μ_G und der Varianz σ_G^2 durchgeführt. Mit Hilfe der AMD Core Math Library (ACML)⁴ wird hierzu mit dem Mersenne-Twister-Algorithmus, einem Pseudozufallszahlengenerator mit einer Periode von $2^{19,937} - 1$, eine Reihe von gleichverteilten Pseudo-Zufallszahlen erzeugt und daraus mit Hilfe der vorgegebenen Parameter der Verteilung eine Reihe von entsprechenden Zufallszahlen erstellt. Diese Zufallszahlen stehen dabei für die Schwingfestigkeiten der einzelnen Proben. Das Ergebnis des „virtuellen“ Versuchs, also ob Bruch oder Durchläufer bzw. Erfolg oder Misserfolg, ergibt sich durch den Vergleich der Schwingfestigkeit der Probe mit einem vorgegebenen Wert für die aufgebrachte Beanspruchung. Je nach Ergebnis des Vergleichs wird die Beanspruchung des nächsten Versuchs um eine Stufe erhöht oder vermindert, also ein größerer oder kleinerer Vorgabewert gewählt. Ausgehend vom vorgegebenen Startwert werden so nacheinander die verschiedenen zufälligen Schwingfestigkeiten der Proben mit den festgelegten Beanspruchungen verglichen, bis die vorgegebene Anzahl an gültigen Versuchen erreicht ist. Für die Simulation der nächsten Versuchsreihe wird als Beanspruchung wieder der Startwert x_0 und für die Schwingfestigkeit der Proben die nächste noch nicht verwendete Zahl aus der Reihe der ermittelten Zufallszahlen gewählt. Dies geschieht so lange, bis die vorgegebene Anzahl an zu simulierenden Versuchsreihen erreicht ist.

Anschließend wird die simulierte Versuchsreihe nach dem gewählten Auswertalgorithmus (nach DIXON/MOOD, nach HÜCK, nach DEUBELBEISS, der PROBIT-METHODE sowie der MAX.-LIKELIHOOD-METHODE) ausgewertet und das Ergebnis, nämlich die Schätzer für eine mittlere Schwingfestigkeit und die Standardabweichung der Schwingfestigkeit festgehalten. Andere Ergebnisse wie z. B. die Anzahl der Versuche im Anschnitt des Treppenstufenversuches n_u , die Anzahl der mit gültigen Versuchen belegten Lasthorizonte n_{Hor} oder die Anzahl der Versuche auf dem untersten und obersten Lasthorizont, auf denen nur Durchläufer bzw. Brüche ermittelt wurden, werden für weitere Untersuchungen ebenfalls dokumentiert.

Für die Simulation von Horizontversuchen werden neben den schon bekannten Parametern x_0 , d und n noch die Anzahl der Horizonte h und deren Spreizung z als Eingangsparameter benötigt. Die Bestimmung der Pseudo-Zufallszahlen für die Festigkeiten der einzelnen Proben erfolgt auf die gleiche Weise wie bei Treppenstufenversuchen. Ebenso beginnt jede Versuchsreihe auf einem festgelegten Startwert x_0 und die Beanspruchung wird solange um den Stufenabstand d vermindert oder erhöht, bis erstmalig das entsprechende Gegenereignis eintritt. Auf dem so ermittelten ersten Lasthorizont wird eine Teilmenge der zu Verfügung

³Die bei einer praktischen Durchführung von Treppenstufenversuchen ebenfalls festzulegende Grenzlastwechselzahl N_G wird nicht benötigt, da sich das Ergebnis des virtuellen Einzelversuchs einzig daraus ergibt, ob die Festigkeit der „virtuellen“ Probe größer oder kleiner der vorgegebenen Beanspruchung ist.

⁴AMD Core Math Library (ACML). Version 5.2.0, Advanced Micro Devices, Inc., Numerical Algorithms Group Ltd., 2012

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

Tabelle 3.1: Übersicht über Eingabegrößen der Treppenstufenversuche.

Eingabegröße	Formel	Vorgegebene Werte
Verteilung der Grundgesamtheit	$X \sim \mathcal{N}(\mu_G; \sigma_G^2)$	$\mathcal{N}(100; 100), \mathcal{N}(100; 49)$
bezogener Startwert	$x'_0 = \frac{x_0 - \mu_G}{\sigma_G}$	{0,0; 0,126; 0,253; 0,385; 0,524; 0,674; 0,842; 1,036; 1,282; 1,645; 1,960; 2,326; 2,576; 3,090; 3,719; 4,265}
bez. Stufenabstand	$d' = \frac{d}{\sigma_G}$	{0,126; 0,253; 0,385; 0,524; 0,674; 0,842; 1,036; 1,282; 1,645; 1,960; 2,326; 2,576; 3,090; 3,719; 4,265; 4,753}
Anzahl nötiger gültiger Versuche	n	{5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 17; 19; 21; 24; 29; 39; 49; 99}
Anschnittsdefinition		{nach HÜCK; nach DIXON}

Tabelle 3.2: Übersicht über Eingabegrößen der Horizontversuche.

Eingabegröße	Formel	Vorgegebene Werte
Verteilung der Grundgesamtheit	$X \sim \mathcal{N}(\mu_G; \sigma_G^2)$	$\mathcal{N}(100; 100)$
bezogener Startwert	$x'_0 = \frac{x_0 - \mu_G}{\sigma_G}$	{0,0; 0,2; 0,5; 0,8; 1,2; 1,6; 2,0; 2,5; 3,0; 4,0}
bez. Stufenabstand	$d' = \frac{d}{\sigma_G}$	{0,25; 0,35; 0,50; 0,75; 1,00; 1,25; 1,55; 1,90; 2,40; 3,00}
bez. Horizontspreizung	$z' = \frac{z}{\sigma_G}$	{0,5; 1,0; 1,5; 2,0; 2,5; 3,0; 3,5; 4,0; 5,0}
Anzahl Horizonte	h	{2; 3; 4; 5; }
Anzahl nötiger gültiger Versuche	n	{10; 12; 15; 18; 21; 25; 29; 34; 39; 44; 49; 99}

stehenden Proben geprüft. Je nach Ergebnis dieser Prüfung berechnet sich mit der vorgegebenen Spreizung z die Lage der weiteren Beanspruchungshorizonte, auf denen die verbliebenen Proben aufgeteilt werden. Auf diese Weise wird die vorgegebene Anzahl an Versuchsreihen simuliert und im Anschluss entweder nach der MAX.-LIKELIHOOD-METHODE oder der PROBIT-METHODE ausgewertet.

3.3.2 Ein- und Ausgabeparameter

Die für die Simulation der Treppenstufenversuche verwendeten Eingangsparameter sind in Tabelle 3.1 zusammengefasst. Für eine Vergleichbarkeit der Ergebnisse erfolgt die Darstellung immer bezogen auf die Verteilungsparameter der Grundgesamtheit. Die Werte für den bezogenen Stufenabstand sowie den bezogenen Startwert orientieren sich an den Quantilen

Tabelle 3.3: Übersicht über Ergebnisgrößen aus Treppenstufen- und Horizontversuchen.

Ergebnisgröße	Formel
bezogener Schätzer Erwartungswert	$m' = \frac{m - \mu_G}{\sigma_G}$
bezogener Schätzer Standardabweichung	$s' = \frac{s}{\sigma_G}$
auf s bez. Schätzer Erwartungswert	$m^* = \frac{m - \mu_G}{s}$
bezogener Schätzer 2,5%-Quantil	$x' = \frac{(m - 1,96s) - (\mu_G - 1,96\sigma_G)}{\sigma_G}$
auf s bez. Schätzer 2,5%-Quantil	$x^* = \frac{(m - 1,96s) - (\mu_G - 1,96\sigma_G)}{s}$
Anzahl Versuche Anschnitt	n_u
Anzahl Horizonte (nur Treppenst.vers.)	n_{Hor}

der Standardnormalverteilung, um möglichst den Gesamtbereich der zugrunde gelegten Verteilung der Festigkeit zu überdecken. So sind die größten Startwerte mit $x'_0 = 3,719$ und $x'_0 = 4,265$ so gewählt, dass die Ausfallwahrscheinlichkeit der Grundgesamtheit bei diesem Lastniveau gerade 99,99 % bzw. 99,999 % beträgt. Für die Horizontversuche ist zusätzlich zu den bereits aus den Treppenstufenversuchen bekannten Parametern noch die Horizontspreizung als Eingangsgröße nötig. Die für die Simulation der Horizontversuche nötigen Eingabeparameter sind in Tabelle 3.2 dargestellt.

Für beide Versuchsmethoden werden die ermittelten Schätzer m und s mit den tatsächlichen Parametern der normalverteilten Grundgesamtheit normiert und können so unabhängig von den für die Simulation gewählten Verteilungsparametern betrachtet werden. Die bezogenen Schätzer m' und s' geben durch ihre jeweilige Abweichung vom Wert 0 oder 1 den Fehler der Schätzung in Vielfachen der tatsächlichen Standardabweichung σ_G wieder. Der bezogene Schätzer m^* drückt die Verschätzung bei der Bestimmung des Erwartungswertes der Grundgesamtheit in Vielfachen des im selben Versuch ermittelten Schätzers der Standardabweichung der Grundgesamtheit aus. Ebenso gibt der bezogene Schätzer x' die Differenz zwischen dem geschätzten und dem tatsächlichen Quantil für eine Ausfallwahrscheinlichkeit von 2,5% in Vielfachen der tatsächlichen Standardabweichung an, während der Schätzer x^* die Differenz zwischen dem geschätzten und dem tatsächlichen Quantil in Vielfachen der im selben Versuch geschätzten Standardabweichung angibt. Alle Größen mit der Bezeichnung „ $'$ “ sind also auf die tatsächliche Standardabweichung bezogen, alle Größen mit „ $*$ “ entstehen durch Normierung auf die im selben Versuch ermittelte Standardabweichung. Weiterhin wird die Anzahl der ungültigen Versuche, also jener Versuche die zwar geprüft, aber nicht in die Auswertung einbezogen werden, ausgegeben. Für die Auswertung der Treppenstufen wird außerdem die Anzahl der mit Versuchen belegten Horizonte festgehalten. Eine Übersicht der erfassten Größen liefert die Tabelle 3.3.

3.3.3 Statistische Auswertung der Simulationsergebnisse

Für jede mögliche Kombination der angegebenen Eingangsgrößen⁵ werden für Treppenstufenversuche jeweils 20.000 Versuchsreihen und für Horizontversuche jeweils 40.000 Versuchsreihen⁶ simuliert und nach den in Abschnitt 2.2.2 beschriebenen Auswerteverfahren nach DIXON/MOOD, HÜCK, DEUBELBEISS, der PROBIT-METHODE sowie der MAX.-LIKELIHOOD-METHODE ausgewertet.⁷ Für jeden Simulationspunkt liegt dann eine Vielzahl von Werten für die verschiedenen, oben beschriebenen, bezogenen Schätzer vor. Diese Werte sind jeweils eine durch die Simulation gezogene Stichprobe einer unbekanntes Grundgesamtheit und werden durch ihre Momente und empirischen Quantile charakterisiert. Es werden also für jede Stichprobe die Größen: Mittelwert bzw. Erwartungswert $\mathbf{E}[\cdot]$, Varianz $\mathbf{V}[\cdot]$ bzw. Standardabweichung $\mathbf{S}[\cdot] = \sqrt{\mathbf{V}[\cdot]}$, Schiefe $\mathbf{H}[\cdot]$, Wölbung $\mathbf{W}[\cdot]$ sowie die empirischen Quantile für 5 %, 10 %, 50 % (Median $\mathbf{M}[\cdot]$), 90 % und 95 % bestimmt, wie es in Abschnitt 2.1.2 dargestellt ist. Aus den berechneten Stichprobenmomenten können systematische und zufällige Schätzfehler abgeleitet werden. Aus den ermittelten Quantilen werden empirische Vertrauensbereiche für die einzelnen Schätzgrößen bestimmt.

Die ermittelten Stichprobenkennwerte, also die Momente und Quantile der jeweiligen Schätzwerte in jedem Simulationspunkt, werden in einer Datenbank abgelegt. Auf die Möglichkeit, nicht nur die Stichprobenkennwerte, sondern jedes Einzelergebnis in der Datenbank abzulegen, wird in Anbetracht des dafür nötigen Speicherplatzes verzichtet.⁸

3.3.4 Ermittlung von Korrelationskoeffizienten

Der Versuch, durch Betrachtung von Korrelationskoeffizienten zwischen den abgelegten Stichprobenkennwerten und den Eingabeparametern der Simulation, direkt die zu erwartenden Treffsicherheiten abzuleiten, war nicht erfolgreich. Auch durch Kombinationen oder gängige Transformationen, wie Quadrieren oder Logarithmieren der verschiedenen Eingangsgrößen und Stichprobenkennwerte, konnten keine Korrelationskoeffizienten ermittelt werden, die einen Rückschluss auf kausale Zusammenhänge ermöglicht hätten. Die Ursache hierfür ist vor allem in dem recht großen Simulationsgebiet sowie den Wechselwirkungen zwischen vorgegebenem Stufenabstand und Startwert der Versuchsreihen zu suchen.⁹ Ersteres könnte durch Verkleinerung und damit einhergehender Verringerung der allgemeingültigen Aussagekraft beseitigt werden. Bei letzteren kommt es durch günstige bzw. ungünstige Wahl der Kombination

⁵= 9.728 Parameterkombinationen für Treppenstufen-, 43.200 für Horizontversuche

⁶= 194.560.000 Versuchsreihen für Treppenstufen-, 1.728.000.000 Horizontversuche

⁷= 972.800.000 Auswertungen für Treppenstufen-, 3.456.000.000 für Horizontversuche

⁸Der kleinstmögliche Datensatz für ein Einzelergebnis besteht aus mindestens 2 Gleitkommazahlen für m und s zu je 4 Byte und 3 ganzzahligen Werten für n , n_{Hor} und n_{u} zu je 3 Byte. Damit ergibt sich eine Mindestgröße der Datenbank von mehr als 75 GB. Sollen zusätzlich noch weitere Größen, z. B. die Varianzkennzahl k bei Treppenstufenversuchen oder auf die Verteilungsparameter bezogene Größen abgelegt werden, verdoppelt oder verdreifacht sich die Datenmenge.

⁹Sowohl eine sehr kleiner Stufenabstand als auch ein sehr großer Stufenabstand führen zu weniger treffsicheren Ergebnissen. Außerdem ergeben sich z. B. bei den Auswertungen nach DIXON/MOOD, wenn Lasthorizonte genau mit dem Erwartungswert der Grundgesamtheit zusammenfallen andere Treffsicherheiten, als wenn zwei Horizonte den Erwartungswert der Grundgesamtheit mittig einschließen, vgl. Abbildung 2.8. Ein ähnliches Verhalten ist auch bei den anderen Auswertemethoden zu erwarten.

immer wieder zu sprunghaften Veränderungen der ermittelten Stichprobenkennwerte. Um diese genauer zu beschreiben, müsste das Simulationsgebiet mit mehr Simulationspunkten ausgefüllt werden, ohne Garantie die Ergebnisse anschließend besser beurteilen zu können.

3.4 Aufbereitung der Simulationsergebnisse

Für die weitere Auswertung werden die ermittelten Stichprobenkennwerte einzeln und in Abhängigkeit der Eingabeparameter betrachtet und grafisch dargestellt. Es werden dabei die in der Simulation maximal und minimal auftretenden Stichprobenkennwerte, sowie daraus gemittelte Stichprobenkennwerte ermittelt. Da die betrachteten bezogenen Schätzer zum einen durch Normierung auf die tatsächliche Standardabweichung und zum anderen durch die Normierung auf die im selben Versuch ermittelte Standardabweichung berechnet wurden, können aus den Ergebnissen verschiedene Erkenntnisse gewonnen werden. Aus den zuerst genannten bezogenen Schätzern können objektive Schätzfehler ermittelt werden, da hier die Basis der Normierung immer derselbe Wert der bekannten oder als bekannt vorausgesetzten Standardabweichung ist. Die aus diesen Größen abgeleiteten Schätzfehler können also benutzt werden, um die verschiedenen Auswerteverfahren untereinander zu vergleichen. Im Gegensatz dazu können die Schätzfehler, die aus den mittleren Stichprobenkennwerten der auf die im selben Versuch ermittelten Standardabweichung bezogenen Schätzer gebildet werden nur eine subjektive Einschätzung der Fehler liefern. Die Größe der so berechnete Schätzfehler ist immer von der im selben Versuch beobachteten Standardabweichung abhängig. Ein Vergleich der verschiedenen Auswertemethoden anhand dieser Fehler ist nicht möglich.

Wie die Ergebnisse im Einzelnen ermittelt werden, wird im Folgenden anhand von Versuchsreihen mit $n = 14$ oder $n = 15$ Einzelversuchen exemplarisch erläutert. Alle anderen Ergebnisse sind im Anhang in Form von grafischen Darstellungen zusammengefasst. Auf die aus den gesammelten Ergebnissen ableitbaren Aussagen für die Unsicherheiten der ermittelten Parameter der Festigkeitsverteilung wird erst im Kapitel 4 eingegangen. Es wird gezeigt, dass sich die Veränderung des Versuchsreihenstartwertes x_0 nur im begrenzten Maß auf die Schätzfehler auswirkt. Die Lage des in praktischen Versuchen unbekanntes Startwertes wird durch mittlere Schätzfehler erfasst. Die untersuchte Wahl der Anschnittsdefinition besitzt, wie sich im Laufe der Arbeit herausstellt, nur einen geringen Einfluss auf die ermittelten Fehlergrößen. Daher werden die Simulationsergebnisse für die Untersuchung der Anschnittsdefinitionen nach DIXON nicht im Anhang aufgeführt. Bei der im Abschnitt 4 folgenden Interpretation der Ergebnisse werden sie aber wieder berücksichtigt.

3.4.1 Vom Auswerteverfahren unabhängige Ergebnisse

3.4.1.1 Anzahl belegter Horizonte in Treppenstufenversuchen

Für jeden Simulationspunkt wird die Anzahl der Versuchsreihen bei einer festgelegten Anzahl von Horizonten ermittelt und auf die Gesamtanzahl an Versuchsreihen dieses Simulationspunktes bezogen. Die so ermittelten relativen Häufigkeiten, mit denen bei der Simulation der Treppenstufenversuche eine Versuchsreihe mit einer bestimmten Horizontanzahl erreicht wird,

sind in den Abbildungen A.1 bis A.6 im Anhang ab S. 180 sowie für $n = 14$ in der Abbildung 4.1 auf S. 96 grafisch dargestellt.

3.4.1.2 Anzahl ungültiger Versuche einer Versuchsreihe

Die Anzahl n_u , der beim **Treppenstufenversuch** als ungültig betrachteten, also im Anschnitt der Versuchsreihe liegenden Einzelversuche, wird analog zum obigen Vorgehen auf die Gesamtanzahl der im jeweiligen Simulationspunkt vorliegenden Anzahl der Versuchsreihen bezogen. Für die Anschnittsdefinition nach HÜCK sind die Ergebnisse in den Abbildungen A.7 bis A.12 im Anhang ab S. 186 sowie für $n = 14$ in der Abbildung 4.2 auf S. 98 angegeben.

Bei **Horizontversuchen** sind die Versuche, die zum Auffinden des ersten Lasthorizontes verwendet werden, keine ungültigen Versuche im Sinne der Auswertung. Entweder können sie mit der Begründung, dass sie dort ebenfalls versagt bzw. nicht versagt hätten, einem höheren bzw. niedrigeren Horizont zugeordnet werden, oder sie können als einzelne Versuche auf jeweils neuen Horizonten formal in die Auswertung einbezogen werden. Im letzteren Fall liefern diese Versuche aber keinerlei zusätzliche Information für die Schätzung der Verteilungsparameter und sind in diesem Sinne unnütz bzw. ungültig. Es ist davon auszugehen, dass der Anteil unnützer Proben in etwa genauso groß ausfällt, wie bei der Auswertung von Treppenstufenversuchen, da bei beiden Versuchsführungen die anfänglichen Versuche dazu genutzt werden müssen, den Bereich der mittleren Festigkeit zu finden. Sowohl bei Horizontversuchen als auch bei Treppenstufenversuchen mit der Anschnittsdefinition nach HÜCK werden Versuche, deren Lasthöhe zu einem späteren Zeitpunkt der Versuchsreihe von einer anderen Probe wieder erreicht wird, als gültig bewertet.

3.4.2 Auswertung der Schätzung des Erwartungswertes bei bekannter Standardabweichung

3.4.2.1 Anteil auswertbarer Versuchsreihen

Bis auf die Auswertung nach der PROBIT-METHODE liefern alle Auswertemethoden der **Treppenstufenversuche** stets Schätzwerte für den Erwartungswert der Grundgesamtheit. Da für die Ziehung der Versuchsreihen nicht die Anzahl der zur Verfügung stehenden Proben, sondern die Anzahl der zu erreichenden gültigen Einzelversuche vorgegeben wird, entstehen keine ungültigen Versuchsreihen und es kann immer wenigstens der Erwartungswert der Grundgesamtheit geschätzt werden. Nur bei der PROBIT-METHODE werden beide Parameter der Grundgesamtheit gemeinsam im Wahrscheinlichkeitsnetz geschätzt. Als gültige Schätzungen des Erwartungswertes werden nur jene betrachtet, die gleichzeitig eine positive Standardabweichung schätzen. Für die Auswertung der Treppenstufenversuche nach der PROBIT-METHODE wird daher die Anzahl der auswertbaren Versuchsreihen in jedem Simulationspunkt auf die Anzahl der in diesem Punkt simulierten Versuchsreihen bezogen. Es ergeben sich je nach Lage des Versuchsreihenstartwertes recht unterschiedliche Anteile auswertbarer Versuchsreihen. Daher werden alle ermittelten Anteile für einen konstanten bezogenen Stufenabstand d' und einer vorgegebenen Versuchsanzahl n , aber verschiedenen Versuchsreihenstartwerten x'_0 zusammengefasst und der ungünstigste in den Simulationen

auftretende Fall, also jeweils der minimale Anteil bestimmt. Die so ermittelten Anteile auswertbarer Treppenstufenversuchsreihen sind für die Schätzung des Erwartungswertes und der Standardabweichung identisch und werden daher in Abschnitt 3.4.3.1 auf S. 73 mit dargestellt.

Auch bei der Auswertung von **Horizontversuchen** liefert die Auswertung mittels PROBIT-METHODE nicht in jedem Fall Schätzwerte für den Erwartungswert der Grundgesamtheit. Die Anzahl der nicht auswertbaren Versuchsreihen für die Schätzung des Erwartungswertes ist wieder mit den Ergebnissen der Standardabweichungsschätzung identisch und Abschnitt 3.4.3.1 zu entnehmen. Mit der MAX.-LIKELIHOOD-METHODE werden stets Erwartungswerte der Grundgesamtheit geschätzt.

3.4.2.2 Systematischer und zufälliger Schätzfehler

Ist die Standardabweichung σ_G der Grundgesamtheit bzw. der Schwingfestigkeit bekannt, dann kann mit Hilfe des bezogenen Schätzers m' die Abweichung der Schätzung in jedem Simulationspunkt angegeben werden. Mit dem geschätzten Erwartungswert $\mathbf{E}[m']$ wird die systematische Abweichung und mit der Standardabweichung $\mathbf{S}[m']$ die zufällige Abweichung jeweils in Vielfachen der tatsächlichen Standardabweichung σ_G angegeben. Um begriffsbedingte Verwirrungen zu vermeiden wird der Erwartungswert des Schätzfehlers oft auch als Bias oder als Verzerrung des Schätzers bezeichnet. Die Standardabweichung des Schätzers wird gern Standardfehler der Schätzung genannt. Im Sinne einer Messung einer physikalischen Größe muss das Messergebnis, also der geschätzte Erwartungswert m , um den Bias korrigiert werden. Die zufällige und nicht korrigierbare Unsicherheit der Messung wird durch den Standardfehler ausgedrückt und muss für eine Messung der Vollständigkeit halber mit angegeben werden. Das korrigierte, vollständige Messergebnis m_k lautet:

$$m_k = m - \mathbf{E}[m'] \cdot \sigma_G \pm \mathbf{S}[m'] \cdot \sigma_G . \quad (3.1)$$

Die aus der Simulation ermittelten $\mathbf{E}[m']$ und $\mathbf{S}[m']$ variieren in Abhängigkeit der für die Simulation gewählten Eingangsparameter x'_0 , d' und n . Wie in Abbildung 3.2 exemplarisch für eine Auswertung und Anschnittsdefinition nach HÜCK dargestellt, ändern sich sowohl Bias als auch Standardfehler mit den gewählten Startwerten.¹⁰ Bei kleinen Stufenabständen vergrößert sich der systematische Fehler gleichmäßig mit zunehmender Startwertverschiebung. Im Bereich großer Stufenabstände ist die Genauigkeit der Schätzung davon abhängig, ob für die gewählten d' und x'_0 der Erwartungswert μ_G mit einem Lashorizont übereinstimmt oder genau dazwischen zum Liegen kommt. Diese Beobachtungen treffen auch bei Simulationsergebnissen mit der Anschnittsdefinition nach DIXON (Abbildung 3.2, c und d) und bei anderen Auswertemethoden, wie sie in Abbildung 3.3 exemplarisch dargestellt sind, zu. In jedem Fall können für beide Fehlerarten untere und obere Grenzen $G_{m,\min}$ und $G_{m,\max}$ bzw. $C_{m,\min}$ und $C_{m,\max}$ beschrieben werden, innerhalb deren sich alle Werte befinden. Hierzu werden bei festgehaltenem Stufenabstand und fester Anzahl gültiger Versuche, aber beliebiger Startwertverschiebungen, die Maximalwerte und Minimalwerte aller ermittelten $\mathbf{E}[m']$ bzw.

¹⁰Jede Strichlinie in Abbildung 3.2 zeigt das Simulationsergebnis für einen anderen Startwert x'_0 .

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

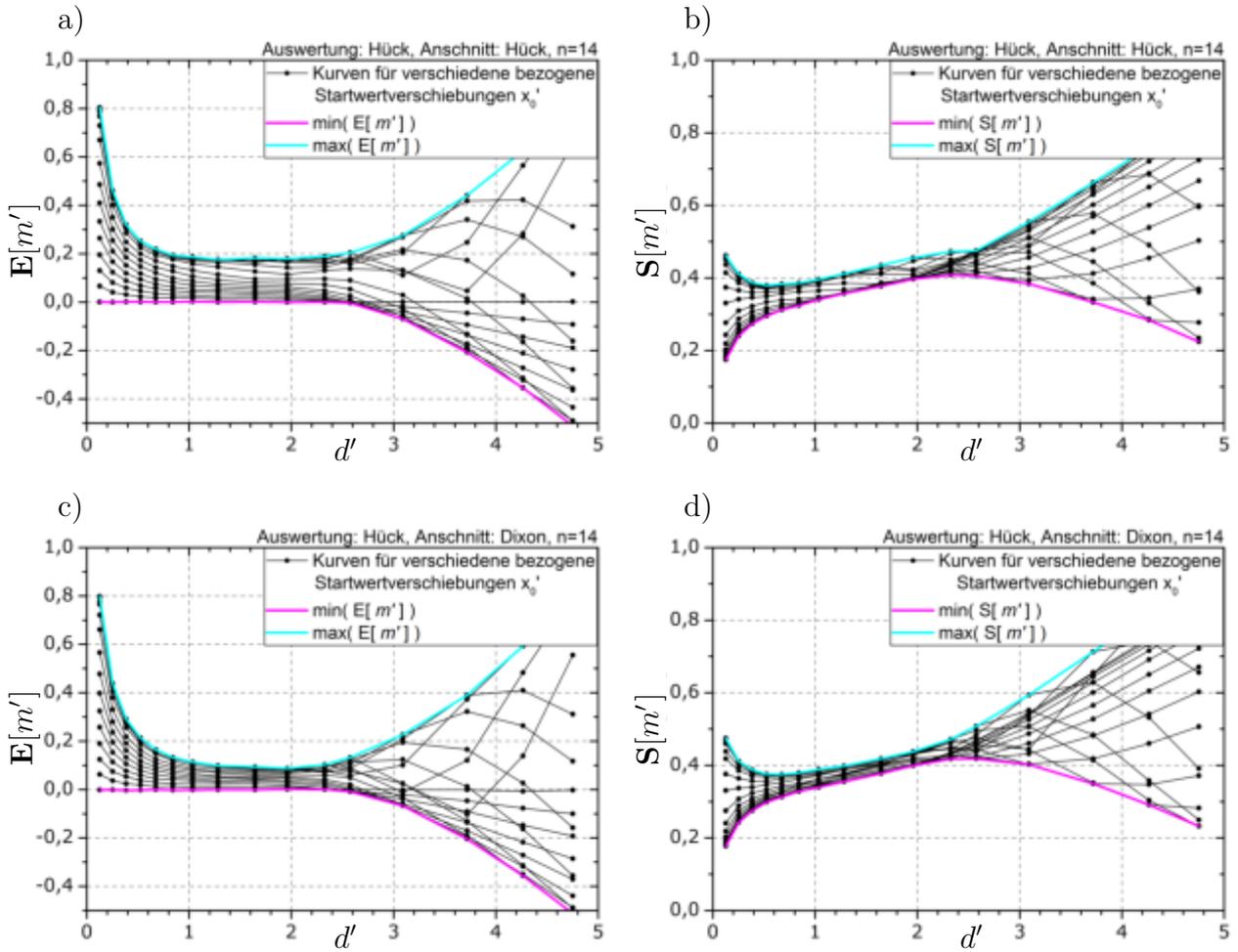


Abb. 3.2: Erwartungswert und Standardfehler des Schätzers m' für Treppenstufenversuche mit $n = 14$ gültigen Versuchen, ausgewertet nach HÜCK für unterschiedlich¹⁰ gewählte Startwerte x_0 und eine Anschnittsdefinition nach HÜCK in a) und b) oder nach DIXON in c) und d).

$S[m']$ bestimmt, also:

$$\begin{aligned}
 G_{m',\min}(d', n) &= \min_{\forall x'_0}(\mathbf{E}[m'(d', n, x'_0)]), \\
 G_{m',\max}(d', n) &= \max_{\forall x'_0}(\mathbf{E}[m'(d', n, x'_0)]), \\
 C_{m',\min}(d', n) &= \min_{\forall x'_0}(\mathbf{S}[m'(d', n, x'_0)]), \\
 C_{m',\max}(d', n) &= \max_{\forall x'_0}(\mathbf{S}[m'(d', n, x'_0)]).
 \end{aligned}$$

Welche Werte sowohl der systematische als auch zufällige Schätzfehler annehmen, hängt von der tatsächlichen Kombination aus x'_0 und d' ab und ist in praktischen Versuchen unbekannt. Es erscheint daher zweckmäßig, einem mittleren systematischen Fehler nach Gleichung 3.2 anzunehmen und die Unsicherheit über den tatsächlichen systematischen Fehler der mittleren zufälligen Messunsicherheit nach Gleichung 3.3 zuzuschlagen. Dies ist gleichbedeutend mit der Behauptung, dass eine zufällige und unbekannte Startwertverschiebung den systematischen

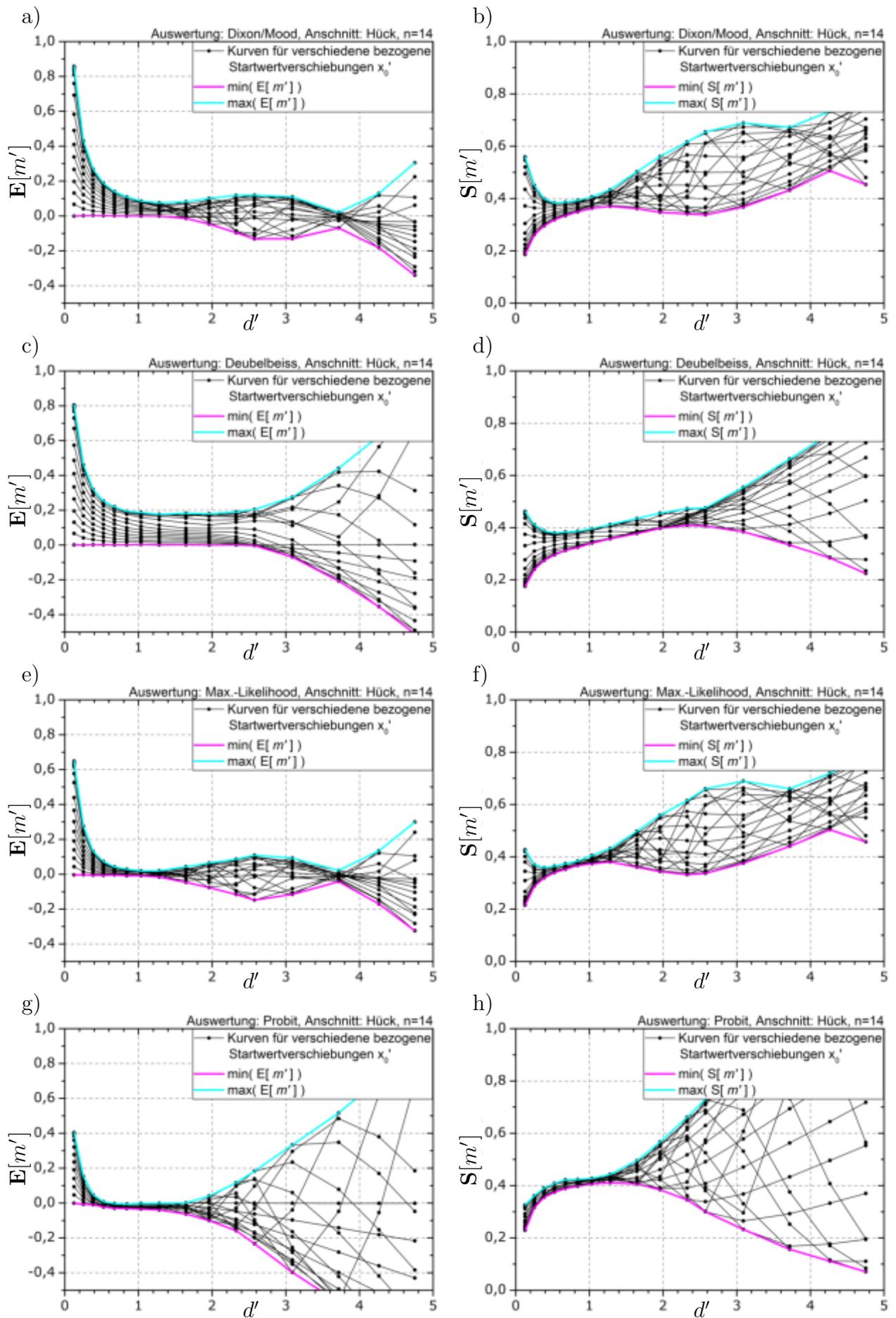


Abb. 3.3: Erwartungswert und Standardfehler des Schätzers m' für Treppenstufenversuche mit $n = 14$ für unterschiedlich¹⁰ gewählte Startwerte x_0 und Anschnittsdefinition nach HÜCK. Auswertemethode nach DIXON/MOOD in a) und b), DEUBELBEISS in c) und d), MAX.-LIKELIHOOD-METHODE in e) und f) sowie PROBIT-METHODE in g) und h).

Fehler des Schätzers m' gleichmäßig im Intervall $[G_{m',\min}, G_{m',\max}]$ verteilt.¹¹ Die Schwankung des zufälligen Schätzfehlers aufgrund verschiedener Startwertverschiebungen wird vernachlässigt und nur der mittlere zufällige Schätzfehler als Messunsicherheit berücksichtigt. Dies ist natürlich nur sinnvoll, wenn die Unsicherheit des zufälligen Fehlers klein gegenüber dem mittleren zufälligen Fehler ist. Wie in den Abbildungen 3.2 und 3.3 dargestellt, ist dies für die betrachteten Auswerteverfahren, im für die praktischen Anwendung verwendeten Bereich des bezogenen Stufenabstandes von $0,5 < d' < 2$, gegeben.

$$\bar{G}_{m'} = \frac{G_{m',\max} + G_{m',\min}}{2} \quad (3.2)$$

$$\bar{C}_{m'} = \sqrt{\frac{(G_{m',\max} - G_{m',\min})^2}{12} + \frac{(C_{m',\max} + C_{m',\min})^2}{4}} \quad (3.3)$$

Die berechneten mittleren systematischen und zufälligen Fehler $\bar{G}_{m'}$ und $\bar{C}_{m'}$ sind im Anhang in den Abbildungen A.13 bis A.14, Seite 192 ff. dargestellt. Das vollständige Messergebnis kann dann in der Form

$$m_k = m - \sigma_G \cdot \bar{G}_{m'} \pm \sigma_G \cdot \bar{C}_{m'} \quad (3.4)$$

angegeben werde. Dabei muss allerdings beachtet werden, dass die Simulationsrechnungen nur mit positiven bezogenen Startwertverschiebungen durchgeführt wurden. Hierdurch nimmt auch der systematische Schätzfehler im Allgemeinen positive Werte an. Ist in der praktischen Durchführung bekannt, dass der Versuchsreihenstartwert unterhalb des ermittelten Erwartungswertes liegt, dann muss das Ergebnis mit dem entsprechenden negativen Bias korrigiert werden, wie es in der folgenden Gleichung angegeben ist:

$$m_k = m + \sigma_G \cdot \bar{G}_{m'} \pm \sigma_G \cdot \bar{C}_{m'}$$

Auf den Standardfehler der Schätzung hat nur der Betrag der Startwertverschiebung einen geringen und hier vernachlässigten Einfluss.

Auch für die Ergebnisse der **Horizontversuche** zeigt sich bei einem möglichst klein gewählten Stufenabstand ein begrenzter Einfluss der Startwertverschiebung auf die Schätzfehler. Hierbei scheinen sinnvolle Schätzungen nur möglich, wenn auf jedem Lasthorizont eine ausreichende Anzahl an Versuchen vorliegt.¹² Für die Versuchsauswertungen wird daher gefordert, dass es zum einen die vorgegebene Anzahl von Einzelversuchen ermöglichen muss, auf jedem Horizont mindesten 5 Einzelversuche zu platzieren, also $n/h \geq 5$ gilt. Zum anderen soll der Stufenabstand nicht größer als die Standardabweichung der Grundgesamtheit gewählt werden, d. h. $d' < 1$. Die statistische Auswertung der Simulationsergebnisse erfolgt dann nach der

¹¹Eine unbekannter, aber zwischen bekannten Grenzen liegender, systematischer Messfehler erzeugt, wie es in DIN 1319-3 S. 6 [6] beschrieben ist, einen mittleren systematischen Schätzfehler und erhöht die Messunsicherheit des Ergebnisses. Die Größe der zusätzlichen Unsicherheit wird aus der Standardabweichung der Gleichverteilung abgeleitet. Für eine zwischen den Grenzen a und b gleichverteilte Zufallsgröße X ist der Erwartungswert gerade $\mathbf{E}[X] = (a + b)/2$ und die Standardabweichung $\mathbf{S}[X] = (b - a)/\sqrt{12}$.

¹²Bei den durchgeführten Simulationen wird die Lage der Horizonte aus der vorläufig geschätzten Ausfallwahrscheinlichkeit des ersten Lasthorizontes abgeleitet. Alle Horizonte werden etwa mit der gleichen Anzahl an Proben belegt. Ein möglichst kleiner Stufenabstand sorgt dafür, dass der erste Horizont nicht zu weit am Rand der Festigkeitsverteilung liegt. Die Mindestanzahl an Proben ist vor allem dafür nötig, die Ausfallwahrscheinlichkeit des ersten Horizontes gut zu schätzen, damit die anderen Horizonte sinnvoll verteilt werden können.

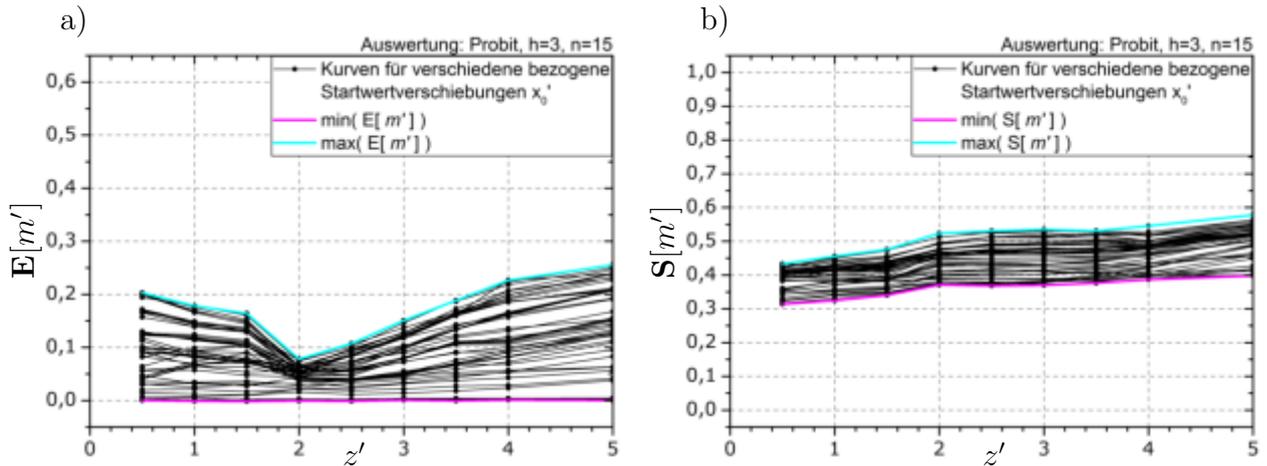


Abb. 3.4: Beispiel für Erwartungswert a) und Standardfehler b) des Schätzers m' aus Horizontversuchen mit $n = 15$ gültigen Versuchen auf $h = 3$ Horizonten, ausgewertet nach der PROBIT-METHODE, für unterschiedlich gewählte Startwerte x_0 .

selben Methodik, wie sie bei der Auswertung der Treppenstufenversuche angewendet wird. Es ist wieder möglich, die minimalen und maximalen auftretenden Schätzfehler zu ermitteln, vgl. Abbildung 3.4.

$$\begin{aligned}
 G_{m',\min}(z', h, n) &= \min_{\forall x'_0, d' \leq 1} (\mathbf{E}[m'(z', x'_0, d', h, n)]), \\
 G_{m',\max}(z', h, n) &= \max_{\forall x'_0, d' \leq 1} (\mathbf{E}[m'(z', x'_0, d', h, n)]), \\
 C_{m',\min}(z', h, n) &= \min_{\forall x'_0, d' \leq 1} (\mathbf{S}[m'(z', x'_0, d', h, n)]), \\
 C_{m',\max}(z', h, n) &= \max_{\forall x'_0, d' \leq 1} (\mathbf{S}[m'(z', x'_0, d', h, n)]).
 \end{aligned}$$

Aus diesen minimalen und maximalen systematischen und zufälligen Schätzfehlern werden mittlere Schätzfehler nach den Gleichungen 3.2 und 3.3 bestimmt. Die damit berechneten mittleren systematischen und zufälligen Fehler des Schätzers m' für die Auswertung von Horizontversuchen sind im Anhang in den Abbildungen A.15 bis A.18, S. 194 ff. dargestellt. Auch hier sind, da alle Simulationen nur mit positiven Startwertverschiebungen durchgeführt wurden, die aufgeführten systematischen Fehler bei einer Versuchsreihe mit einem Startwert kleiner dem Erwartungswert der Grundgesamtheit mit einem negativem Vorzeichen zu versehen.

3.4.2.3 Allgemeine Messunsicherheit

Ist über die Lage des Startwertes einer Versuchsreihe nichts bekannt, muss die Unsicherheit über die tatsächliche systematische Messabweichung, die ja dann sowohl positiv als auch negativ ausfallen kann, dem zufälligen Fehler zugeordnet werden. Hierzu wird wieder die Annahme getroffen, dass der aus einer Startwertverschiebung resultierende Bias des Schätzers m' gleichmäßig über das Intervall

$$[-\max(|G_{m',\max}|, |G_{m',\min}|), \max(|G_{m',\max}|, |G_{m',\min}|)] \quad (3.5)$$

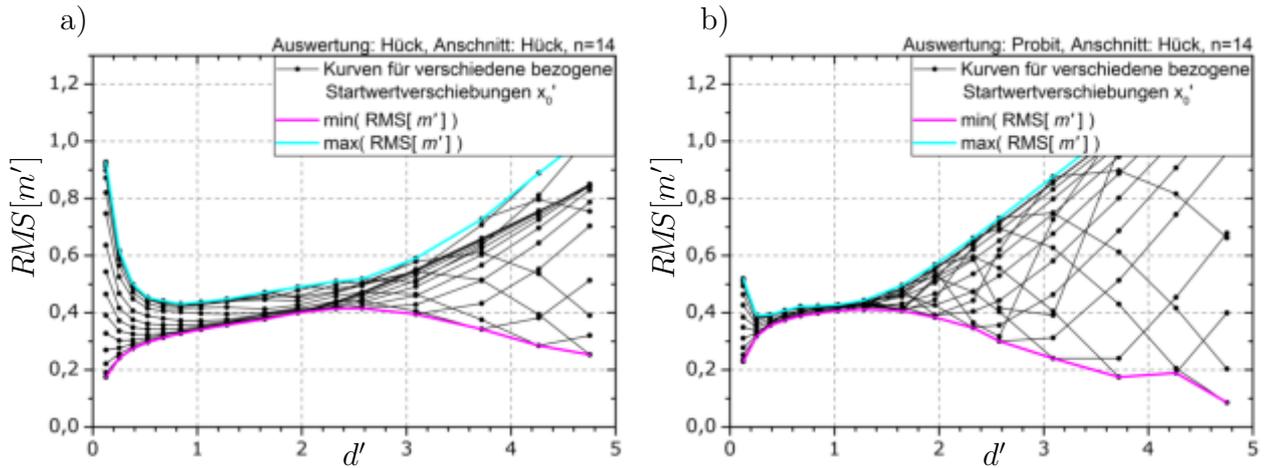


Abb. 3.5: RMS-Schätzfehler des Schätzers m' für Treppenstufenversuche mit $n = 14$ gültigen Versuchen, unterschiedlichen Startwerten x_0' und einer Anschnittsdefinition nach HÜCK. Auswertung nach HÜCK in a) sowie mittels PROBIT-METHODE in b).

verteilt ist. Dieser wird dann mit Gleichung 3.6, analog zu obigem Vorgehen bei der Ermittlung des Standardfehlers, zu einem verallgemeinerten Standardfehler zusammengefasst.

$$U_{m'} = \sqrt{\frac{\max(|G_{m',\max}|, |G_{m',\min}|)^2}{3} + \frac{(C_{m',\max} + C_{m',\min})^2}{4}} \quad (3.6)$$

Das korrigierte vollständige Messergebnis lautet:

$$m_k = m \pm \sigma_G \cdot U_{m'}$$

Die ermittelten verallgemeinerten Standardfehler $U_{m'}$ sind für **Treppenstufenversuche** im Anhang in der Abbildung A.19 auf S. 197 dargestellt. Für die Auswertung der **Horizontversuche** können die ermittelten verallgemeinerten Standardfehler aus den Abbildungen A.20 und A.21, S. 198 ff. im Anhang entnommen werden.

3.4.2.4 RMS-Fehler

Eine andere Möglichkeit die Genauigkeit der experimentell ermittelten mittleren Schwingfestigkeit zu beurteilen, wird durch die Auswertung des mittleren quadratischen Fehlers (MSE) bzw. dessen Quadratwurzel (RMS) nach Gleichung 2.5 erreicht. Damit geht die oben beschriebene Aufteilung des Fehlers in einen systematischen und zufälligen Anteil verloren. Die ermittelten Fehlergrößen sind aber anschaulicher für einen Vergleich zugänglich. Es ist ebenso wie bei der Betrachtung des systematischen und zufälligen Fehlers möglich, ein Intervall anzugeben, in dessen Grenzen sich der quadratische Fehler bei Variation des Versuchsreihenstartwertes befindet, siehe Beispiel in Abbildung 3.5. Die daraus berechneten Intervallmitten

$$\overline{RMS}_{m'} = \frac{RMS_{m',\max} + RMS_{m',\min}}{2}$$

eignen sich dann, um die verschiedenen Auswerteverfahren zu vergleichen. Die Intervallbreiten $\Delta RMS_{m'} = RMS_{m',\max} - RMS_{m',\min}$ beschreiben die Empfindlichkeit, mit der die jeweiligen

Auswerteverfahren auf die Veränderung des Versuchsreihenstartwertes reagieren. Die Mittelwertbildung ist nur dann angebracht, wenn der Intervallmittelwert deutlich größer als die Breite des Intervalls ausfällt. Für die Auswertungen der **Treppenstufenversuche** sind daher in den Abbildung A.22 und A.23 ab S. 200 im Anhang den mittleren RMS-Fehlern die halben Intervallbreiten, bezogen auf die Intervallmittelwerte, gegenübergestellt. Ebenso sind die Ergebnisse für die **Horizontversuche** den Darstellungen A.24 und A.27 ab S. 202 zu entnehmen.

3.4.2.5 Empirischer Vertrauensbereich

Neben den Momenten werden in jedem Simulationspunkt auch die empirischen Quantile des Schätzers m' ermittelt. Zwischen dem linksseitigen Quantil $m'_{5\%}$ und dem rechtsseitigen Quantil $m'_{95\%}$ liegen 90 % aller ermittelten bezogenen Schätzwerte eines Simulationspunktes.

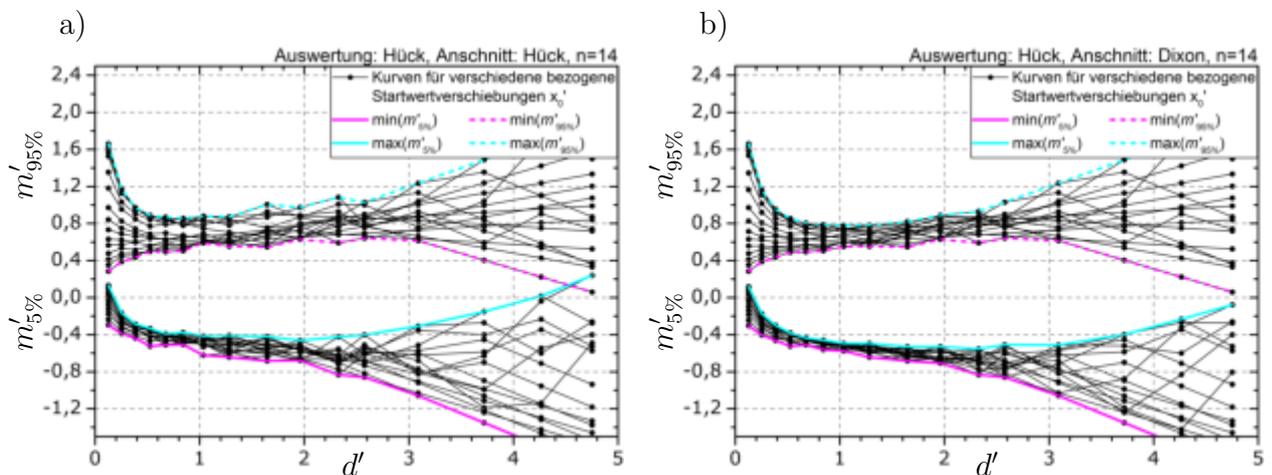


Abb. 3.6: 5%- bzw. 95%-Quantile des Schätzers m' für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und unterschiedlich gewählten Startwerten x'_0 für Anschnittsdefinition nach HÜECK a) und DIXON b).

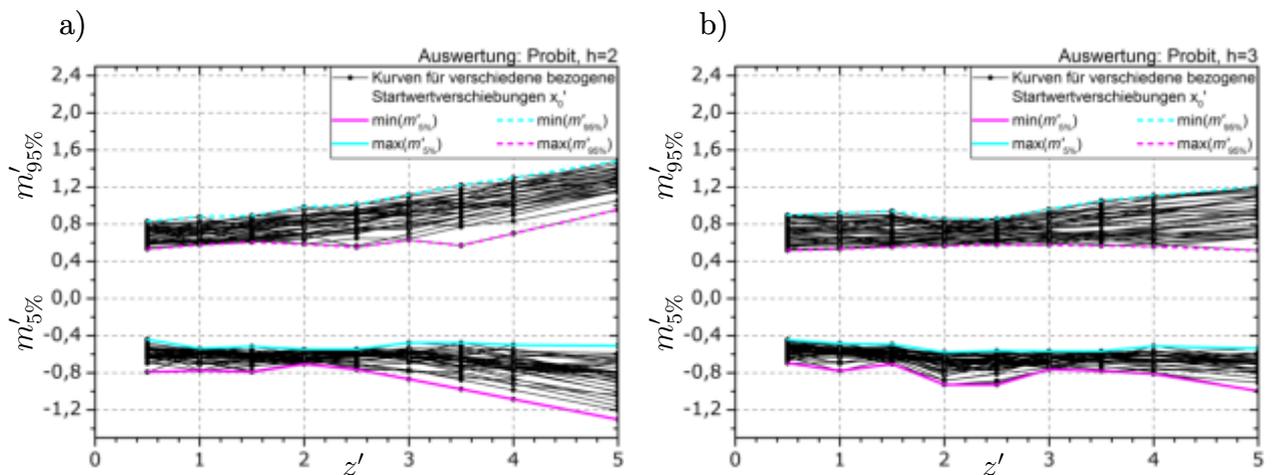


Abb. 3.7: 5%- bzw. 95%-Quantile des Schätzers m' für Horizontversuche mit $n = 15$ gültigen Versuchen und unterschiedlich gewählten Startwerten x'_0 . Auswertung nach der PROBIT-METHODE für Versuche auf 2 Horizonten in a) und 3 Horizonten in b).

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

Wird in einem Experiment der Schätzer m bestimmt, dann liegt der wahre Wert μ_G mit einer relativen Irrtumshäufigkeit von 10 % in dem Intervall $[m - m'_{95\%} \cdot \sigma_G, m - m'_{5\%} \cdot \sigma_G]$.¹³

Sowohl für Auswertungen der Treppenstufenversuche als auch für die Auswertungen der Horizontversuche¹⁴ zeigt sich wieder, wie in Abbildung 3.6 und 3.7 exemplarisch dargestellt, der beschränkte Einfluss der Startwertverschiebung auf die Größe der ermittelten Quantile. Daher werden für die Auswertung der Treppenstufenversuche die minimale und maximale links- und rechtsseitige Quantile ermittelt:

$$\begin{aligned} m'_{5\%,\min}(d', n) &= \min_{\forall x'_0} (m'_{5\%}(d', n, x'_0)), \\ m'_{5\%,\max}(d', n) &= \max_{\forall x'_0} (m'_{5\%}(d', n, x'_0)), \\ m'_{95\%,\min}(d', n) &= \min_{\forall x'_0} (m'_{95\%}(d', n, x'_0)), \\ m'_{95\%,\max}(d', n) &= \max_{\forall x'_0} (m'_{95\%}(d', n, x'_0)). \end{aligned}$$

Ebenso erfolgt die Ermittlung der minimalen und maximalen links- und rechtsseitigen Quantile für die Ergebnisse der Horizontmethode nach:

$$\begin{aligned} m'_{5\%,\min}(z', h, n) &= \min_{\forall x'_0, d' \leq 1} (m'_{5\%}(z', x'_0, d', h, n)), \\ m'_{5\%,\max}(z', h, n) &= \max_{\forall x'_0, d' \leq 1} (m'_{5\%}(z', x'_0, d', h, n)), \\ m'_{95\%,\min}(z', h, n) &= \min_{\forall x'_0, d' \leq 1} (m'_{95\%}(z', x'_0, d', h, n)), \\ m'_{95\%,\max}(z', h, n) &= \max_{\forall x'_0, d' \leq 1} (m'_{95\%}(z', x'_0, d', h, n)). \end{aligned}$$

Mit diesen Maximal- bzw. Minimalwerten können für beide Arten der Versuchsführung mittlere empirische Quantile $\overline{m}'_{5\%}$ und $\overline{m}'_{95\%}$ nach den Gleichungen 3.7 und 3.8 angegeben werden.

$$\overline{m}'_{5\%}(d', n) = \frac{m'_{5\%,\max} + m'_{5\%,\min}}{2} \quad (3.7)$$

$$\overline{m}'_{95\%}(d', n) = \frac{m'_{95\%,\max} + m'_{95\%,\min}}{2} \quad (3.8)$$

Dies ermöglicht es, einen mittleren Vertrauensbereich zu einer Irrtumswahrscheinlichkeit von 10 % durch das folgende Intervall abzuschätzen:

$$[m - \overline{m}'_{95\%} \cdot \sigma_G, m - \overline{m}'_{5\%} \cdot \sigma_G].$$

Wie bereits oben beschrieben, wurden die Simulationsrechnungen ausschließlich mit positiven Startwertverschiebungen durchgeführt. Daher ist bei unbekanntem Startwert der Versuchsreihe für die Bestimmung des Vertrauensbereiches das betragsmäßig größere Quantil zu wählen. Nach Gleichung 3.9 wird mit

$$\overline{m}'_{|5\%|} = \max(|\overline{m}'_{5\%}|, |\overline{m}'_{95\%}|) \quad (3.9)$$

¹³Zu beachten ist, dass mit dem rechtsseitigen Quantil $m'_{95\%}$ die untere Grenze und mit dem linksseitigen Quantil $m'_{5\%}$ die obere Grenze des Vertrauensbereichs gebildet wird.

¹⁴Für Horizontversuche gilt wieder: $n/h \geq 5$ und $d' \leq 1$

ein verallgemeinertes 5 %-Quantil $\overline{m}'_{|5\%|}$ berechnet. Der Vertrauensbereich zu einer Irrtumswahrscheinlichkeit von 10 % kann dann durch das Intervall

$$[m - \overline{m}'_{|5\%|} \cdot \sigma_G, m + \overline{m}'_{|5\%|} \cdot \sigma_G]$$

abgeschätzt werden. Für die **Treppenstufenversuche** sind die so berechneten verallgemeinerten mittleren Quantile $\overline{m}'_{|5\%|}$ im Anhang in der Abbildung A.28 auf S. 205 aufgeführt. Die für die **Horizontversuche** bestimmten verallgemeinerten mittleren Quantile $\overline{m}'_{|5\%|}$ sind im Anhang in den Abbildungen A.29 und A.30 ab S. 206 zu finden.

3.4.3 Schätzung der Standardabweichung

Die in der Simulation ermittelten Standardabweichungen s werden auf die tatsächliche Standardabweichung σ_G der Grundgesamtheit normiert. Die so ermittelten Schätzwerte s' sind die Stichprobe einer unbekanntem Verteilung und besitzen in jedem Simulationspunkt eine schiefe Häufigkeitsverteilung. Diese unbekanntem Verteilung kann nach LIU in [26] S. 46 mit einer logarithmischen Normalverteilung approximiert werden. Daher wird der nach Gleichung 3.10 gebildete Schätzer s'_{\lg} als eine normalverteilte Zufallsgröße betrachtet und ausgewertet.

$$s'_{\lg} = \lg \frac{s}{\sigma_G} \quad (3.10)$$

3.4.3.1 Anteil auswertbarer Versuchsreihen

Bei der Auswertung der Treppenstufenversuche liefern nur die Auswertungen nach DEUBELBEISS und mittels der MAX.-LIKELIHOOD-METHODE stets Schätzwerte für die Standardabweichung. Die Auswertung nach DIXON/MOOD liefert nur dann einen Schätzwert für die Standardabweichung, wenn die Varianzkennzahl nach Gleichung 2.41 einen Wert $k_{DM} \geq 0,3$ annimmt. Andernfalls muss die entsprechende Versuchsreihe nach der MAX.-LIKELIHOOD-METHODE ausgewertet werden. Damit sind auch bei der Auswertung nach DIXON/MOOD für alle Versuchsreihen Schätzer der Standardabweichung vorhanden. Bei der Auswertung nach HÜCK kann eine Schätzung der Standardabweichung nicht vorgenommen werden, wenn die Varianzkennzahl Werte annimmt, mit denen sich aus dem Korrekturdiagramm 2.9 a) auf S. 50 keine Werte mehr ablesen lassen. Für die Simulationsrechnungen wird dieses Korrekturdiagramm mit den Kurvenzügen nach HANTSCHKE [15], siehe Abschnitt B.1.1.1, regressiert. Dabei werden Versuchsreihen als ungültig erklärt, deren Varianzkennzahl außerhalb des Bereiches $0,3 \leq k_H \leq 6$ liegen oder deren berechnetes Verhältnis aus Standardabweichung der Grundgesamtheit und Stufenabstand den Wert 16 übersteigt. Bei der Auswertung nach der PROBIT-METHODE werden nur Versuchsreihen als gültig betrachtet, deren Schätzwert der Standardabweichung einen positiven Wert annimmt. Für alle Auswertemethoden werden in jedem Simulationspunkt die Anzahl der auswertbaren Versuchsreihen auf die Anzahl der in diesem Punkt simulierten Versuchsreihen bezogen. Es ergeben sich je nach Lage des Versuchsreihenstartwertes unterschiedliche Anteile auswertbarer Versuchsreihen. Alle ermittelten Anteile werden für einen Stufenabstand d' und eine Versuchsanzahl n , aber verschiedene

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

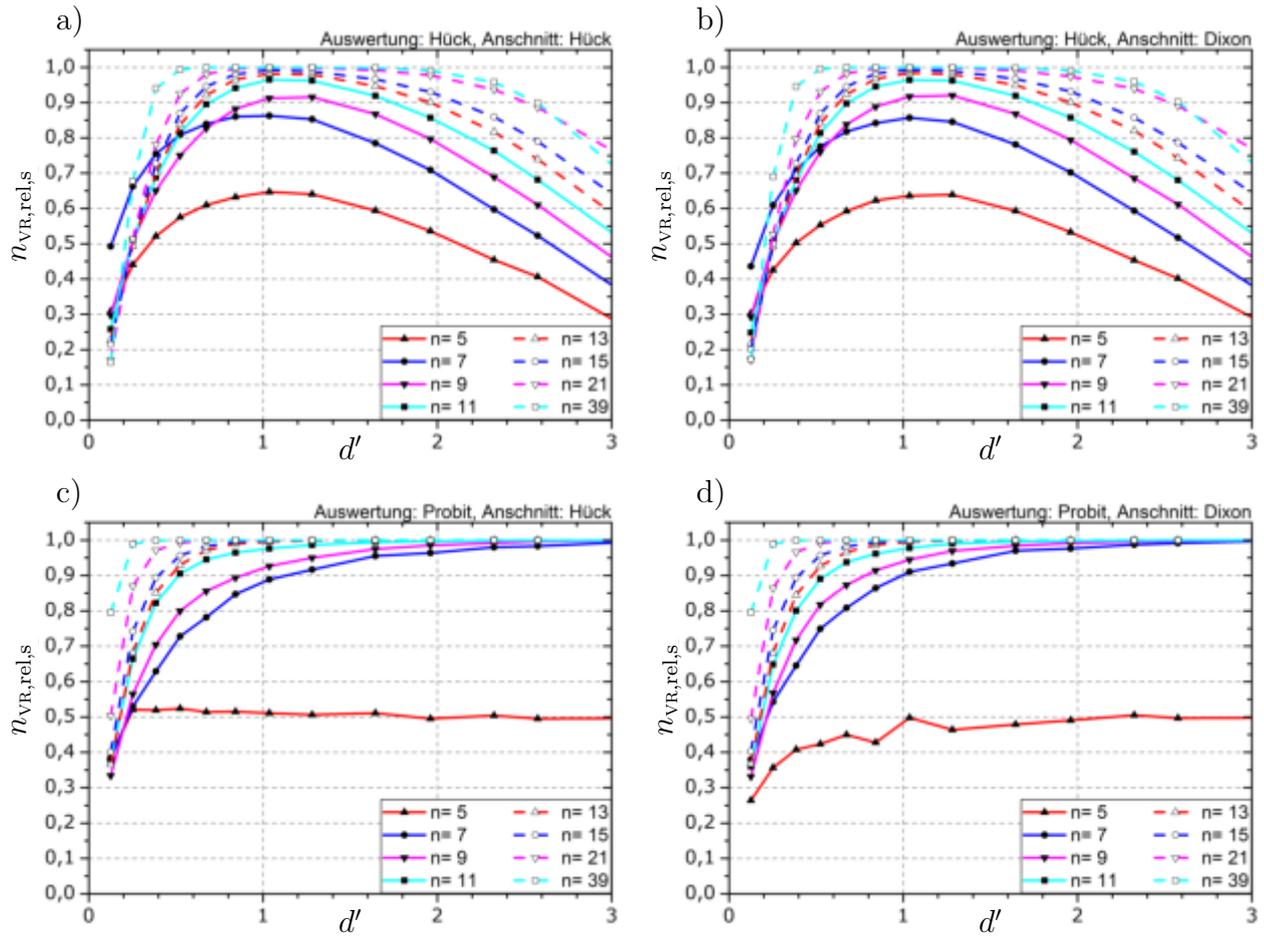


Abb. 3.8: Anteil der auswertbaren Treppenstufenversuchsreihen mit verschiedenen Anschnittsdefinitionen für die Auswertungen nach HÜCK und nach der PROBIT-METHODE.

Startwerte x'_0 , zusammengefasst. Der ungünstigste in den Simulationen auftretende Fall, also jeweils der minimale Anteil, wird dann als die relative Anzahl auswertbarer Versuchsreihen $n_{VR,rel,s}$ für jede Auswertemethode bestimmt. In Abbildung 3.8 sind die so ermittelten Anteile der auswertbaren Versuchsreihen für die Auswertungen von Treppenstufenversuchen nach HÜCK und der PROBIT-METHODE grafisch dargestellt. Die Abbildungen 3.9 und 3.10 zeigen die Anteile auswertbarer Versuchsreihen bei der Auswertung von Horizontversuchen mit der PROBIT-METHODE und der MAX.-LIKELIHOOD-METHODE.

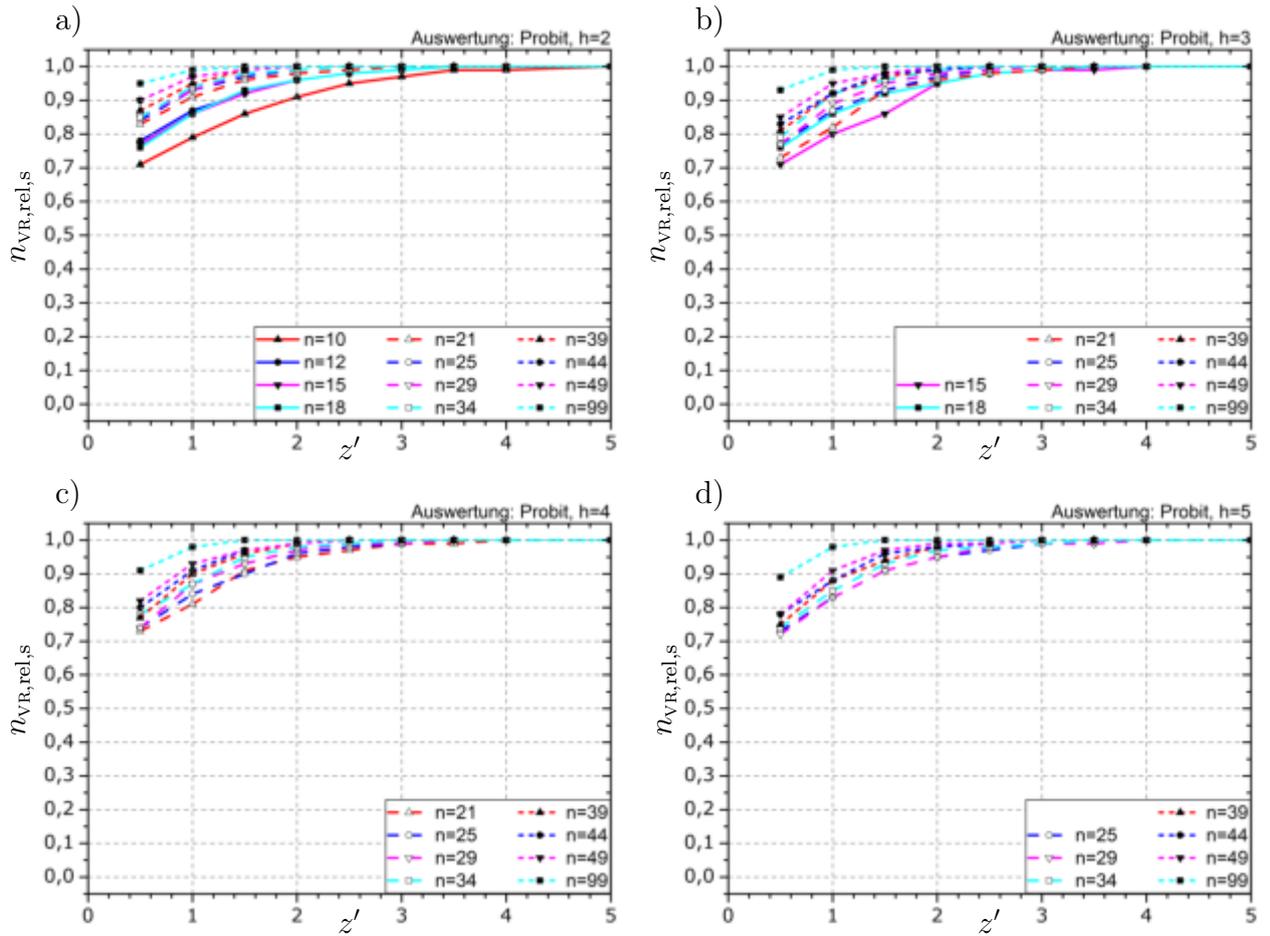


Abb. 3.9: Anteil auswertbarer Versuchsreihen für Horizontversuche bei Auswertung mittels PROBIT-METHODE. Für Versuche auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

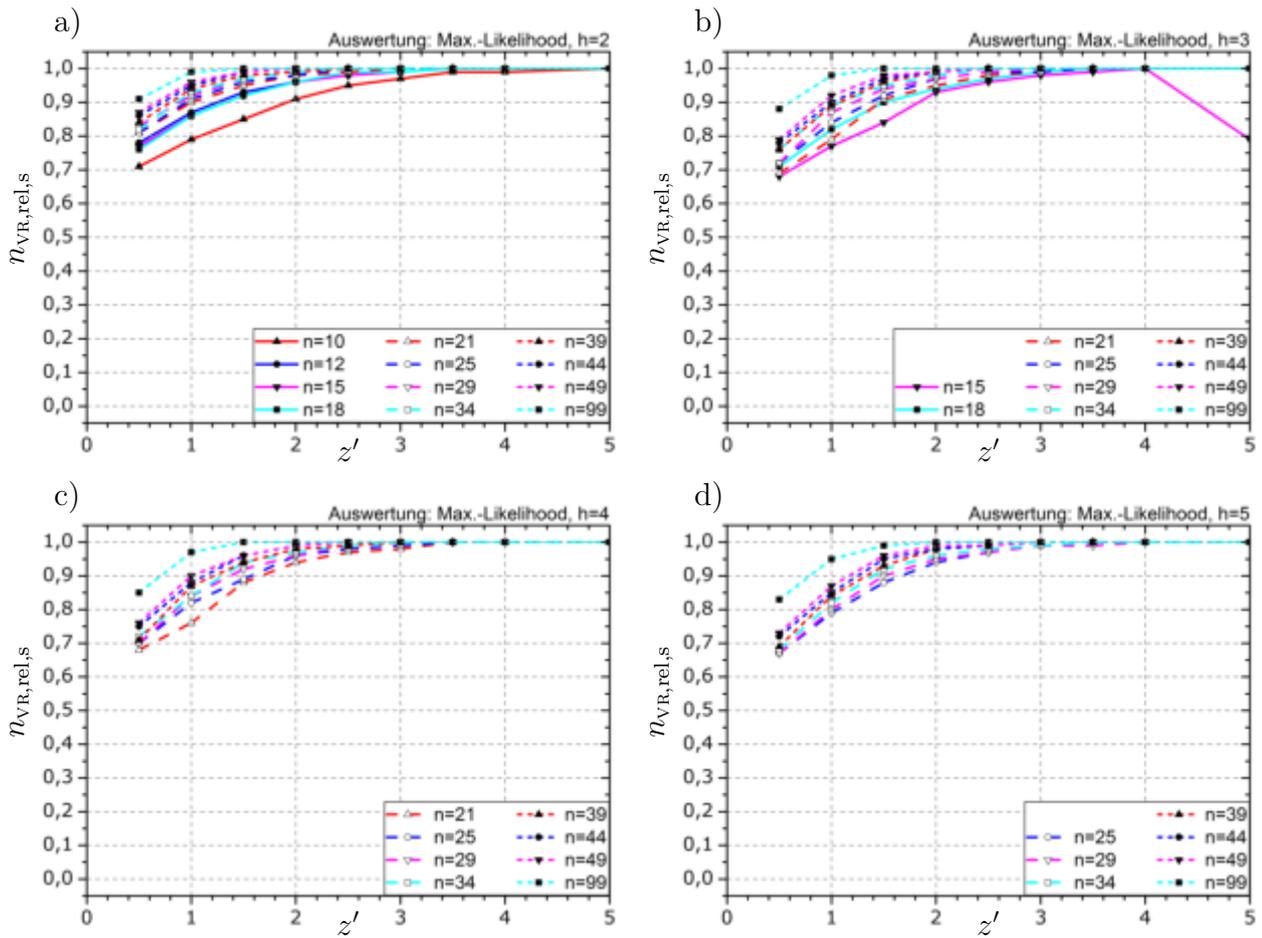


Abb. 3.10: Anteil auswertbarer Versuchsreihen für Horizontversuche bei Auswertung mittels MAX.-LIKELIHOOD-METHODE. Für Versuche auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

3.4.3.2 Systematischer und zufälliger Schätzfehler

Der Erwartungswert $\mathbf{E}[s'_{\text{lg}}]$ gibt die Verzerrung des Schätzers s'_{lg} an und ist für eine erwartungstreue Schätzung gerade Null. Die Standardabweichung $\mathbf{S}[s'_{\text{lg}}]$ des Schätzers beschreibt die zufällige Abweichung in jedem Simulationspunkt. Wird die Bestimmung der Standardabweichung der Schwingfestigkeit im Treppenstufenversuch als eine Messung verstanden, dann kann das vollständige Messergebnis nach Gleichung 3.11 angegeben werden.

$$s_k = s \cdot 10^{-\mathbf{E}[s'_{\text{lg}}] \pm \mathbf{S}[s'_{\text{lg}}]} \quad (3.11)$$

Die ermittelten Erwartungswerte und Standardabweichungen des Schätzers s'_{lg} variieren je nach Wahl des Versuchsreihenstartwertes, vgl. Abbildung 3.11 für Treppenstufenversuche und Abbildung 3.12 für Horizontversuche. Auch hier zeigt sich, dass die ermittelten systematischen und zufälligen Anteile der Schätzfehler gegen untere und obere Grenzen $G_{s',\text{lg},\text{max}}$ und $G_{s',\text{lg},\text{min}}$ bzw. $C_{s',\text{lg},\text{max}}$ und $C_{s',\text{lg},\text{min}}$ laufen. Daher werden aus den Ergebnissen der Treppenstufenversuche analog zum obigen Vorgehen, minimale und maximale Schätzfehler nach den folgenden Gleichung berechnet:

$$\begin{aligned} G_{s',\text{lg},\text{min}}(d', n) &= \min_{\forall x'_0} \left(\mathbf{E}[s'_{\text{lg}}(d', n, x'_0)] \right), \\ G_{s',\text{lg},\text{max}}(d', n) &= \max_{\forall x'_0} \left(\mathbf{E}[s'_{\text{lg}}(d', n, x'_0)] \right), \\ C_{s',\text{lg},\text{min}}(d', n) &= \min_{\forall x'_0} \left(\mathbf{S}[s'_{\text{lg}}(d', n, x'_0)] \right), \\ C_{s',\text{lg},\text{max}}(d', n) &= \max_{\forall x'_0} \left(\mathbf{S}[s'_{\text{lg}}(d', n, x'_0)] \right). \end{aligned}$$

Ebenso berechnen sich die minimalen und maximalen Schätzfehler für die Horizontversuche zu:

$$\begin{aligned} G_{s',\text{lg},\text{min}}(z', h, n) &= \min_{\forall x'_0, d' \leq 1} \left(\mathbf{E}[s'_{\text{lg}}(z', x'_0, d', h, n)] \right), \\ G_{s',\text{lg},\text{max}}(z', h, n) &= \max_{\forall x'_0, d' \leq 1} \left(\mathbf{E}[s'_{\text{lg}}(z', x'_0, d', h, n)] \right), \\ C_{s',\text{lg},\text{min}}(z', h, n) &= \min_{\forall x'_0, d' \leq 1} \left(\mathbf{S}[s'_{\text{lg}}(z', x'_0, d', h, n)] \right), \\ C_{s',\text{lg},\text{max}}(z', h, n) &= \max_{\forall x'_0, d' \leq 1} \left(\mathbf{S}[s'_{\text{lg}}(z', x'_0, d', h, n)] \right). \end{aligned}$$

Aus den so ermittelten Größen können mittlere Verzerrungen $\overline{G}_{s',\text{lg}}$ und Standardfehler $\overline{C}_{s',\text{lg}}$ des Schätzers s'_{lg} nach den Gleichungen 3.12 und 3.13 bestimmt werden.

$$\overline{G}_{s',\text{lg}} = \frac{G_{s',\text{lg},\text{max}} + G_{s',\text{lg},\text{min}}}{2} \quad (3.12)$$

$$\overline{C}_{s',\text{lg}} = \sqrt{\frac{(G_{s',\text{lg},\text{max}} - G_{s',\text{lg},\text{min}})^2}{12} + \frac{(C_{s',\text{lg},\text{max}} + C_{s',\text{lg},\text{min}})^2}{4}} \quad (3.13)$$

Die mittleren Verzerrungen und Standardfehler für die verschiedenen Auswertemethoden sind im Anhang für die **Treppenstufenversuche** in den Abbildungen A.31 bis A.32 ab S. 208 und für die **Horizontversuche** in den Abbildungen A.33 bis A.36 ab S. 210 aufgeführt.

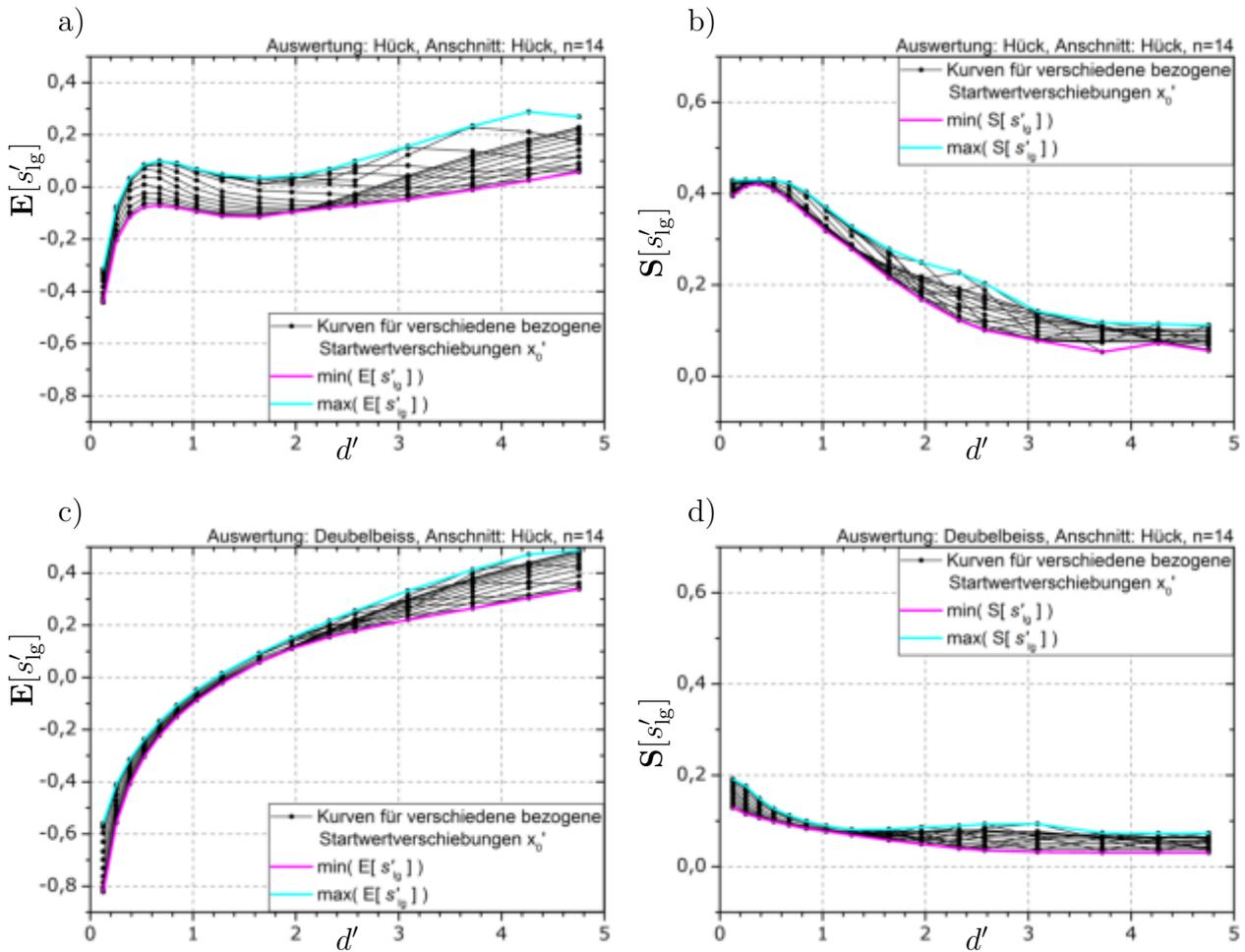


Abb. 3.11: Erwartungswert und Standardfehler des Schätzers s'_{lg} für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und unterschiedlich gewählten Startwerten x_0 . Auswertung nach HÜCK in a) und b) sowie nach DEUBELBEISS in c) und d).

3.4.3.3 RMS-Fehler

Nach Gleichung 2.5 werden in jedem Simulationspunkt für den Schätzer s'_{lg} die RMS-Fehler bestimmt. Analog zu Abschnitt 3.4.2.4 können aus den maximalen und minimalen RMS-Fehlern, vgl. Abbildungen 3.13 und 3.14, mittlere RMS-Fehler nach Gleichung 3.14 gebildet werden.

$$\overline{RMS}_{s',lg} = \frac{RMS_{s',lg,max} + RMS_{s',lg,min}}{2} \quad (3.14)$$

Die für die Auswertungen der **Treppenstufenversuche** berechneten mittleren RMS-Fehler sind in den Abbildungen A.37 und A.38 ab S. 213 dargestellt. Die mittleren RMS-Fehler der **Horizontversuche** sind in den Abbildungen A.39 und A.42 im Anhang auf S. 215 dargestellt. Die Mittelwertbildung ist nur dann angebracht, wenn der Intervallmittelwert größer als die Breite des Intervalls ausfällt. Daher sind den mittleren RMS-Fehlern die halben Intervallbreiten bezogen auf die Intervallmittelwerte gegenübergestellt.

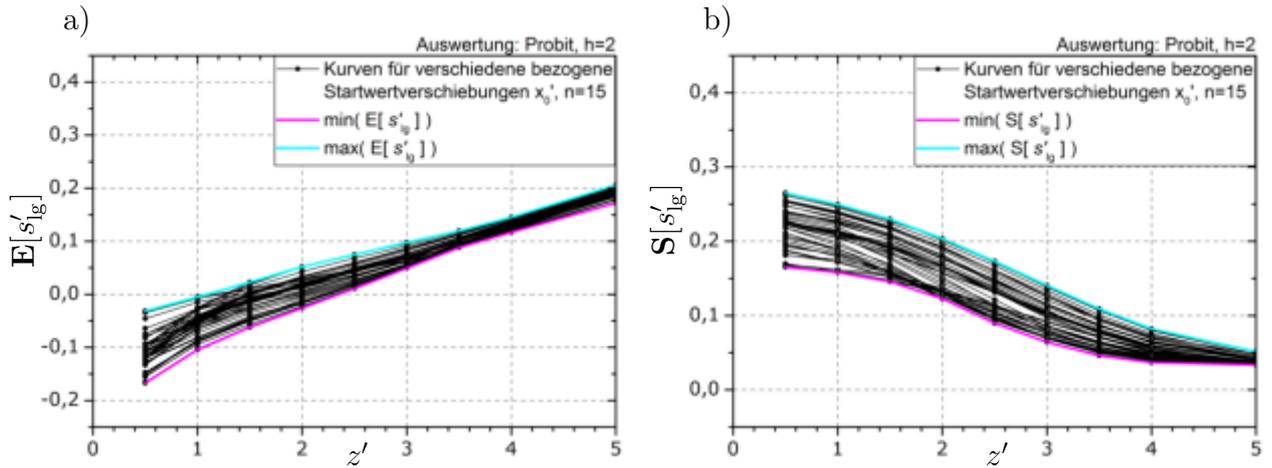


Abb. 3.12: Beispiel für Erwartungswert a) und Standardfehler b) des Schätzers s'_{lg} für Horizontversuche mit $n = 15$ gültigen Versuchen, ausgewertet mittels PROBIT-METHODE, für unterschiedlich gewählte Startwerte x_0 .

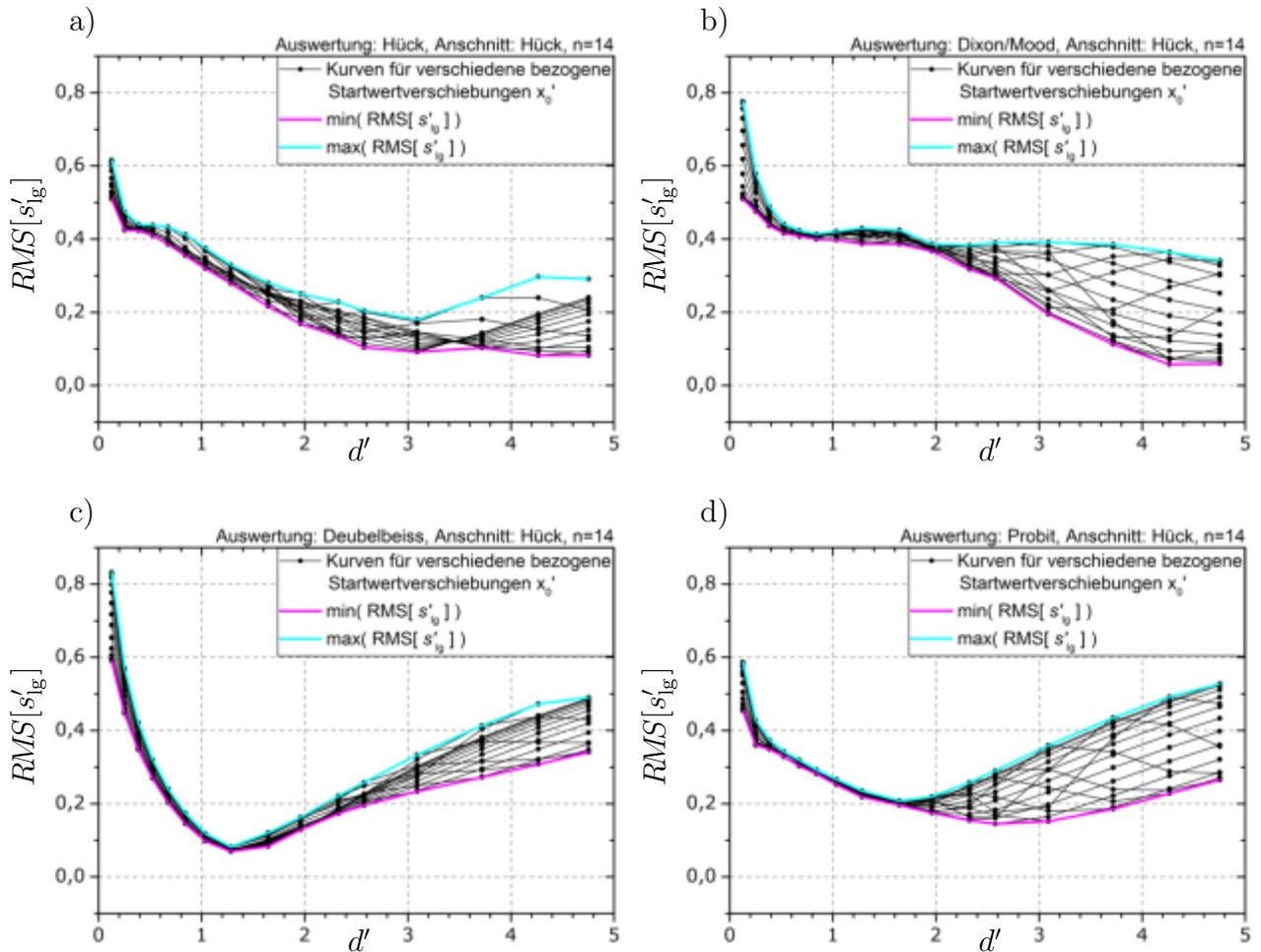


Abb. 3.13: Beispiel für RMS-Fehler des Schätzers s'_{lg} für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und unterschiedlich gewählten Startwerten x_0 . Auswertung nach HÜCK in a) , DIXON/MOOD in b), DEUBELBEISS in c) sowie mittels PROBIT-METHODE in d).

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

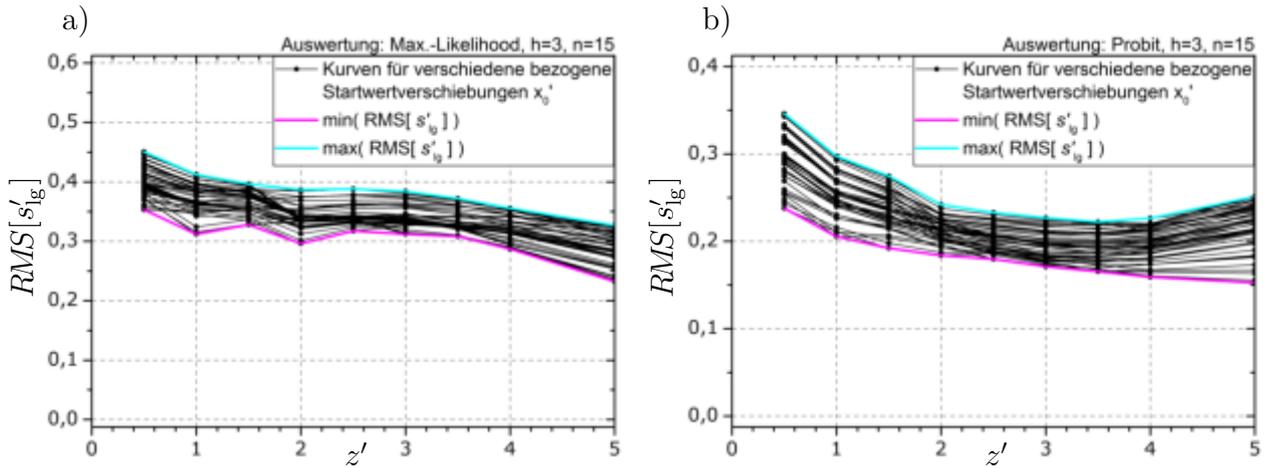


Abb. 3.14: Beispiel für RMS-Fehler des Schätzers s'_{lg} für Horizontversuche mit $n = 15$ gültigen Versuchen und unterschiedlich gewählten Startwerten x_0 . Auswertung mittels MAX.-LIKELIHOOD-METHODE in a) und PROBIT-METHODE in b).

3.4.3.4 Empirischer Vertrauensbereich des Schätzers s'_{lg}

In jedem Simulationspunkt werden aus der gezogenen Stichprobe die empirischen linksseitigen Quantile $s'_{lg,5\%}$ und die rechtsseitigen Quantile $s'_{lg,95\%}$ des Schätzers s'_{lg} bestimmt. Analog zum bereits oben beschriebenen Vorgehen werden aus diesen die minimal auftretenden Quantile $s'_{lg,5\%,min}$ und $s'_{lg,95\%,min}$ sowie die maximalen auftretenden Quantile $s'_{lg,5\%,max}$ und $s'_{lg,95\%,max}$ bestimmt. Diese begrenzen die in den Simulationen für verschiedene Startwertverschiebungen ermittelten empirischen Quantile, siehe exemplarisch in Abbildung 3.15. Mithilfe der Gleichungen 3.15 und 3.16 können dann die mittleren Quantile $\bar{s}'_{lg,5\%}$ und $\bar{s}'_{lg,95\%}$ für die

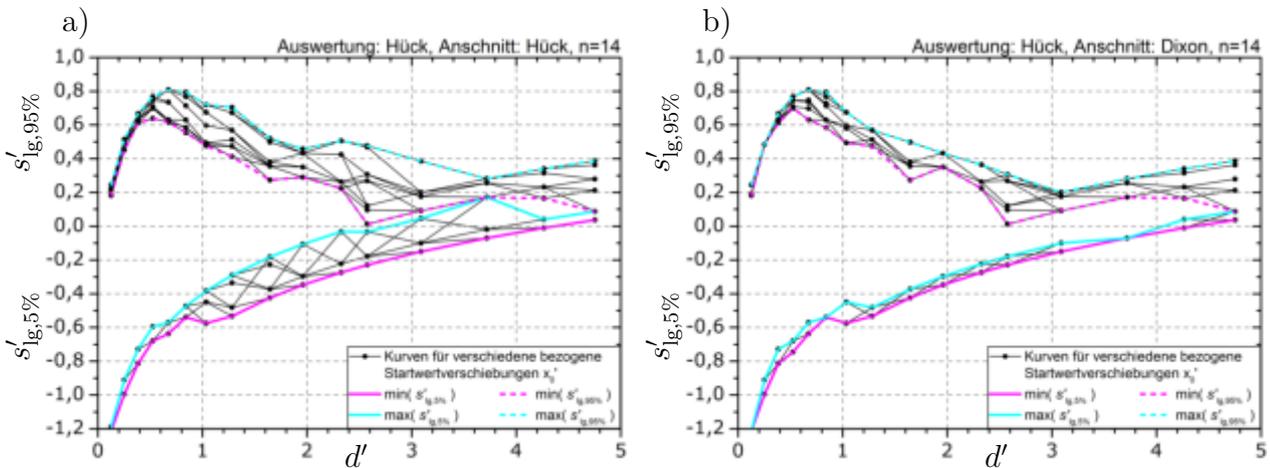


Abb. 3.15: Empirische 5%- und 95%- Quantile des Schätzers s'_{lg} für Treppenstufenversuche mit $n = 14$ gültigen Einzelversuchen, ausgewertet nach HÜCK für unterschiedlich gewählte Startwerte x_0 und eine Anschnittsdefinition nach HÜCK in a) oder nach DIXON in b).

verschiedenen Auswerteverfahren ermittelt werden.

$$\bar{s}'_{\text{lg},5\%} = \frac{s'_{\text{lg},5\%,\text{min}} + s'_{\text{lg},5\%,\text{max}}}{2} \quad (3.15)$$

$$\bar{s}'_{\text{lg},95\%} = \frac{s'_{\text{lg},95\%,\text{min}} + s'_{\text{lg},95\%,\text{max}}}{2} \quad (3.16)$$

Entsprechend wird der mittlere Vertrauensbereich für eine aus Treppenstufen- oder Horizontversuchen geschätzte Standardabweichung s zu einer Irrtumswahrscheinlichkeit von $\alpha = 10\%$ durch das Intervall

$$\left[s \cdot 10^{-\bar{s}'_{\text{lg},95\%}}, s \cdot 10^{-\bar{s}'_{\text{lg},5\%}} \right] \quad (3.17)$$

angegeben. Für die Bewertung der Ergebnisse ist die obere Grenze des Vertrauensbereiches wichtig. Sie erlaubt es, die höchstens zu erwartende Streuung einzugrenzen. Für den tatsächlichen Parameter σ_G kann mit

$$\sigma_G < 10^{-\bar{s}'_{\text{lg},5\%}} \cdot s$$

eine obere Vertrauensgrenze zu einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$ angegeben werden. Die aus den Simulationsergebnissen berechneten mittleren linksseitigen 5%-Quantile $\bar{s}'_{\text{lg},5\%}$ des Schätzers s'_{lg} sind für **Treppenstufenversuche** im Anhang in der Abbildung A.43 auf S. 218 aufgeführt. Für **Horizontversuche** sind die Ergebnisse in den Abbildungen A.44 und A.45 ab S. 219 dargestellt.

3.4.4 Auswertung der Schätzung des Erwartungswertes bei unbekannter Standardabweichung

3.4.4.1 Systematischer und zufälliger Schätzfehler

Wird der in der jeweiligen Versuchsreihe ermittelte Schätzwert für den Erwartungswert auf den im selben Versuch ermittelten Schätzwert der Standardabweichung normiert, d. h. der bezogene Schätzer m^* gebildet, lassen sich ebenso, wie im vorhergehenden Abschnitt beschrieben, systematische und zufällige Schätzfehler ableiten. Der Erwartungswert $\mathbf{E}[m^*]$ und die Standardabweichung $\mathbf{S}[m^*]$ des Schätzers m^* geben den systematischen und den zufälligen Schätzfehler für den Erwartungswert der Grundgesamtheit in Vielfachen der im selben Versuch geschätzten Standardabweichung an. Das korrigierte und vollständige Messergebnis kann nach Gleichung 3.18 angegeben werden.

$$m_k = m - \mathbf{E}[m^*] \cdot s \pm \mathbf{S}[m^*] \cdot s \quad (3.18)$$

Sowohl der Erwartungswert als auch die Standardabweichung des Schätzers m^* laufen bei Variation der Startwertverschiebung gegen obere und untere Grenzen, vgl. Abbildung 3.16, sodass sich die jeweils minimal und maximal auftretenden Verzerrungen und Standardfehler des Schätzers m^* bestimmen lassen. Aus diesen Gleichungen werden, analog zum Vorgehen in Abschnitt 3.4.2.2, mittlere Verzerrungen \bar{G}_{m^*} und Standardfehler \bar{C}_{m^*} gebildet:

$$\bar{G}_{m^*} = \frac{G_{m^*,\text{max}} + G_{m^*,\text{min}}}{2},$$

$$\bar{C}_{m^*} = \sqrt{\frac{(G_{m^*,\text{max}} - G_{m^*,\text{min}})^2}{12} + \frac{(C_{m^*,\text{max}} + C_{m^*,\text{min}})^2}{4}}.$$

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

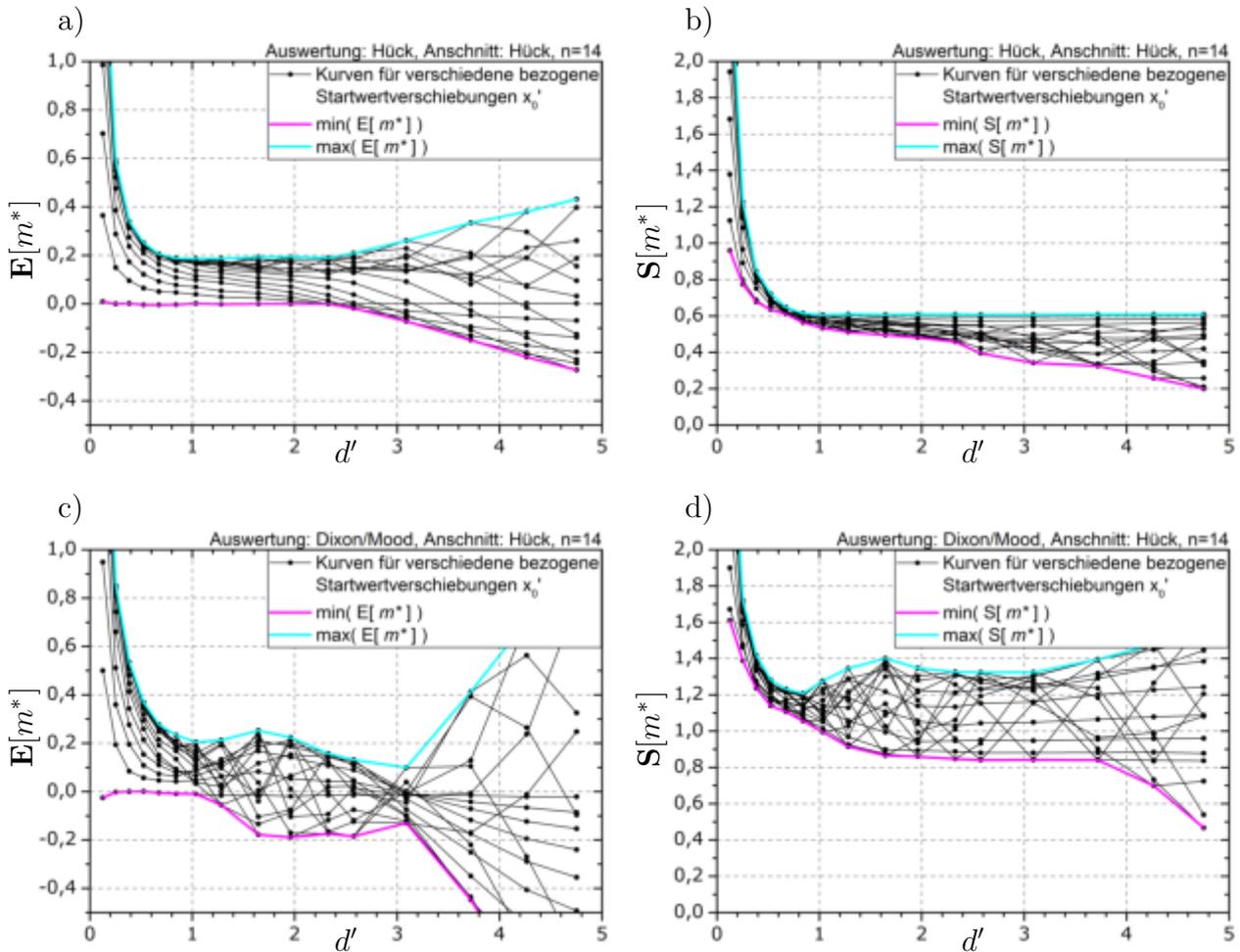


Abb. 3.16: Erwartungswert und Standardfehler des Schätzers m^* für Treppenstufenversuche mit $n = 14$ gültigen Versuchen, für unterschiedlich gewählte Startwerte x_0 und eine Anschnittsdefinition nach HÜCK. Auswertung nach HÜCK in a) und b) oder nach DIXON/MOOD in c) und d).

Die für die **Treppenstufenversuche** ermittelten mittleren Schätzfehler können den Abbildungen A.46 bis A.47 ab S. 221 im Anhang entnommen werden. Die Ergebnisse der **Horizontversuche** sind im Anhang in den Abbildungen A.48 bis A.51 ab S. 223 grafisch dargestellt. Durch die zusätzliche Schätzung der Standardabweichung streut der Schätzer m^* vor allem bei geringer Probenanzahl deutlich stärker als der Schätzer m' . Wie auch schon im vorherigen Abschnitt erscheint daher eine Betrachtung von Ergebnissen mit $n < 10$ gültigen Einzelversuchen nicht sinnvoll. Bei der Auswertung der Horizontversuche werden darüber hinaus nur Ergebnisse von Versuchsreihen mit $n/h \geq 5$ berücksichtigt. Die hier dargestellten mittleren Schätzfehler gelten für die Annahme eines unbekanntes, aber beliebig oberhalb des Erwartungswertes der Grundgesamtheit liegenden Versuchsreihenstartwertes.

3.4.4.2 Allgemeine Messunsicherheit

Ist über den Startwert der Versuchsreihe nichts bekannt, kann für die vorliegenden Simulationsergebnisse aus den minimalen und maximalen Verzerrungen und Standardfehlern des Schätzers x' ein verallgemeinerter Standardfehler nach Gleichung 3.19 bestimmt werden.

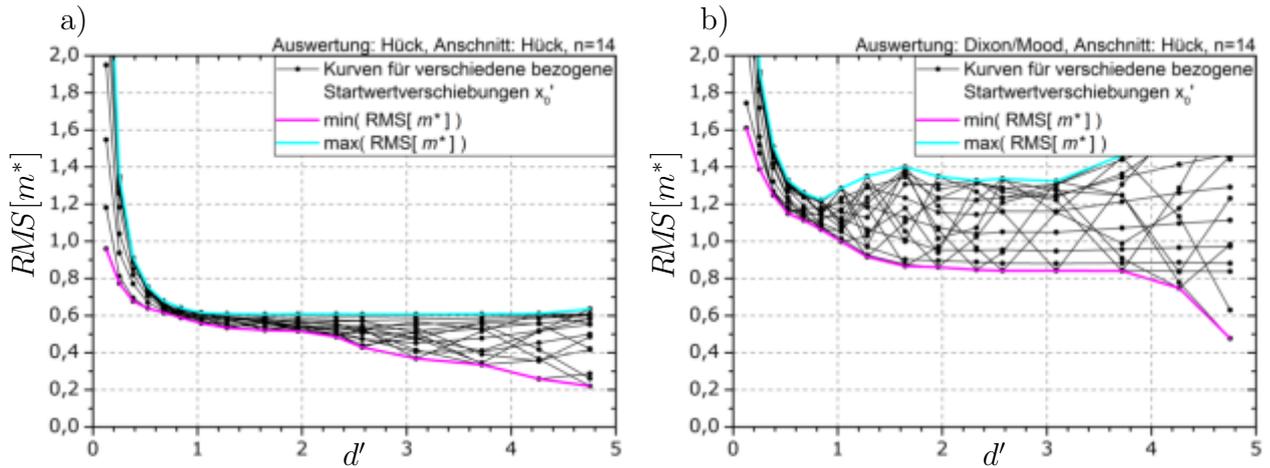


Abb. 3.17: RMS-Fehler des Schätzers m^* für Treppenstufenversuche mit $n = 14$ gültigen Versuchen, für unterschiedlich gewählte Startwerte x_0 und eine Anschnittsdefinition nach HÜCK. Auswertung nach HÜCK in a) und nach DIXON/MOOD in b).

Die grafische Darstellung des verallgemeinerten Standardfehlers bei unbekannter Standardabweichung ist für **Treppenstufenversuche** in der Abbildung A.52 auf S. 226 und für **Horizontversuche** in den Abbildungen A.53 bis A.54 ab S. 227 aufgeführt.

$$U_{m^*} = \sqrt{\frac{\max(|G_{m^*,\max}|, |G_{m^*,\min}|)^2}{3} + \frac{(C_{m^*,\max} + C_{m^*,\min})^2}{4}} \quad (3.19)$$

3.4.4.3 RMS-Fehler

Werden die ermittelten Schätzfehler nicht nach systematischen und zufällig auftretenden Abweichungen unterschieden, sondern nach Gleichung 2.5 die RMS-Fehler für den Schätzer m^* bestimmt, können ebenso maximale und minimale RMS-Fehler ermittelt werden, vgl. Abbildung 3.17. Mit den daraus berechneten Mittelwerten bzw. mittleren RMS-Fehlern ist es möglich, die Genauigkeit der verschiedenen Auswertemethoden zu vergleichen. Für die Ergebnisse der **Treppenstufenversuche** sind die ermittelten mittleren RMS-Fehler in den Abbildungen A.55 bis A.56 ab S. 229 dargestellt. Für **Horizontversuche** sind ebenjene Ergebnisse den Abbildungen A.57 bis A.60 ab S. 231 zu entnehmen.

3.4.4.4 Empirischer Vertrauensbereich

Aus den empirischen links- und rechtsseitigen 5%-Quantilen $m_{5\%}^*$ und $m_{95\%}^*$ des Schätzers m^* werden mittlere empirische Quantile $\overline{m}_{5\%}^*$ und $\overline{m}_{95\%}^*$ bestimmt, vgl. exemplarisch in Abbildung 3.18 und 3.19. Für jede Auswertemethode kann dann die Lage des so ermittelten, zu einer Irrtumswahrscheinlichkeit von 10% gehörenden, Vertrauensbereichs

$$[m - \overline{m}_{95\%}^* \cdot s, m - \overline{m}_{5\%}^* \cdot s]$$

abgeschätzt werden. Bei unbekannter Lage des Versuchsreihenstartwertes ist das betragsmäßig größere Quantil für die Bestimmung des Vertrauensbereiches zu wählen. Es wird daher mit

$$\overline{m}_{|5\%|}^* = \max(|\overline{m}_{5\%}^*|, |\overline{m}_{95\%}^*|)$$

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

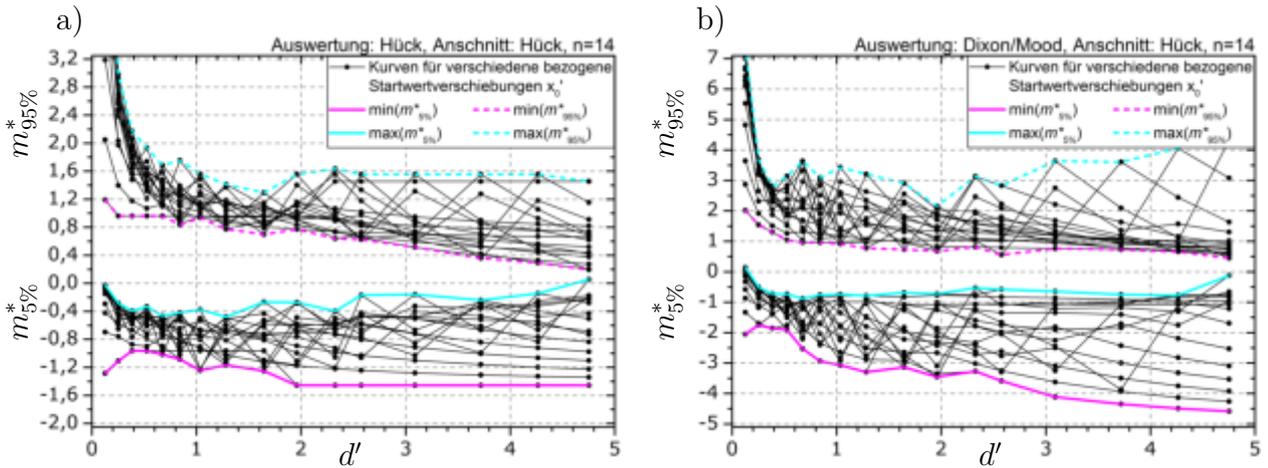


Abb. 3.18: 5 %- bzw. 95 %-Quantile des Schätzers m^* für Treppenstufenversuche mit $n = 14$ gültigen Versuchen, für unterschiedlich gewählte Startwerte x_0 und eine Anschnittsdefinition nach HÜCK. Auswertung nach HÜCK in a) und nach DIXON/MOOD in b).

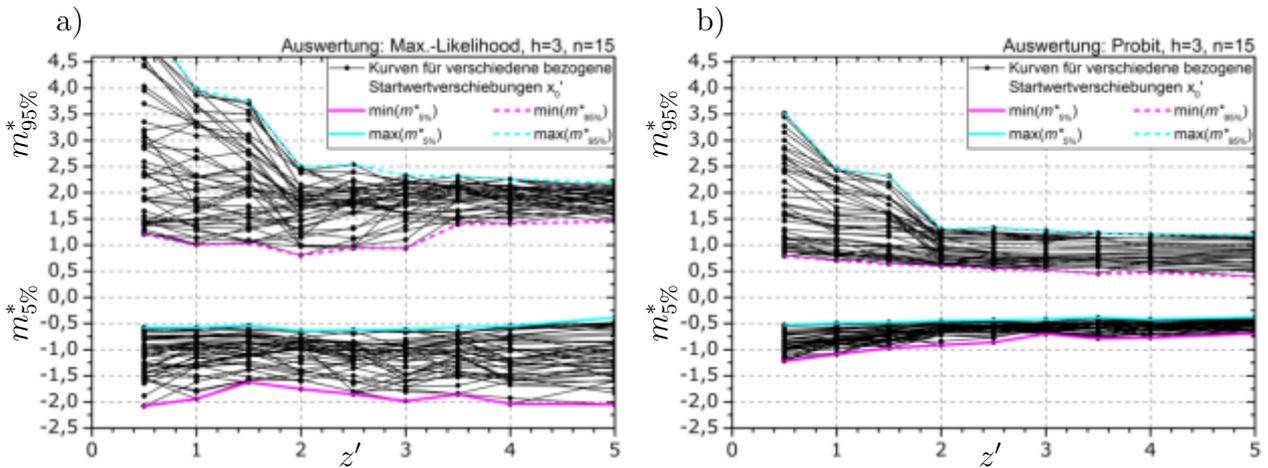


Abb. 3.19: 5 %- bzw. 95 %-Quantile des Schätzers m^* für Horizontversuche mit $n = 14$ gültigen Versuchen, für unterschiedlich gewählte Startwerte x_0 . Auswertung mittels MAX.-LIKELIHOOD-METHODE in a) und PROBIT-METHODE in b).

ein verallgemeinertes mittleres 5 %-Quantil $\overline{m^*}_{|5\%|}$ berechnet und der Vertrauensbereich zu $\alpha = 10 \%$ durch das folgende Intervall abgeschätzt

$$[m - \overline{m^*}_{|5\%|} \cdot s, m + \overline{m^*}_{|5\%|} \cdot s].$$

Für **Treppenstufenversuche** sind die verallgemeinerten mittleren 5 %-Quantile im Anhang in der Abbildung A.61 auf Seite 234 dargestellt. Ebenso können für **Horizontversuche** den Abbildungen A.62 und A.63 ab Seite 235 die verallgemeinerten mittleren 5 %-Quantile entnommen werden.

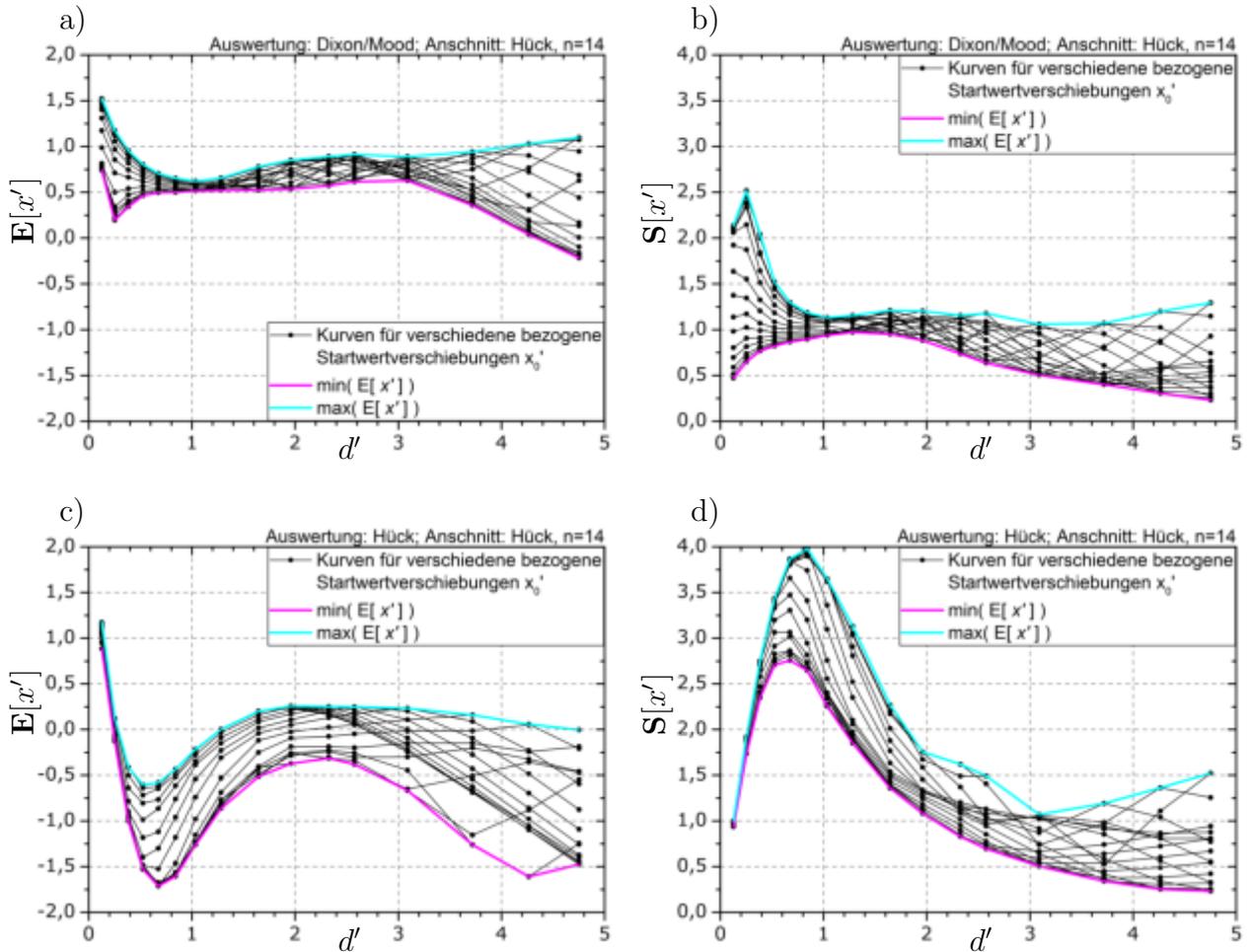


Abb. 3.20: Erwartungswert und Standardfehler des Schätzers x' für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und Anschnittsdefinition nach HÜCK für unterschiedlich gewählte Startwerte x_0 . Auswertung nach DIXON/MOOD in a) und b) sowie nach HÜCK in c) und d).

3.4.5 Auswertung der Schätzung des 2,5 %-Quantils bei bekannter Standardabweichung

3.4.5.1 Systematischer und zufälliger Schätzfehler

Für die Bauteildimensionierung bzw. den Festigkeitsnachweis wird allgemein eine Festigkeit geringer Ausfallwahrscheinlichkeit gewählt. In den Standardregelwerken wird meist das zu einer Ausfallwahrscheinlichkeit von $P_A = 2,5\%$ korrespondierende Quantil $X_{2,5\%}$ der Festigkeitsverteilung verwendet, vgl. [35] S. 10. Dessen direkte experimentelle Ermittlung ist für hohe Lastspielzahlen praktisch ausgeschlossen. Daher wird es unter Annahme einer Verteilung und den zugehörigen geschätzten Verteilungsparametern bestimmt. Unter der Annahme einer normalverteilten Festigkeit mit den geschätzten Parametern m und s erfolgt die Bestimmung durch

$$x_{2,5\%} = m - 1,96 \cdot s. \quad (3.20)$$

Sind die Messunsicherheiten bzw. Standardfehler s_m und s_s bekannt und die Schätzer m und s unkorreliert und normalverteilt, kann der Standardfehler des so geschätzten Quantils nach

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

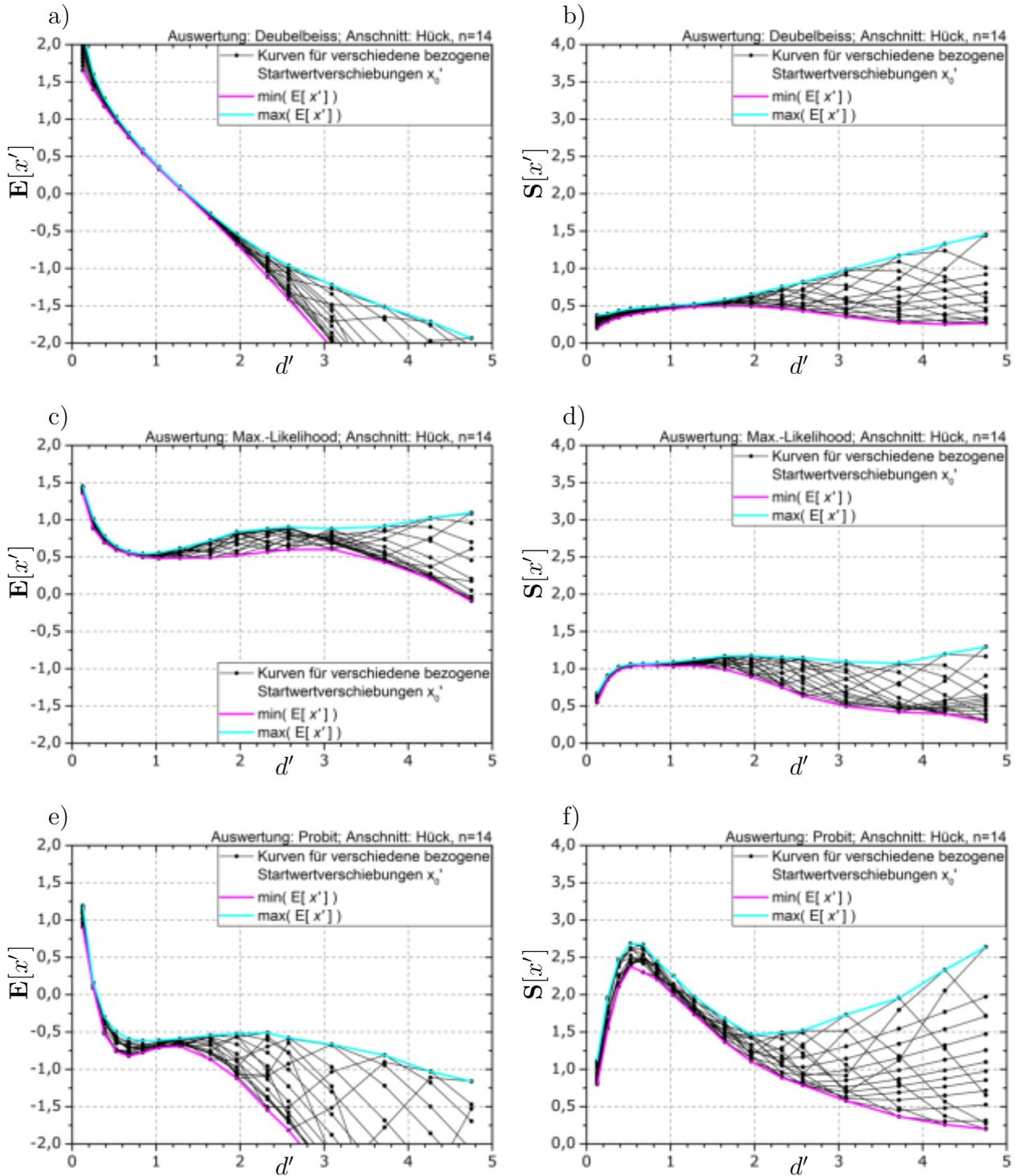


Abb. 3.21: Erwartungswert und Standardfehler des Schätzers x' für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und An schnittsdefinition nach HÜCK für unterschiedlich gewählte Startwerte x_0 . Auswertung nach DEUBELBEISS in a) und b), nach MAX.-LIKELIHOOD-METHODE in c) und d) sowie nach PROBIT-METHODE in e) und f).

Gleichung 3.21 ermittelt werden.

$$s_{x_{2,5\%}} = \sqrt{s_m^2 + 1,96^2 s_s^2} \quad (3.21)$$

Alternativ ist es möglich, die Schätzfehler auch aus den Simulationsergebnissen abzuschätzen. Dazu wird für jede Versuchsreihe aus den geschätzten Parametern das Quantil nach Gleichung 3.20 ermittelt und die Abweichung zum tatsächlichen Wert des 2,5 %-Quantils in Vielfachen der tatsächlichen Standardabweichung, also

$$x'_{2,5\%} = \frac{(m - 1,96 \cdot s) - (\mu_G - 1,96 \cdot \sigma_G)}{\sigma_G} \quad (3.22)$$

festgehalten. Für $x'_{2,5\%}$ wird der Einfachheit halber im Folgenden die Bezeichnung x' gewählt. Der Erwartungswert $\mathbf{E}[x']$ gibt die systematische Verschätzung, die Standardabweichung $\mathbf{S}[x']$ die zufällige Verschätzung des 2,5 %-Quantils an. Die Auswertung folgt demselben schon mehrfach oben beschriebenen Schema. Wie in den Abbildungen 3.20 bis 3.21 exemplarisch dargestellt, variieren beide Schätzfehler mit der Wahl des Versuchsreihenstartwertes in einem beschränkten Umfang. Aus den oberen und unteren Schranken $G_{x',\max}$ und $G_{x',\min}$ bzw. $C_{x',\max}$ und $C_{x',\min}$ lassen sich nach Gleichung 3.23 und 3.24 die mittleren systematischen Schätzfehler $\overline{G}_{x'}$ und mittleren zufällige Schätzfehler $\overline{C}_{x'}$ ableiten.

$$\overline{G}_{x'} = \frac{G_{x',\max} + G_{x',\min}}{2} \quad (3.23)$$

$$\overline{C}_{x'} = \sqrt{\frac{(G_{x',\max} - G_{x',\min})^2}{12} + \frac{(C_{x',\max} + C_{x',\min})^2}{4}} \quad (3.24)$$

Die berechneten mittleren Schätzfehler sind für **Treppenstufenversuche** in den Abbildungen A.64 bis A.65 gegenübergestellt. Für die untersuchten **Horizontversuche** werden die Ergebnisse in den Abbildungen A.66 und A.69 grafisch aufbereitet.

3.4.5.2 RMS-Fehler

In bekannter Art und Weise wird für jeden Simulationspunkt der RMS-Fehler des Schätzers x' nach Gleichung 2.5 ermittelt. Sowohl für die Ergebnisse der Treppenstufenversuche als auch der Horizontversuche werden aus den minimalen und den maximalen RMS-Fehlern, die bei der Variation der Startwertverschiebung auftreten, wieder mittlere RMS-Fehler des Schätzers x' gebildet, vgl. Abbildung 3.22 und 3.23.

Für die **Treppenstufenversuche** sind die abgeleiteten mittleren RMS-Fehler in den Abbildungen A.70 bis A.71 ab S. 242 dargestellt. Die entsprechenden Ergebnisse der **Horizontversuche** sind in den Abbildungen A.72 und A.75 ab S. 244 grafisch aufbereitet. Wie stark die mittleren RMS-Fehler durch Veränderung der Startwertverschiebung schwanken, wird durch die ebenfalls in den genannten Abbildungen dargestellten relativen RMS-Fehler angegeben.

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

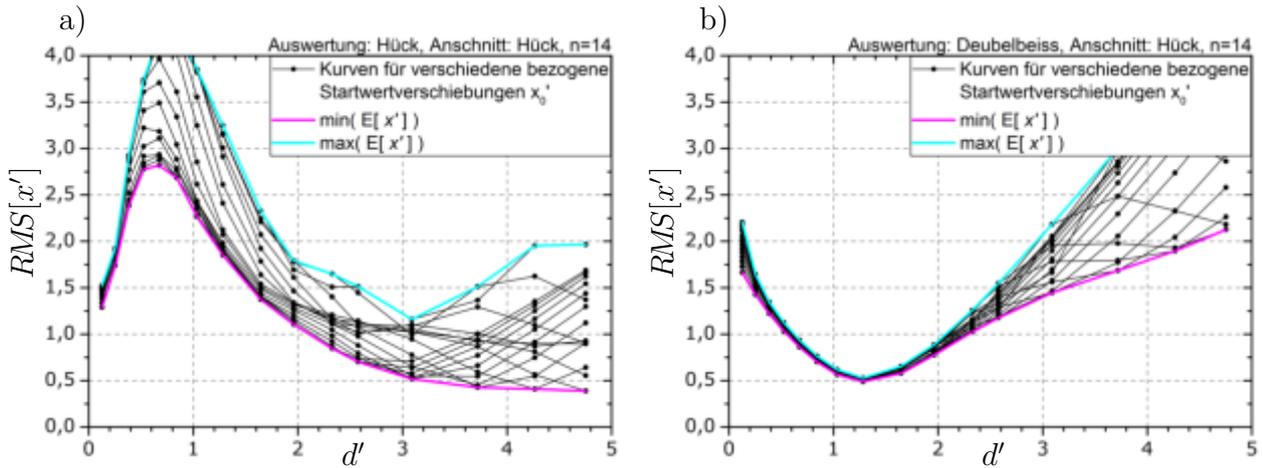


Abb. 3.22: RMS-Fehler des Schätzers x' für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und Anschnittsdefinition nach HÜCK für unterschiedlich gewählte Startwerte x_0 . Auswertung nach HÜCK in a) sowie DEUBELBEISS in b) .

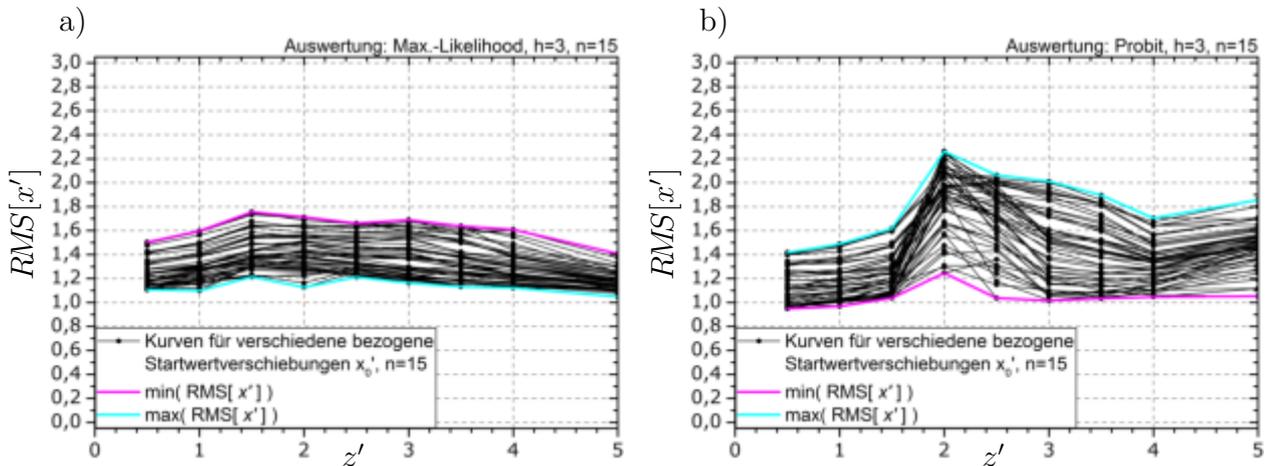


Abb. 3.23: RMS-Fehler des Schätzers x' für Horizontversuche mit $n = 15$ gültigen Versuchen bei unterschiedlich gewählten Startwerten x_0 . Auswertung nach MAX.-LIKELIHOOD-METHODE in a) sowie PROBIT-METHODE in b) .

3.4.5.3 Empirischer Vertrauensbereich des Schätzers x'

Entscheidend für die praktische Anwendung ist es, eine Überschätzung der Festigkeit auszuschließen. Anhand der ermittelten empirischen 95 %-Quantile des Schätzers x' kann eine untere Grenze für den Vertrauensbereich zu $\alpha = 0,05$ angegeben werden:

$$X_{2,5\%,G} \geq x_{2,5\%} - \bar{x}'_{95\%} \cdot \sigma_G .$$

Der wahre Wert $X_{2,5\%,G}$ des 2,5 %-Quantils wird dann im Mittel nur in einer von 20 Wiederholungen des Experiments unterhalb dieser Grenze liegen. Das mittlere 95 %-Quantil $\bar{x}'_{95\%}$ berechnet sich hierbei wieder aus den ermittelten minimalen und maximalen 95 %-Quantilen. Die so bestimmten mittleren 95 %-Quantile sind in den Abbildungen A.76 bis A.78 ab Seite 247 grafisch dargestellt.

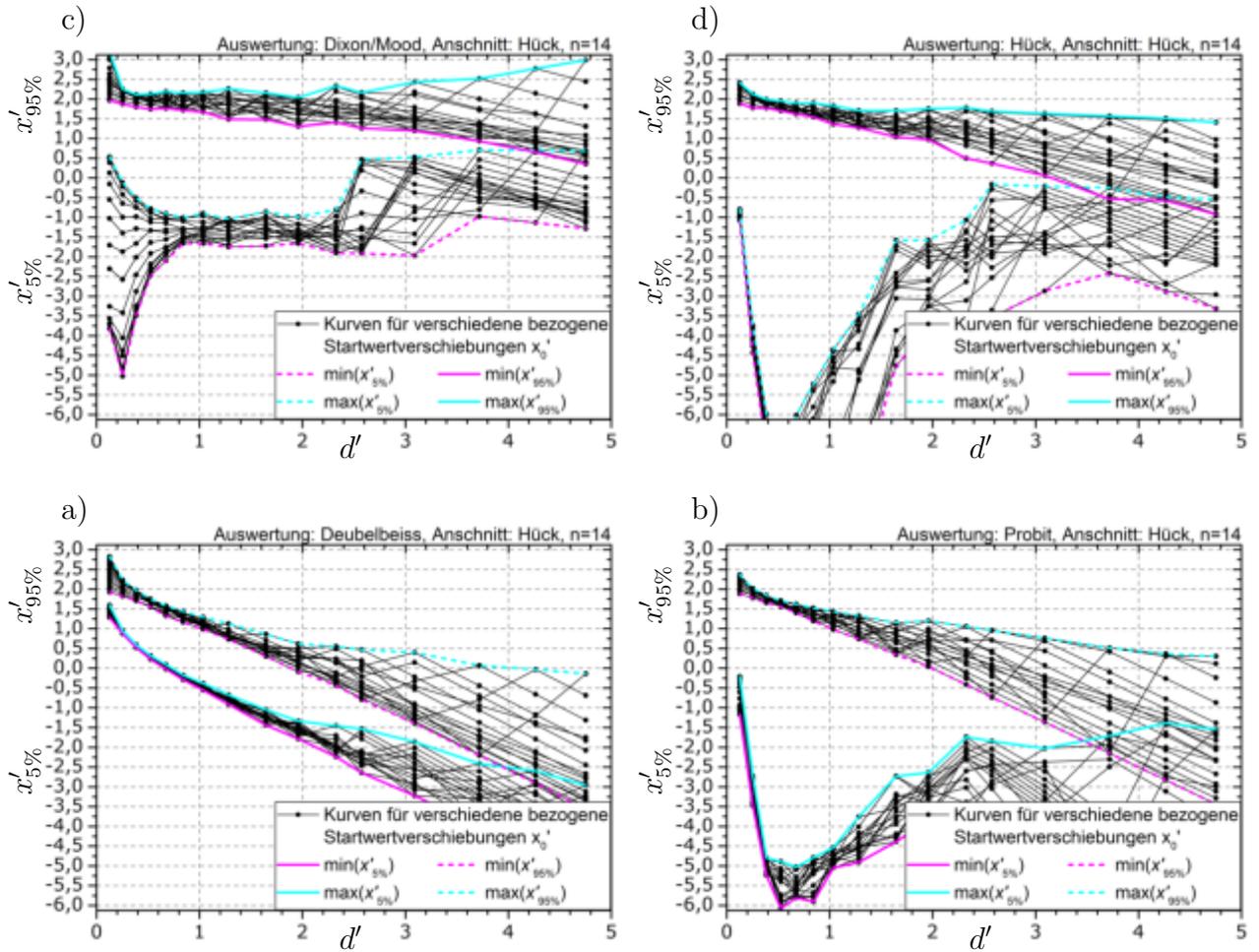


Abb. 3.24: Empirische 5%- und 95%-Quantile des Schätzers x' für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und Anschnittsdefinition nach HÜCK für unterschiedlich gewählte Startwerte x_0 . Auswertung nach DIXON/MOOD in a), nach HÜCK in b), nach DEUBELBEISS in c) sowie mittels PROBIT-METHODE in d).

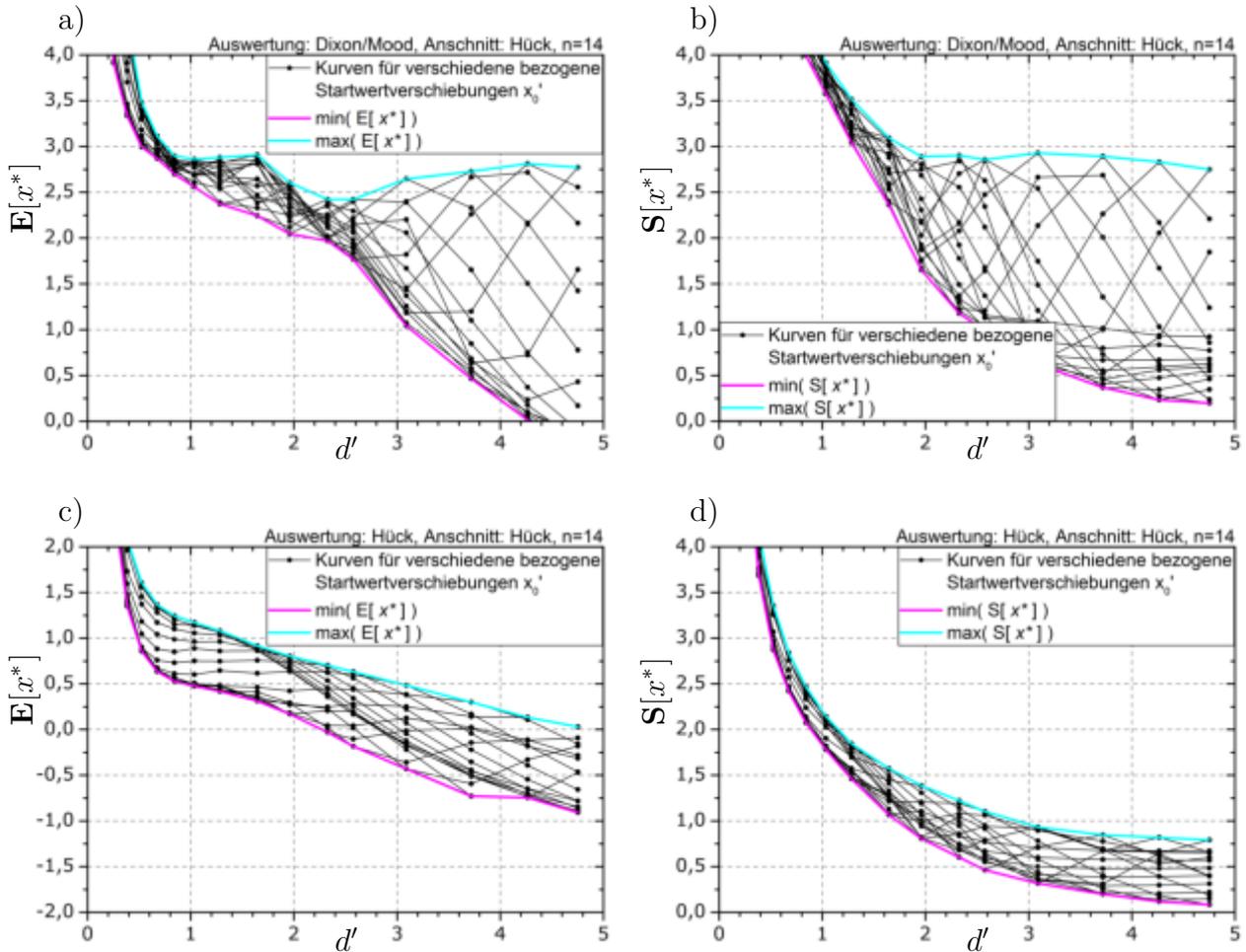


Abb. 3.25: Erwartungswert und Standardfehler des Schätzers x^* für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und Anschnittsdefinition nach HÜCK für unterschiedlich gewählte Startwerte x_0 . Auswertung nach DIXON/MOOD in a) und b) sowie nach HÜCK in c) und d).

3.4.6 Auswertung der Schätzung des 2,5 %-Quantils bei unbekannter Standardabweichung

3.4.6.1 Systematischer und zufälliger Schätzfehler

Für die Beurteilung des Schätzfehlers anhand der im selben Versuch ermittelten Standardabweichung wird der Abstand zwischen dem geschätzten und dem tatsächlichen 2,5 %-Quantil auf die im selben Versuch geschätzte Standardabweichung normiert, d. h. es wird der Schätzer

$$x^* = \frac{(m - 1,96 \cdot s) - (\mu_G - 1,96 \cdot \sigma_G)}{s} = \frac{x_{2,5\%} - X_{2,5\%,G}}{s}$$

gebildet. Dessen Erwartungswert $E[x^*]$ gibt die systematische Verschätzung bzw. die Verzerrung der Schätzung an. Die Standardabweichung $S[x^*]$ beschreibt den zufälligen Anteil der Verschätzung des 2,5 %-Quantils. Aus den Simulationsergebnissen werden die bei Variation der Startwertverschiebung auftretenden minimalen und maximalen Schätzfehler ermittelt, vgl. Abbildung 3.25 und 3.26. Aus diesen oberen und unteren Grenzen werden, wie schon

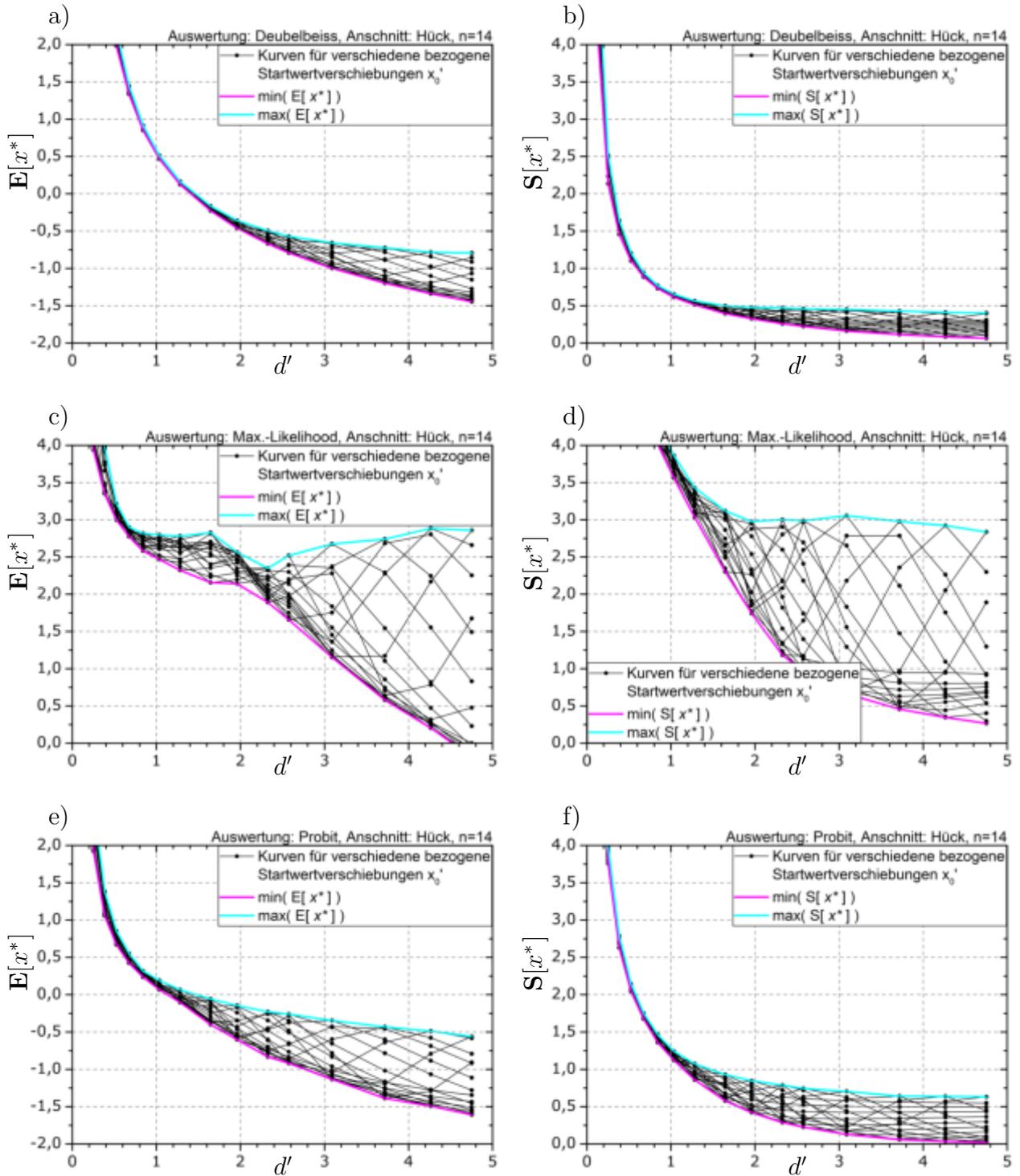


Abb. 3.26: Erwartungswert und Standardfehler des Schätzers x^* für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und Anschnittsdefinition nach HÜCK für unterschiedlich gewählte Startwerte x_0 . Auswertung nach DEUBELBEISS in a) und b), nach MAX.-LIKELIHOOD-METHODE in c) und d) sowie nach PROBIT-METHODE in e) und f).

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

mehrfach oben beschriebenen, die mittlere Schätzfehler

$$\bar{G}_{x^*} = \frac{G_{x^*,\max} + G_{x^*,\min}}{2}$$

$$\bar{C}_{x^*} = \sqrt{\frac{(G_{x^*,\max} - G_{x^*,\min})^2}{12} + \frac{(C_{x^*,\max} + C_{x^*,\min})^2}{4}}$$

bestimmt. Die daraus berechneten mittleren Schätzfehler werden für **Treppenstufenversuche** in den Abbildungen A.79 bis A.80 gegenübergestellt. Für die untersuchten **Horizontversuche** sind die Ergebnisse in den Abbildungen A.81 und A.84 grafisch aufbereitet.

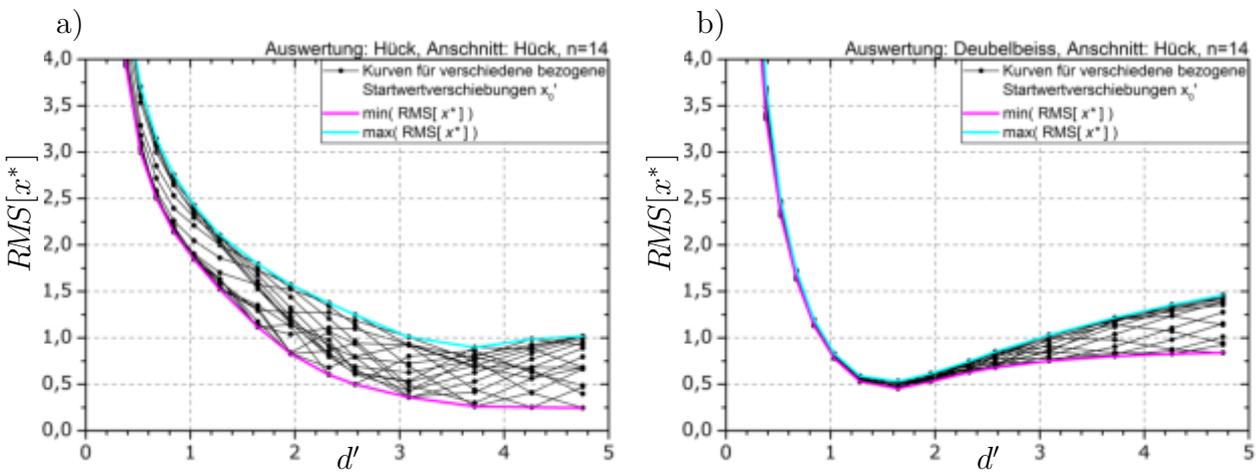


Abb. 3.27: RMS-Fehler des Schätzers x^* für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und Anschnittsdefinition nach HÜCK für unterschiedlich gewählte Startwerte x_0 . Auswertung nach HÜCK in a) sowie DEUBELBEISS in b).

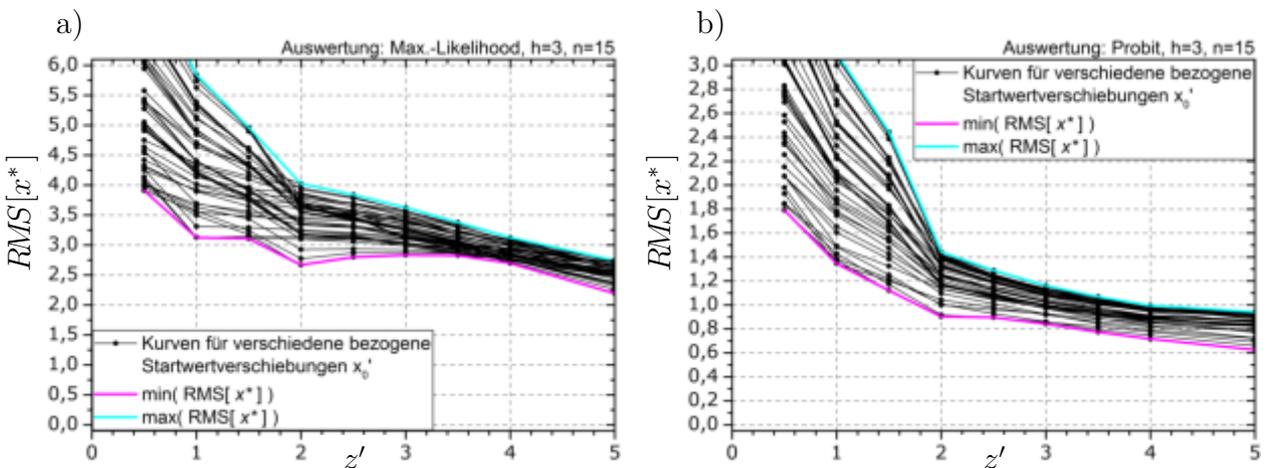


Abb. 3.28: RMS-Fehler des Schätzers x^* für Horizontversuche mit $n = 15$ gültigen Versuchen für unterschiedlich gewählte Startwerte x_0 . Auswertung nach MAX.-LIKELIHOOD-METHODE in a) sowie PROBIT-METHODE in b).

3.4.6.2 RMS-Fehler

Wird der Schätzfehler in jedem Simulationspunkt nicht in einen systematischen und zufälligen Anteil aufgeteilt, sondern zu einem mittleren Schätzfehler nach Gleichung 2.5 zusammengefasst, ist eine einfachere Bewertung der Ergebnisse möglich. Allerdings kann dann die Schätzgüte von gemittelten Ergebnissen nicht eingeschätzt werden. Aus den ermittelten minimalen und maximalen RMS-Fehlern, vgl. Abbildungen 3.27 und 3.28, die bei der Variation der Startwertverschiebung beobachtet werden, wird ein mittlerer RMS-Fehler berechnet:

$$\overline{RMS}_{x^*} = \frac{RMS_{x^*,\max} + RMS_{x^*,\min}}{2}. \quad (3.25)$$

Für die **Treppenstufenversuche** sind die mittleren RMS-Fehler in den Abbildungen A.70 bis A.71 ab S. 242 dargestellt. Die entsprechenden Ergebnisse der **Horizontversuche** sind in den Abbildungen A.87 und A.90 ab S. 257 zusammengefasst. Wie stark der mittlere RMS-Fehler durch Veränderung der Startwertverschiebung schwankt, wird durch den ebenfalls in den genannten Abbildungen dargestellten relativen RMS-Fehler angegeben.

3.4.6.3 Empirischer Vertrauensbereich des Schätzers x^*

Für eine Abschätzung der unteren Vertrauensgrenze des 2,5 %-Quantils $X_{2,5\%,G}$ der Grundgesamtheit werden aus den Simulationsversuchen mittlere 95 %-Quantile $\bar{x}_{95\%}^*$ bestimmt.

$$\bar{x}_{95\%}^* = \frac{x_{95\%,\max}^* + x_{95\%,\min}^*}{2} \quad (3.26)$$

Hierzu erfolgt wieder zunächst die Ermittlung minimaler und maximaler 95 %-Quantile $x_{95\%,\min}^*$ und $x_{95\%,\max}^*$, vgl. Abbildung 3.29 und 3.30. Die mittlere Vertrauensgrenze zu einer Irrtumswahrscheinlichkeit von $\alpha = 0,05$ kann dann durch

$$X_{2,5\%,G} \geq x_{2,5\%} - \bar{x}_{95\%}^* \cdot s \quad (3.27)$$

abgeschätzt werden. Die berechneten mittleren 95 %-Quantile $\bar{x}_{95\%}^*$ des Schätzers x^* sind in den Abbildungen A.91 bis A.93 im Anhang ab S. 260 grafisch dargestellt.

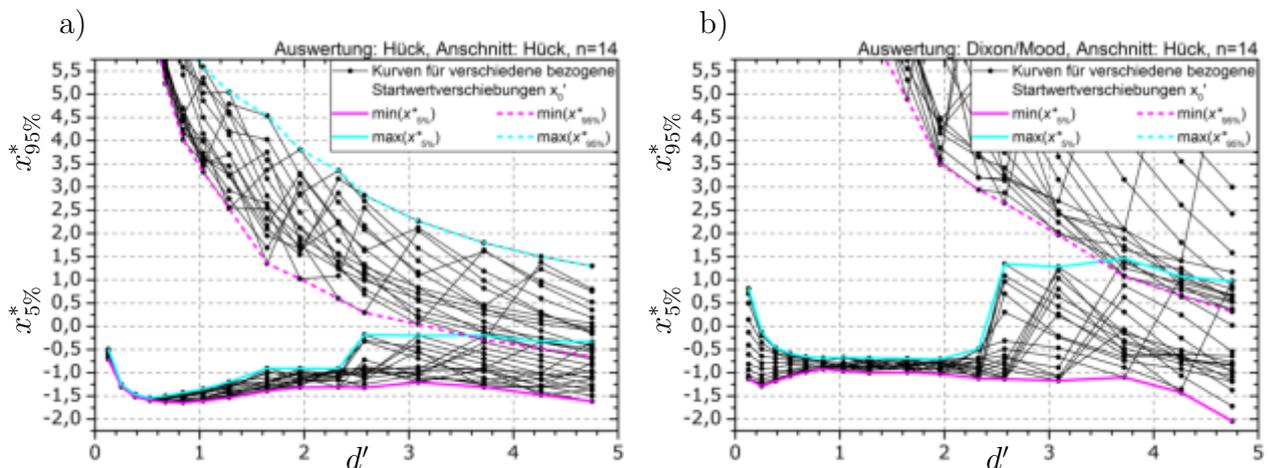


Abb. 3.29: Minimale und maximale Fehler des Schätzers x^* für Treppenstufenversuche mit $n = 14$ gültigen Versuchen und Anschnittsdefinition nach HÜCK für unterschiedlich gewählte Startwerte x_0 . Auswertung nach HÜCK in a) sowie DEUBELBEISS in b).

3 Simulation von Versuchen zur Bestimmung der Dauerfestigkeit

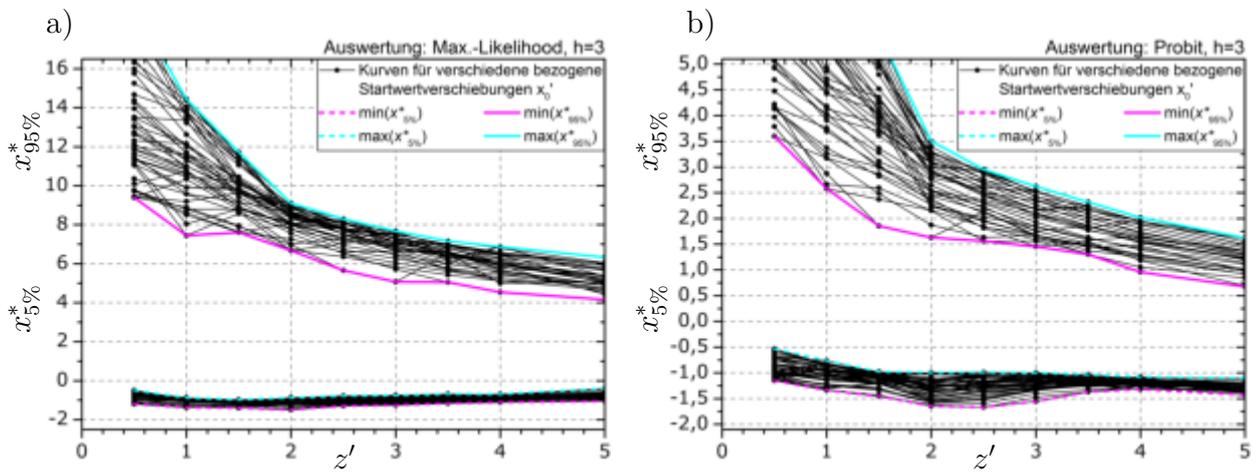


Abb. 3.30: Minimale und maximale Fehler des Schätzers x^* für Horizontversuche mit $n = 15$ gültigen Versuchen für unterschiedlich gewählte Startwerte x_0 . Auswertung nach MAX.-LIKELIHOOD-METHODE in a) sowie PROBIT-METHODE in b).

4 Analyse der Simulationsergebnisse

4.1 Vom Auswerteverfahren unabhängige Ergebnisse

Die Ergebnisse der verschiedenen Treppenstufenauswerteverfahren unterscheiden sich zwar in der Schätzung des Erwartungswertes und der Standardabweichung, nicht aber in der Form der Versuchsabfolge.¹ Die Anzahl der mit Versuchen belegten Horizonte oder die Anzahl von für den Anschnitt aufzuwendenden Proben können daher unabhängig von der Auswertemethode betrachtet werden. Beide Größen liefern dabei bereits Hinweise auf die zu schätzenden Größen der Grundgesamtheit.

Die relativen Häufigkeiten mit denen in der Simulation Versuchsreihen mit einer bestimmten Horizontanzahl erreicht werden, sind für Versuchsreihen mit $n = 14$ gültigen Einzelversuchen in der Abbildung 4.1 exemplarisch dargestellt. Diese relativen Häufigkeiten können dabei als Eintretenswahrscheinlichkeiten interpretiert werden. Beispielsweise kann bei einem tatsächlichen Stufenabstand von $d' = 1$ in etwa der Hälfte aller Versuchsreihen eine Versuchsreihe mit $h = 4$ Horizonten erwartet werden (Teilbild c). Liegt hingegen eine Versuchsreihe mit $n = 14$ auf $h = 6$ Horizonten vor, dann ist von einem deutlich kleineren bezogenen Stufenabstand auszugehen, wie aus dem Teilbild e) der Abbildung zu entnehmen ist. Genauso lässt sich aus den Teilbildern a) und b) erkennen, dass bei 14 gültigen Einzelversuchen und großen Stufenabständen mit Versuchsreihen mit zwei oder drei belegten Horizonten gerechnet werden kann. Darüber hinaus wird im Vergleich dieser Teilbilder auch deutlich, dass Versuchsreihen mit drei Lastniveaus meist dann entstehen, wenn der bezogene Versuchsreihenstartwert x'_0 ein Vielfaches des bezogenen Stufenabstandes d' beträgt, also der tatsächliche Erwartungswert der Grundgesamtheit mit einem Lastniveau zusammenfällt. Versuchsreihen mit zwei Lastniveaus entstehen hingegen dann häufig, wenn der Erwartungswert der Grundgesamtheit genau zwischen diesen zu liegen kommt. Für Versuchsreihen mit einer anderen Anzahl gültiger Einzelversuche sind die entsprechenden Ergebnisse im Anhang A.1.1 ab Seite 180 grafisch dargestellt.

Ebenfalls vom gewählten Auswerteverfahren unabhängig ist die für den Anschnitt der Versuchsreihe im Mittel aufzuwendende Anzahl von Versuchen n_u . In Abbildung 4.2 ist die relative Häufigkeit der erhaltenen Versuchsreihen für unterschiedliche n_u für $n = 14$ gültige Versuche dargestellt. Es wird deutlich, dass bei einer Versuchsreihe mit mehr als 4 Versuchen

¹Es werden an dieser Stelle nur die Ergebnisse für Treppenstufenversuche mit einer Anschnittsdefinition nach HÜCK betrachtet. Ebenso wird die Tatsache vernachlässigt, dass Versuche im Anschnitt der Versuchsreihe bei einer Auswertung nach HÜCK oder DEUBELBEISS durch das Hinzufügen des fiktiven Versuchs wieder für gültig erklärt werden können und sich dadurch die Horizontanzahl und die Anzahl gültiger Versuche erhöht.

4 Vergleich der Simulationsergebnisse

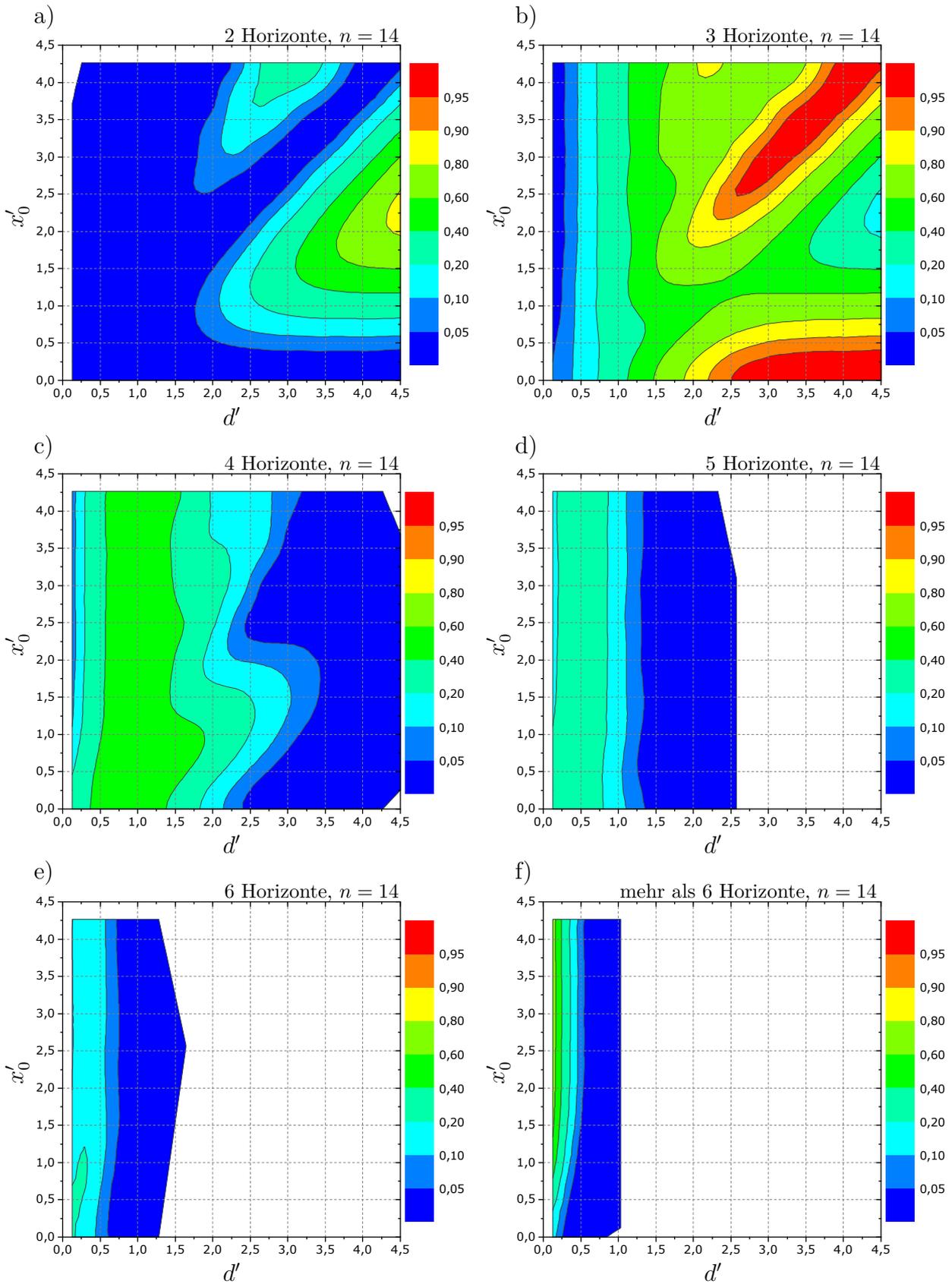


Abb. 4.1: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit $n = 14$ Einzelversuchen und im Ergebnis genau $h = 2$ in a), $h = 3$ in b), $h = 4$ in c), $h = 5$ in d), $h = 6$ in e) oder mehr als 6 belegten Horizonten in f).

im Anschnitt von einem deutlich zu klein gewählten bezogenen Stufenabstand sowie einer deutlichen Verschätzung bei der Wahl des Startwertes ausgegangen werden kann.

Folgt man der Annahme, dass der bei einem Treppenstufenversuch ermittelte Versuchsablauf auch der wahrscheinlichste aller möglichen Versuchsabläufe ist, dann könnte für eine Versuchsreihe mit $n = 14$ gültigen Einzelversuchen bereits mit den Abbildungen 4.1 und 4.2 ein Schätzwert für die Parameter der Grundgesamtheit ermittelt werden. Hierzu wird in Abbildung 4.1 eine Gerade parallel zur Ordinatenachse durch den Bereich der größten Eintretenswahrscheinlichkeit gelegt. An der Abszissenachse kann dann das vermutete Verhältnis von Stufenabstand zu Standardabweichung abgelesen werden. Aus Abbildung 4.2 ist dann bei diesem Verhältnis, im Bereich der größten Eintretenswahrscheinlichkeit, der bezogene Versuchsreihenstartwert x'_0 abzulesen. Aus letzterem kann dann der Erwartungswert der Grundgesamtheit geschätzt werden. Der Schätzer der Standardabweichung wird aus dem bereits zuvor ermittelten bezogenen Stufenabstand d' geschätzt. Da bei dieser Form der Parameterschätzung die tatsächliche Verteilung der Versuche auf die einzelnen Lastniveaus nicht berücksichtigt wird, ist anzunehmen, dass die Güte der Schätzung keine zufriedenstellenden Ergebnisse liefert. Deshalb wird dieses Vorgehen im Weiteren nicht näher untersucht.

Davon unberührt ist aus den oberen Abbildungen ersichtlich, dass bei einer statistischen Auswertung der Simulationsergebnisse der Gültigkeitsbereich der getroffenen Aussagen immer zu überprüfen ist. Wird z. B. die Analyse auf den Bereich $0,25 < d' < 2$ beschränkt, werden fast alle Versuchsreihen herausgefiltert, die im Ergebnis nur zwei belegte Horizonte aufweisen. Umgekehrt ist die Aussagekraft eines Ergebnisses, bei dem nur Versuchsreihen mit 4 belegten Horizonten ausgewertet werden, im Bereich $d' < 2$ aufgrund der geringen Anzahl an Simulationsergebnissen eingeschränkt. Daher wird darauf verzichtet, die Ergebnisse nach Anzahl der belegten Lastniveaus oder Anzahl der ungültigen Einzelversuche zu filtern.

4 Vergleich der Simulationsergebnisse

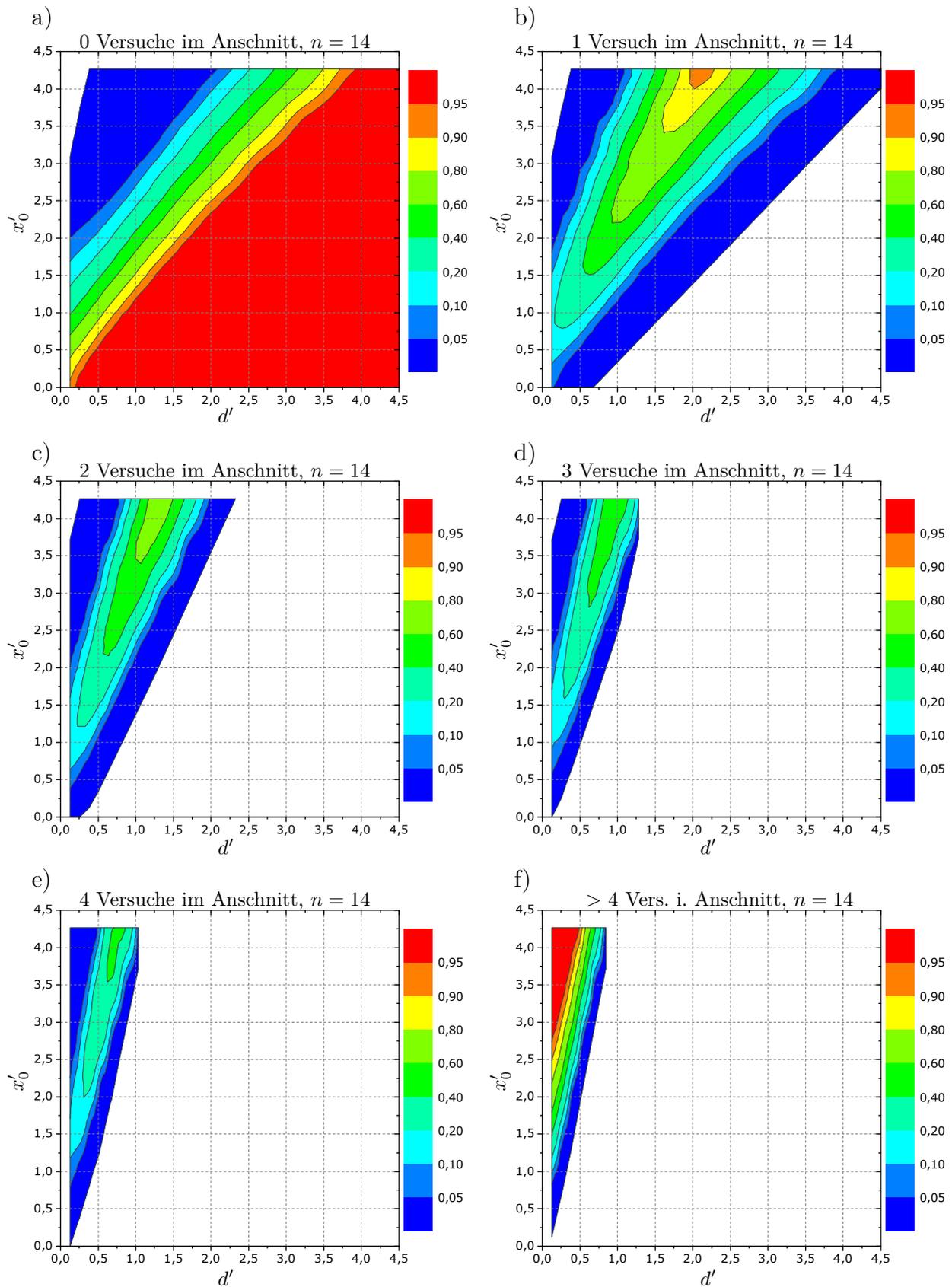


Abb. 4.2: Relative Häufigkeit der Treppenstufenversuchsreihen mit $n_u = 0$ in a), $n_u = 1$ in b), $n_u = 2$ in c), $n_u = 3$ in d), $n_u = 4$ in e) oder mehr als 4 in f) Einzelversuchen im Abschnitt, bei jeweils $n = 14$ gültigen Einzelversuchen.

4.2 Schätzung des Erwartungswertes bei bekannter Standardabweichung

4.2.1 Anteil auswertbarer Versuchsreihen

Bei allen durchgeführten Simulationsrechnungen wurden die einzelnen Versuchsreihen stets für eine vorgegebene Anzahl an zu erreichenden gültigen Einzelversuchen ermittelt. Dadurch liefern fast alle Treppenstufenauswertungen stets Schätzwerte für den Erwartungswert der Grundgesamtheit. Lediglich bei der Auswertung nach der PROBIT-METHODE sind Schätzungen des Erwartungswertes nur dann möglich, wenn gleichzeitig auch eine positive Schätzung der Standardabweichung ($s > 0$) vorliegt. Da dadurch die relative Anzahl auswertbarer Versuchsreihen genauso groß ist wie bei der Schätzung der Standardabweichung, sind die Ergebnisse für die PROBIT-METHODE dem Abschnitt 4.3.1 zu entnehmen.

Bei der Auswertung der Horizontversuche werden sowohl nach der PROBIT-METHODE als auch nach der MAX.-LIKELIHOOD-METHODE beide Parameter der Grundgesamtheit gemeinsam geschätzt. Als gültig werden auch hier nur Schätzungen betrachtet, die eine Standardabweichung größer Null schätzen. Der Anteil gültiger Versuchsreihen ist daher bei der Schätzung des Erwartungswertes und der Standardabweichung gleich groß und wird im Abschnitt 4.3.1 diskutiert.

4.2.2 Mittlerer zufälliger und systematischer Schätzfehler

Werden Treppenstufenversuche und Horizontversuche mit einem Startwert durchgeführt, der genau dem Erwartungswert der Grundgesamtheit entspricht, so ist die Schätzung des Erwartungswertes der Grundgesamtheit erwartungstreu. Liegt der Startwert der Versuchsreihe oberhalb bzw. unterhalb der tatsächlichen mittleren Schwingfestigkeit, wird diese von allen betrachteten Auswertverfahren systematisch über- bzw. unterschätzt. Allerdings führt eine unbegrenzte Vergrößerung der Startwertverschiebung nicht zu einem unbegrenzten Anwachsen des systematischen Schätzfehlers. Vielmehr können mit den Faktoren $G_{m',\max}$ bzw. $G_{m',\min}$, vgl. Abschnitt 3.4.2.2 ab S. 65, obere und untere Grenzen des systematischen Schätzfehlers berechnet werden, die von den vorliegenden Simulationsergebnissen nicht überschritten werden.² In praktischen Versuchen ist die tatsächliche Startwertverschiebung unbekannt. Der tatsächliche systematische Fehler G_m schwankt von Versuchsreihe zu Versuchsreihe und wird daher mit Hilfe des gemittelten Faktors \bar{G}_m des Schätzers m' erfasst. Bei bekannter Standardabweichung σ_G berechnet sich damit der mittlere systematische Fehler des experimentell

²Das liegt vor allem daran, dass nicht die Anzahl der für eine Versuchsreihe nötigen Versuche betrachtet wird, sondern die in einer Versuchsreihe im Sinne der Auswertung gültigen Versuche gezählt werden. Eine beliebig größere Startwertverschiebung führt dann zu einer Vergrößerung des Versuchsreihenanschnitts und damit zur Zunahme der als ungültig betrachteten Einzelversuche, siehe Abschnitt 4.1. Bei der Auswertung der Horizontversuche tritt ein ähnliches Verhalten dann ein, wenn die zum Auffinden des ersten Lasthorizontes nötigen Versuche mit einem möglichst kleinen Stufenabstand (mindestens $d \leq \sigma_G$) positioniert und dann mindestens 5 Versuche auf jedem Lasthorizont durchgeführt werden.

bestimmten Erwartungswertes der Festigkeit durch

$$G_m = \bar{G}_m \cdot \sigma_G$$

Wie stark die mittlere systematische Verschätzung ausfällt, hängt dann hauptsächlich von der Anzahl der gültigen Versuche, bei Treppenstufenversuchen zusätzlich von der Größe des Stufenabstandes und bei den Horizontversuchen von der Spreizung der Horizonte ab. Für Treppenstufenversuche mit geringer Versuchsanzahl und kleinem Stufenabstand liegt die systematische Verschätzung des Erwartungswertes hier im Bereich von 0,1 bis 0,3 Standardabweichungen der Grundgesamtheit, vgl. Abbildung A.13 und A.14 ab S. 192. Bei Horizontversuchen fallen die systematischen Verschätzungen geringer aus und erreichen meist nicht mehr als 0,1 Standardabweichungen der Grundgesamtheit.

Alle untersuchten Auswertemethoden der **Treppenstufenversuche** zeigen, dass sich mit zunehmender Versuchsanzahl die mittlere systematische Verschätzung verringert, die Methoden also den Erwartungswert der Grundgesamtheit asymptotisch erwartungstreu schätzen. Dieses asymptotische Verhalten ist bei den Auswertungen mit der MAX.-LIKELIHOOD-METHODE und der PROBIT-METHODE am stärksten ausgeprägt, d.h. die Erhöhung der Anzahl der Einzelversuche führt hier schneller zu einer nahezu erwartungstreuen Schätzung als nach der Auswertung nach DEUBELBEISS oder HÜCK, vgl. Abbildung 4.3. Der Bias bei der Auswertemethode nach DIXON/MOOD zeigt bei sehr kleinen Stufenabständen ein ähnliches Verhalten wie bei den Auswertungen nach HÜCK oder DEUBELBEISS. Bei großen Stufenabständen folgt er hingegen dem systematischen Fehler der MAX.-LIKELIHOOD-METHODE. Der bei den Treppenstufenversuchen untersuchte Einfluss der Anschnittsdefinition ist bei den Auswertungen nach DIXON/MOOD quasi nicht vorhanden, da die Auswertung des weniger oft eingetretenen Ereignisses nahezu immer dazu führt, dass schlussendlich die Anschnittsdefinition nach DIXON/MOOD verwendet wird. Für die Auswertungen nach HÜCK und DEUBELBEISS verringert sich der Bias bei der Verwendung der Anschnittsdefinition nach DIXON/MOOD. Der Grund hierfür ist vor allem in der Zunahme der Anzahl der Versuche pro Versuchsreihe zu suchen, die bei dieser Definition des Anschnittes nötig ist, um die vorgegebene Anzahl gültiger Versuche zu erreichen. Für die Auswertungen nach der MAX.-LIKELIHOOD-METHODE oder der PROBIT-METHODE vergrößert sich der Bias etwas. Dies ist aber in Anbetracht der geringen Höhe der Verschätzungen zu vernachlässigen.

Auch die Schätzer des Erwartungswertes aus **Horizontversuchen** sind nur asymptotisch erwartungstreu. Wie in den Abbildungen 4.3 zu erkennen ist, sind die mittleren systematischen Verschätzungen für kleine Horizontspreizungen sehr gering. Das liegt auch daran, dass für sinnvolle Auswertungen mit den Horizontverfahren einfach von vornherein eine größere Anzahl an Einzelversuchen nötig ist. Ein Vergleich der Auswertemethoden der Horizontversuche zeigt, dass die Wahl der Auswertemethode kaum Einfluss auf die Größe des Bias besitzt, siehe Abbildung 4.3. Wichtiger ist die richtige Wahl von Horizontspreizung und Horizontanzahl. Während es bei kleinen Horizontspreizungen vorteilhaft ist, die Versuche auf möglichst wenige Horizonte zu verteilen, führt selbiges Vorgehen bei großen Horizontspreizungen zu größeren systematischen Schätzfehlern.

Die zufälligen Fehler der Erwartungswertschätzung ergeben sich aus den berechneten mittleren Standardfehlern \bar{C}_m , des Schätzers m' , vgl. Abschnitt 3.4.2.2 ab S. 65, mit

$$C_m = \bar{C}_m \cdot \sigma_G.$$

4.2 Schätzung des Erwartungswertes bei bekannter Standardabweichung

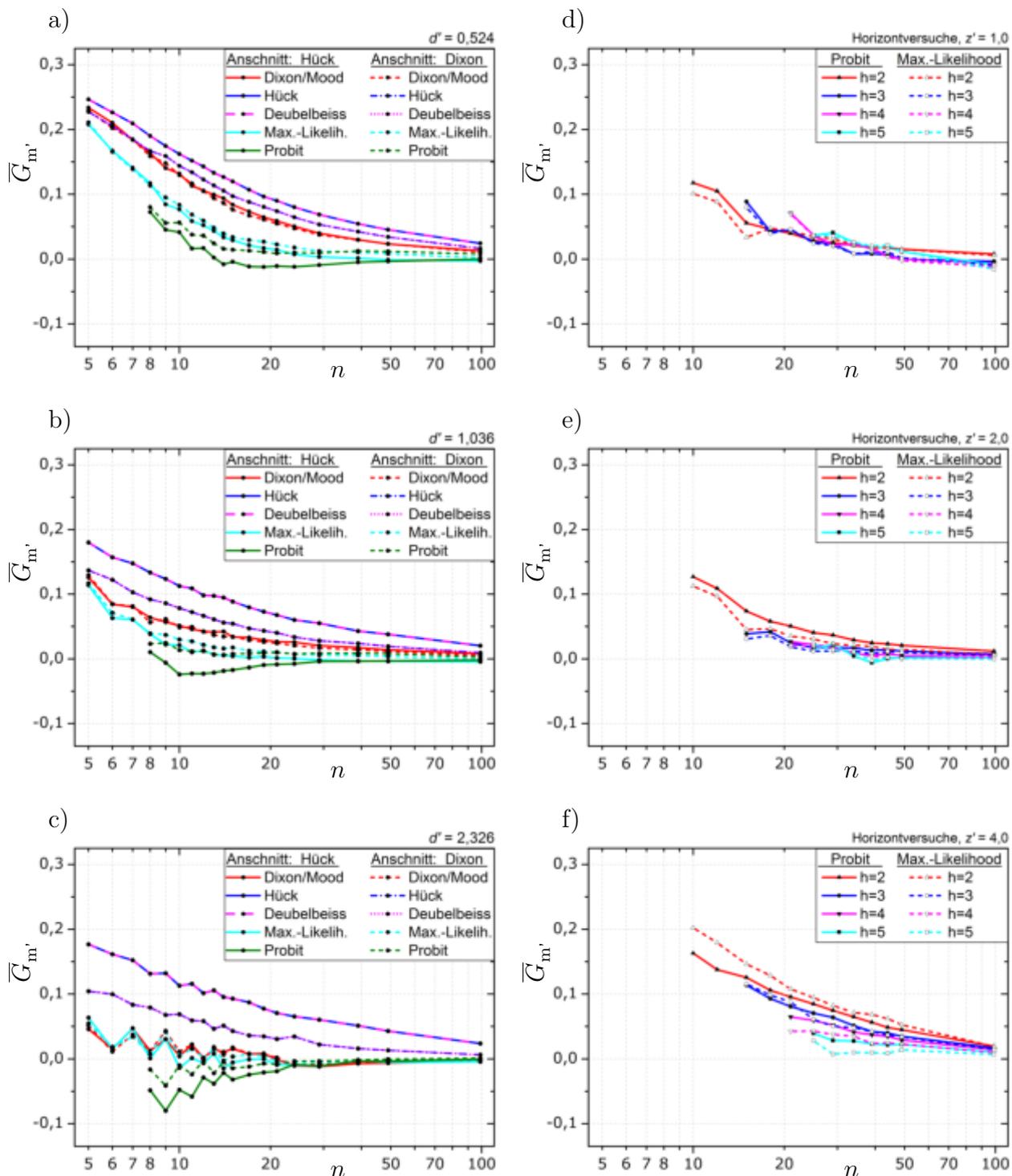


Abb. 4.3: Mittlere Verzerrungen des Schätzers m' für Treppenstufenversuche bei verschiedenen Anschnittsdefinitionen und einem Stufenabstand von $d' = 0,524$ in a), $d' = 1,036$ in b) und $d' = 2,326$ in c) sowie für Horizontversuche mit verschiedenen Spreizungen von $z' = 1$ in d), $z' = 2$ in e) und $z' = 4$ in f).

4 Vergleich der Simulationsergebnisse

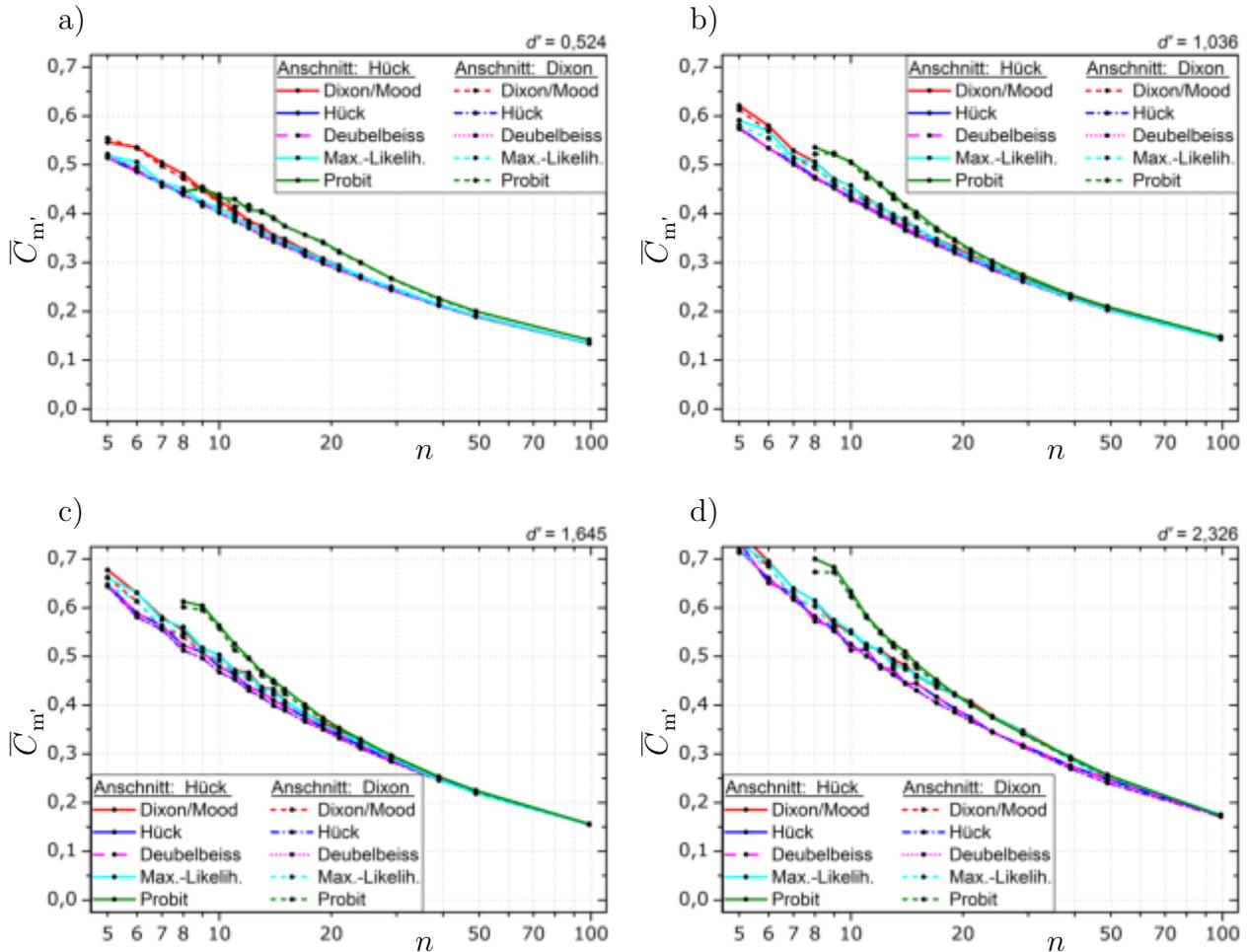


Abb. 4.4: Mittlerer Standardfehler des Schätzers m' für Treppenstufenversuche bei verschiedenen Treppenstufenabständen. Für $d' = 0,524$ in a), $d' = 1,036$ in b), $d' = 1,645$ in c) sowie $d' = 2,326$ in d).

Ebenso wie die Verzerrung schwankt auch der Standardfehler des Schätzers m' bei der Variation der Startwertverschiebung. Die dabei auftretenden Unterschiede sind allerdings gering. Der mittlere Standardfehler wird um die sich aus der Schwankung der Verzerrung ergebende Unsicherheit ergänzt, wie es im Abschnitt 3.4.2.2 beschrieben ist. Die Simulationsergebnisse für die **Treppenstufenversuche** sind in der Abbildung 4.4 dargestellt. Es ist zu erkennen, dass die Auswertung mit der PROBIT-METHODE die größten zufälligen Schätzfehler liefert. Mit allen anderen Treppenstufenauswertemethoden werden geringere, sich kaum unterscheidende Ergebnisse erzielt. Die Schätzungen nach HÜCK sowie DEUBELBEISS zeigen stets übereinstimmende Ergebnisse und die geringsten Streuungen. Eine Anschnittsdefinition nach DIXON führt für alle Auswertemethoden bei größeren Stufenabständen zu einer geringfügigen Verbesserung der Schätzungen. Dieser Einfluss ist jedoch deutlich geringer als die richtige Wahl des Treppenstufenabstandes, welcher möglichst kleiner als die tatsächliche Standardabweichung ($d' < 1$) gewählt werden sollte. Bei der Schätzung des Erwartungswertes mittels **Horizontversuchen** treten die in der Abbildung 4.5 dargestellten mittleren Standardfehler auf. Unabhängig von der tatsächlichen Auswertemethode sind die damit ermittelten zufälligen Fehler größer als bei der Schätzung des Erwartungswertes mit Treppenstufenversuchen. Erwartungsgemäß ist eine bessere Schätzung möglich, wenn bei einer

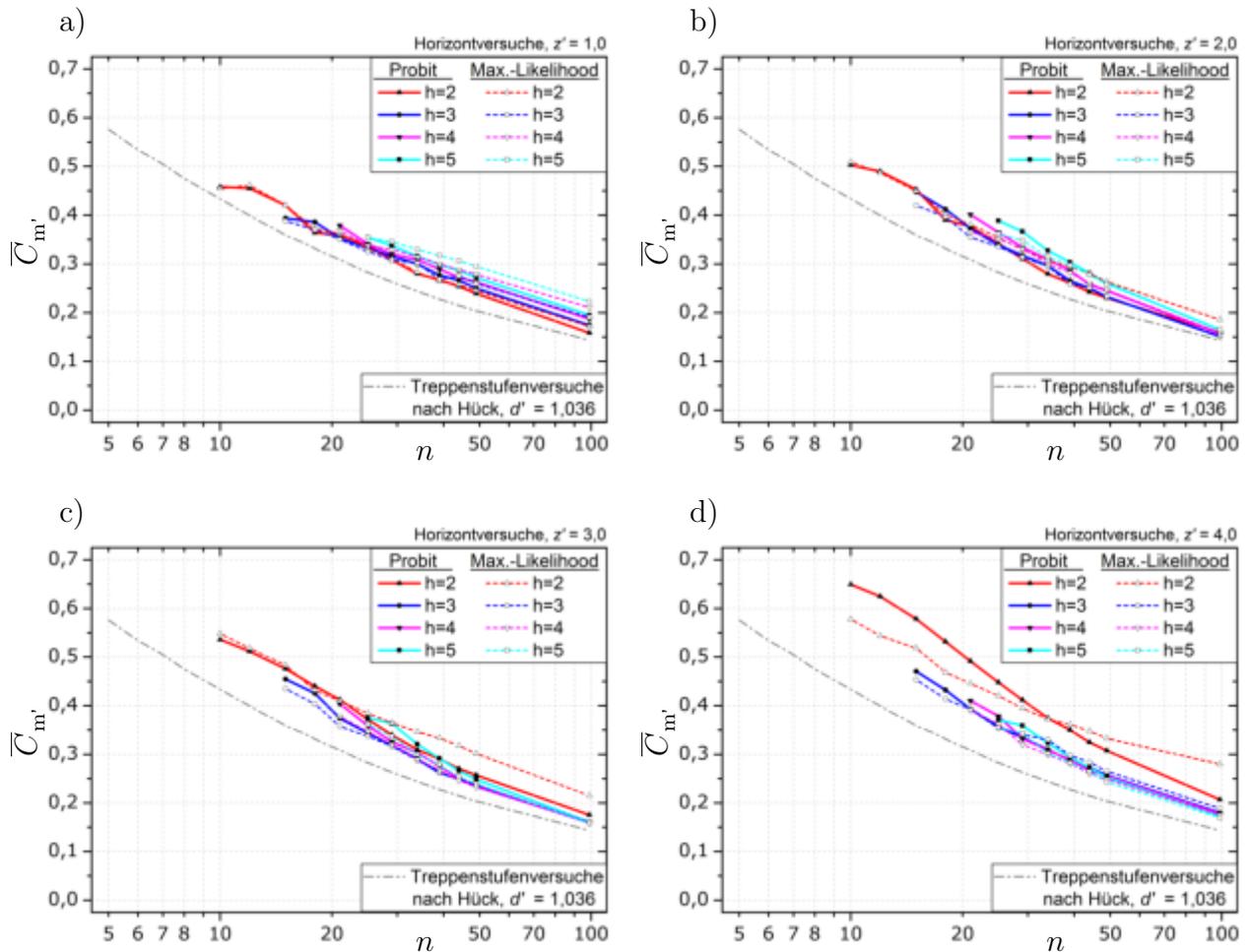


Abb. 4.5: Mittlerer Standardfehler des Schätzers m' für Horizontversuche bei verschiedenen Horizontspreizungen. Für $z' = 1,0$ in a), $z' = 2,0$ in b), $z' = 3,0$ in c) sowie $z' = 4,0$ in d).

kleinen Horizontspreizung die Einzelversuche auf möglichst wenige Horizonte verteilt werden. Bei sehr großen Horizontspreizungen sollten dagegen mehrere Horizonte verwendet werden. Da in den meisten Fällen die Horizontspreizung nicht exakt festgelegt werden kann, ist für die Versuchsplanung die Wahl von $h = 3$ Horizonten zu empfehlen, wenn der Erwartungswert der Grundgesamtheit geschätzt werden soll. Die Wahl der Auswertemethode hat nahezu keinen Einfluss auf die Güte der Schätzung. Um den Erwartungswert mit den Horizontmethoden jedoch ähnlich genau wie in Treppenstufenversuchen nach HÜCK und $d' \approx 1$ zu ermitteln, d. h. ähnlich große zufällige Schätzfehler zu erhalten, sind nach Abbildung 4.5 etwa 30 % bis 50 % mehr Einzelversuche nötig.

4.2.3 Verallgemeinerter Standardfehler

Der systematische und der zufällige Schätzfehler des Schätzers m' können, wie in Abschnitt 3.4.2.3 beschrieben, zu einem verallgemeinerten Standardfehler zusammengefasst werden. Dieser gibt dann die mittlere zu erwartende Abweichung des im Experiment ermittelten Schätzwertes vom tatsächlichen Wert des Erwartungswertes in Vielfachen der tatsächlichen Standardabweichung an, wenn über den Startwert oder den Anschnitt der Versuchsreihe

keine weiteren Informationen bekannt sind. Das ist im Allgemeinen dann der Fall, wenn auf Literaturangaben zurückgegriffen wird, bei denen der tatsächliche Versuchsablauf nicht dokumentiert ist. Mit dem verallgemeinerten Standardfehler ist es möglich, eine Messunsicherheit U_m mit

$$U_m = U_{m'} \cdot \sigma_G$$

für den geschätzten Erwartungswert der Grundgesamtheit anzugeben.

Für die simulierten **Treppenstufenversuche** ergeben sich die in der Abbildung 4.6 dargestellten verallgemeinerten Standardfehler. Für Treppenstufenauswertungen nach HÜCK sind diese direkt mit dem von HÜCK in [21] angegebenen und ebenfalls in der Abbildung eingetragenen Faktor C_m vergleichbar.³ Die Ergebnisse der simulierten Treppenstufenversuche zeigen, dass für sehr kleine Stufenabstände die Wahl der Anschnittsdefinition keine Bedeutung besitzt und die Ergebnisse mit der MAX.-LIKELIHOOD-METHODE am besten geschätzt werden. Bei größeren Stufenabständen ($d' > 1$) liefert hingegen die Auswertung nach HÜCK oder DEUBELBEISS mit der Anschnittsdefinition nach DIXON/MOOD die besten Ergebnisse. Die Auswertung von Treppenstufenversuchen mit der PROBIT-METHODE liefert zu keiner Zeit treffsichere Schätzwerte für den Erwartungswert der Grundgesamtheit⁴. Die von HÜCK in [21] angegebenen Schätzfehler werden nur bei Stufenabständen $d' > 1$ erreicht und müssen nach dieser Auswertung insgesamt als zu niedrig angesehen werden. Auch wird deutlich, dass die angegebenen Faktoren C_m nicht mit der geschätzten, sondern mit der tatsächlichen Standardabweichung multipliziert werden müssen, um den Standardfehler des Erwartungswertschätzers zu erhalten.

Die Auswertung der **Horizontversuche** liefert die in der Abbildung 4.7 gezeigten Ergebnisse. Zum besseren Vergleich ist hier der aus den Simulationen ermittelte verallgemeinerte Standardfehler für die Auswertungen der Treppenstufenversuche nach HÜCK bei einem Stufenabstand von $d' \approx 1$ eingetragen. Es ist zu erkennen, dass mit den Horizontversuchen nicht die gleiche Treffsicherheit bei der Ermittlung des Erwartungswertes wie mit den Treppenstufenversuchen erreicht werden kann. Bei sehr kleinen Horizontspreizungen ($z' < 1,5$) liefert die PROBIT-METHODE bessere Schätzergebnisse als die MAX.-LIKELIHOOD-METHODE. Das Abgrenzungsverfahren, also Horizontversuche auf zwei Lasthorizonten, schätzt dann am genauesten. Für größere Horizontspreizungen ($z' > 2$) ist die Auswertung auf zwei Horizonten hingegen sehr ungenau und bei Versuchen mit mehr als zwei Horizonten sind die Unterschiede zwischen den Auswertemethoden vernachlässigbar. Die Horizontversuche auf drei Horizonten zeigen sich sehr robust gegenüber einer Variation der Spreizung und liefern unabhängig von der Auswertemethode stets recht treffsichere Schätzwerte im Vergleich zu Versuchen mit zwei, vier oder fünf Horizonten.

Wird der verallgemeinerte Standardfehler mit der bekannten tatsächlichen Standardabweichung multipliziert, ergibt sich die „Messunsicherheit“ des geschätzten Erwartungswertes. Wenn nur 10 gültige Versuche für eine Versuchsreihe gefordert werden, fällt diese Unsicherheit

³Nach HÜCK wird der Standardfehler s_m der Erwartungswertschätzung aus der geschätzten Standardabweichung durch $s_m = s \cdot C_m$ berechnet. In der hier verwendeten Betrachtung folgt der Standardfehler aus $s_m = U_{m'} \cdot \sigma_G$.

⁴Die Auswertung nach der PROBIT-METHODE liefert tatsächlich für sehr kleine Stufenabstände und wenige Einzelversuche recht geringe Streuungen des Erwartungswertschätzers. Dies liegt aber daran, dass in diesen Fällen ein Großteil der Versuchsreihen nicht auswertbar, die Stichprobe also zensiert ist.

4.2 Schätzung des Erwartungswertes bei bekannter Standardabweichung

selbst bei günstiger Versuchsführung nicht unter das 0,4-Fache der tatsächlichen Standardabweichung. Sie kann durch die Forderung nach 20 gültigen Versuchen nicht unter das 0,3-fache der tatsächlichen Standardabweichung gedrückt werden. Selbst bei der Forderung nach 50 gültigen Versuchen beträgt die Unsicherheit bei der Bestimmung des Erwartungswertes etwa 20 % der Standardabweichung. Mit Horizontversuchen kann diese Genauigkeit erst mit etwa 70 Einzelversuchen erreicht werden.

4 Vergleich der Simulationsergebnisse

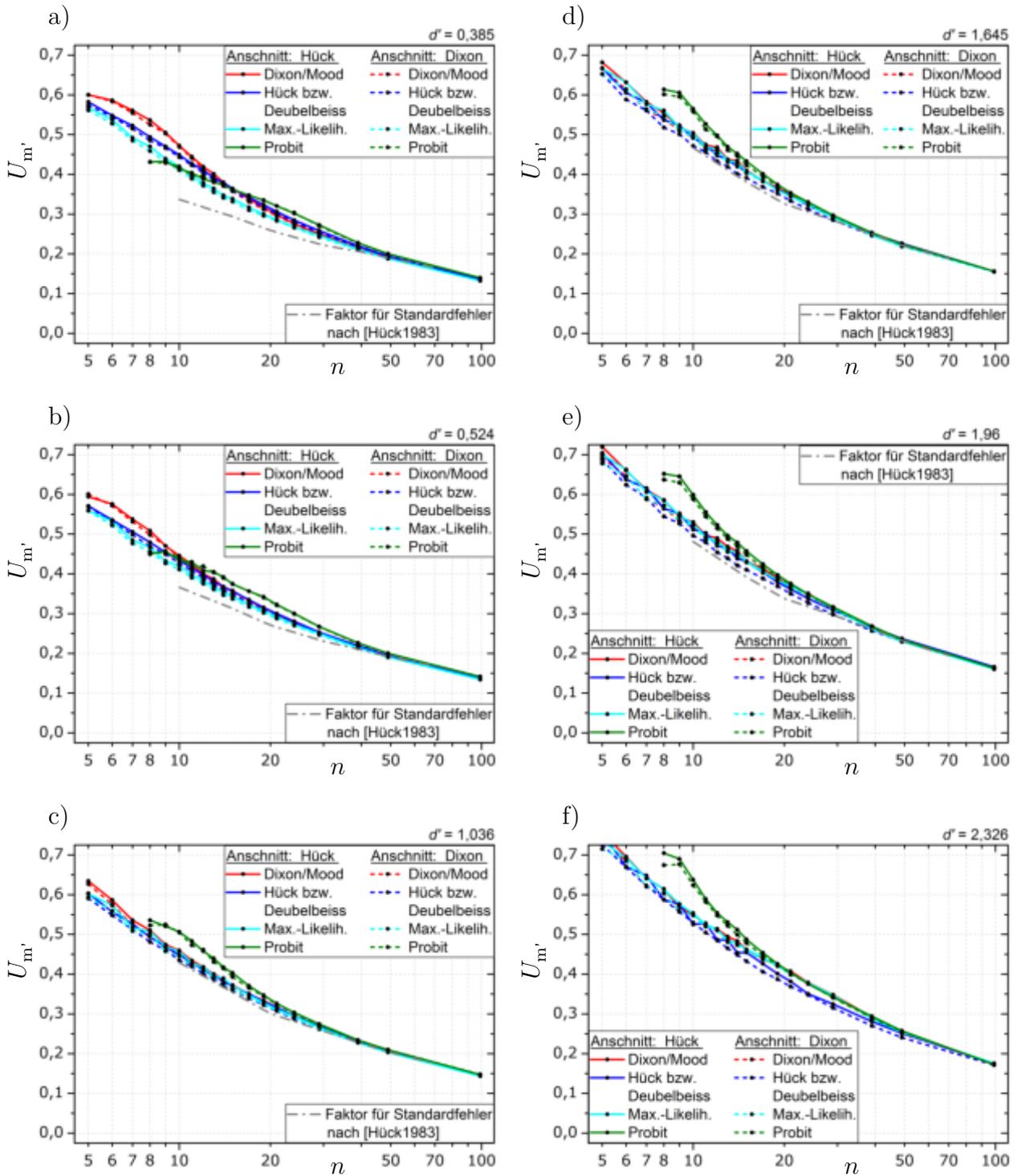


Abb. 4.6: Verallgemeinerter Standardfehler des Schätzers m' für Treppenstufenversuche bei verschiedenen Treppenstufenabständen d' .

4.2 Schätzung des Erwartungswertes bei bekannter Standardabweichung

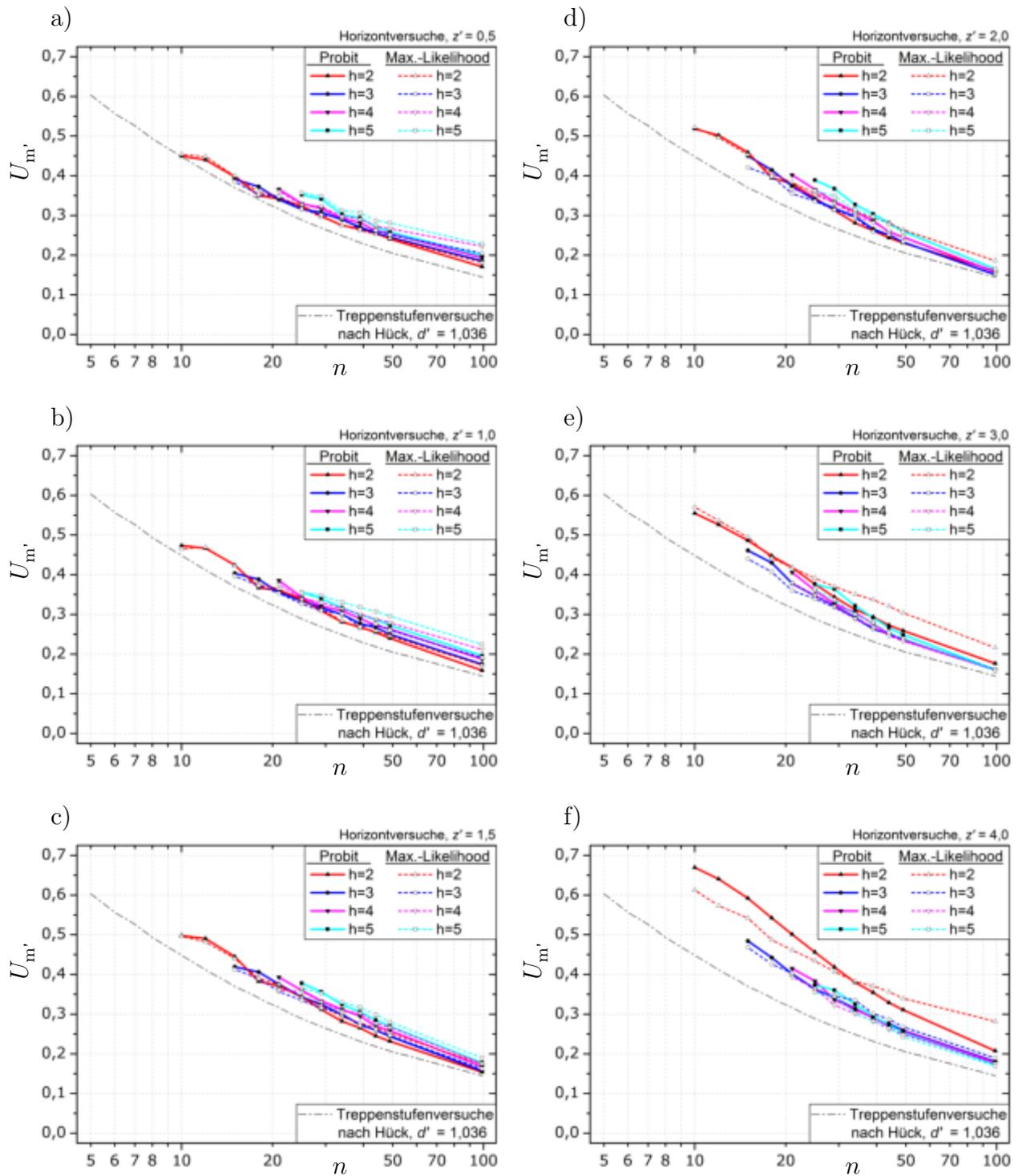


Abb. 4.7: Verallgemeinerter Standardfehler des Schätzers m' für Horizontversuche bei verschiedenen Horizontspreizungen z' .

4.2.4 RMS-Fehler

Die Berechnung des mittleren RMS-Fehlers, vgl. Abschnitt 3.4.2.4 ab S. 70, liefert als Ergebnis nahezu gleiche Zahlenwerte wie die Auswertung des verallgemeinerten Standardfehlers. In den Abbildungen 4.8 und 4.9 sind daher lediglich exemplarische Ergebnisse für Treppenstufen- und Horizontversuche dargestellt. Es gelten die im vorherigen Abschnitt 4.2.3 gezogenen Schlüsse. Treppenstufenversuche liefern für alle Auswertemethoden außer der PROBIT-METHODE bessere Schätzungen des Erwartungswertes der Grundgesamtheit als es mit Horizontversuchen möglich ist.

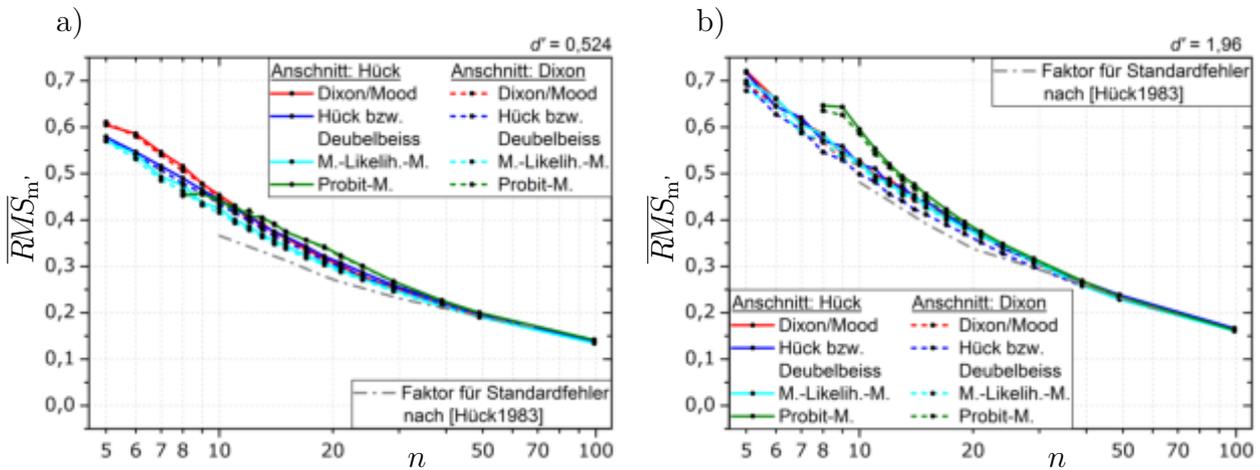


Abb. 4.8: Mittlerer RMS-Fehler des Schätzers m' für Treppenstufenversuche bei einem Stufenabstand von $d' = 0,524$ in a) und $d' = 1,96$ in b).

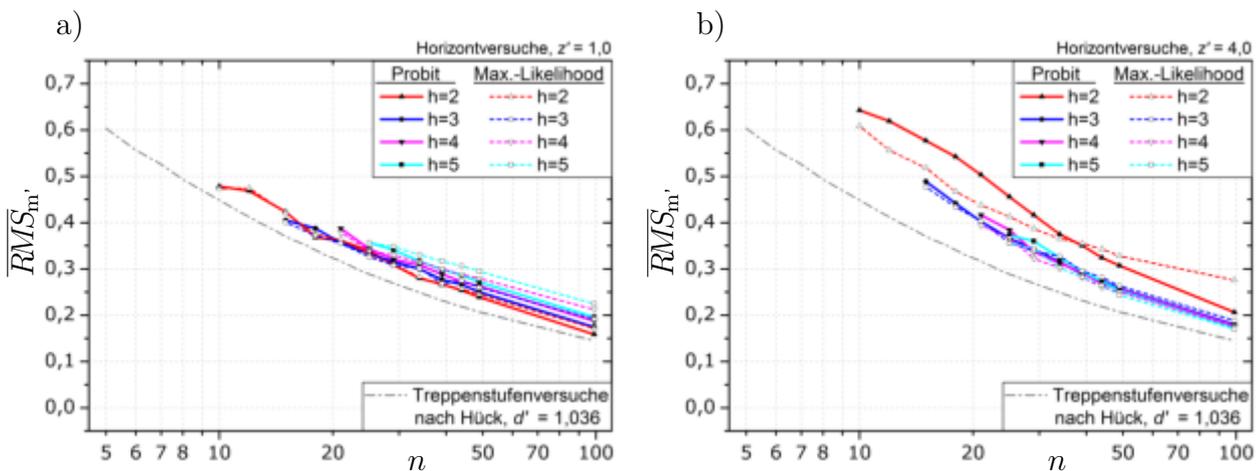


Abb. 4.9: Mittlerer RMS-Fehler des Schätzers m' für Horizontversuche mit Horizontspreizungen von $z' = 1,0$ in a) und $z' = 4,0$ in b).

4.2.5 Empirischer Vertrauensbereich des geschätzten Erwartungswertes

Anhand der ermittelten empirischen Quantile des Schätzers m' kann ein Vertrauensintervall für den Schätzer des Erwartungswertes abgeleitet werden. Die Quantile geben dabei den Abstand zwischen dem tatsächlichen Erwartungswert der Grundgesamtheit und dem 5 %- bzw. 95 %-Quantil aller dafür ermittelten Schätzwerte in Vielfachen der tatsächlichen Standardabweichung an. Ist über den Startwert der Versuchsreihe nichts bekannt, wird mit dem betragsmäßig größeren Quantil ein verallgemeinertes Quantil $m'_{|5\%|}$ für die Bildung des Vertrauensintervalls zu einer Irrtumswahrscheinlichkeit $\alpha = 10\%$ verwendet, vgl. Abschnitt 3.4.2.5 ab S. 71. Für den Schätzer des Erwartungswertes folgt daraus das Intervall

$$[m - m'_{|5\%|} \cdot \sigma_G, m + m'_{|5\%|} \cdot \sigma_G]. \quad (4.1)$$

In Abbildung 4.10 sind exemplarisch die Ergebnisse der **Treppenstufenversuche** dargestellt. Zusätzlich sind hier zum Vergleich die aus dem verallgemeinerten Standardfehler unter Annahme einer Normalverteilung geschätzten Quantile eingezeichnet, wie sie von HÜCK in [21] angegeben werden. Den vorliegenden Ergebnissen folgend, muss der Vertrauensbereich größer angenommen werden, als er in [21] angegeben ist. Für sehr kleine Stufenabstände können für die mittels MAX.-LIKELIHOOD-METHODE geschätzten Erwartungswerte die kleinsten Konfidenzintervalle angegeben werden. Bei größeren Stufenabständen erreichen alle Auswertemethoden etwa gleich große Konfidenzintervalle. Hier sind die Unterschiede in den Ergebnissen der verschiedenen Auswertemethoden so gering, dass eine eindeutige Bevorzugung einer Auswertemethode nicht gerechtfertigt erscheint. Die Wahl der Anschnittsdefinition besitzt keinen beachtenswerten Einfluss auf die Größe der Konfidenzintervalle.

Mit **Horizontversuchen**, vgl. Abbildung 4.11, ist es nicht möglich, bessere Intervallschätzungen als mit Treppenstufenversuchen zu erreichen. Generell sind bei der Verwendung von nur zwei Lasthorizonten nur Spreizungen bis $z' \leq 1,5$ anzustreben. Versuche auf drei Horizonten zeigen sich sehr robust gegenüber einer zu groß gewählten Spreizung der Horizonte. Versuche auf mehr als drei Horizonten durchzuführen, bietet keinen Vorteil, zumindest dann nicht, wenn die Schätzung des Erwartungswertes im Vordergrund steht.

Weiterhin ist für die Auswertung von Versuchen die Fragestellung interessant, ob sich aus dem im Abschnitt 4.2.3 betrachteten verallgemeinerten Standardfehler auch die verallgemeinerten Quantile des Schätzers m' bestimmen lassen. Beispielhaft wird hierzu der verallgemeinerte Standardfehler mit dem 95 %-Quantil $u_{0,95} = 1,645$ der Standardnormalverteilung multipliziert. Es zeigen sowohl die Ergebnisse der Treppenstufen- wie auch der Horizontversuche, dass mit diesem Vorgehen die empirischen Quantile nicht in ausreichendem Maß nachgebildet werden können. Wie in Abbildung 4.12 zu erkennen, nehmen bei Treppenstufenversuchen mit geringen Probenumfängen und kleinen Stufenabständen die empirischen Quantile meist deutlich größere Werte an. Lediglich bei der PROBIT-METHODE können die verallgemeinerten Quantile recht gut abgebildet werden. Bessere Übereinstimmungen können erreicht werden, wenn statt der Quantile der Standardnormalverteilung die Quantile der t-Verteilung (mit der Anzahl der gültigen Einzelversuche als Freiheitsgrad) zur Abschätzung der Quantile des Schätzers m' verwendet werden. Wie in der Abbildung 4.13 dargestellt, liegen die so ermittelten Quantile dann bei geringer Versuchsanzahl eher auf der sicheren Seite und könnten für eine konservative Abschätzung der Konfidenzintervalle verwendet werden.

4 Vergleich der Simulationsergebnisse

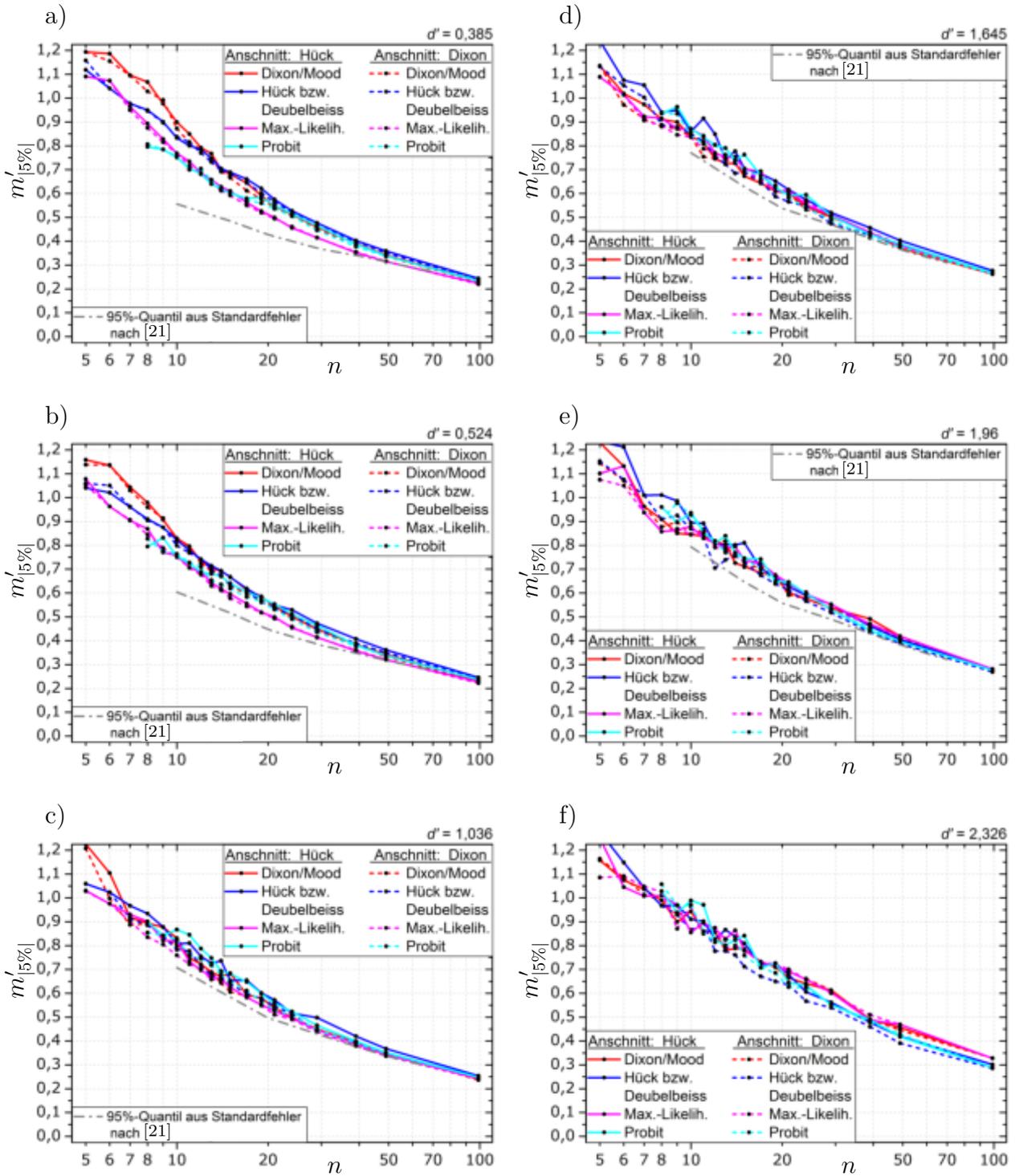


Abb. 4.10: Mittlere verallgemeinerte 5 %-Quantile des Schätzers m' für Treppenstufenversuche bei verschiedenen Treppenstufenabständen d' sowie mit dem nach HÜCK [21] aus dem Standardfehler s_m berechnete 95 %-Quantile. Für $d' = 0,385$ in a), $d' = 0,524$ in b), $d' = 1,036$ in c), $d' = 1,645$ in d), $d' = 1,960$ in e) sowie $d' = 2,326$ in f).

4.2 Schätzung des Erwartungswertes bei bekannter Standardabweichung

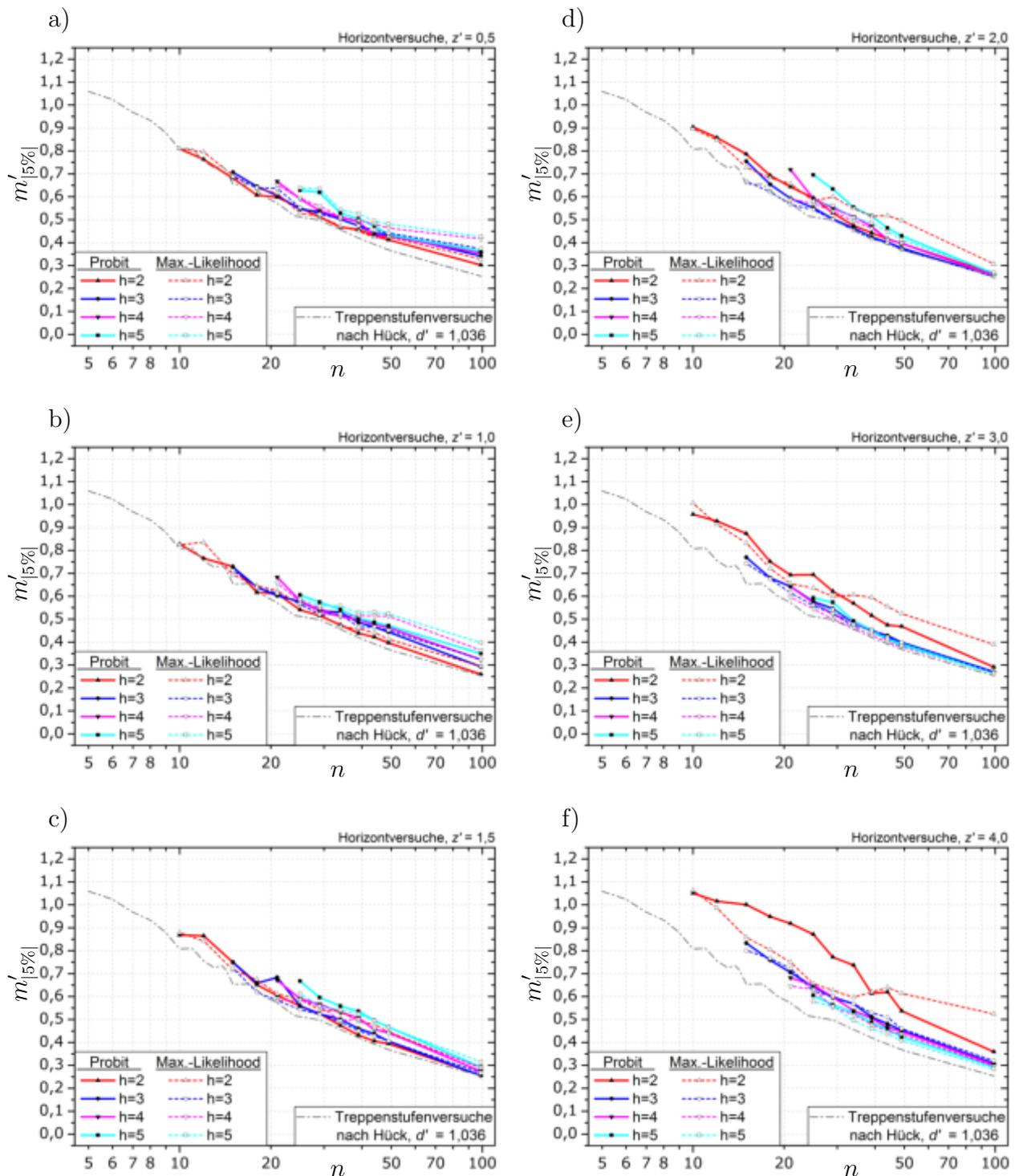


Abb. 4.11: Mittlere verallgemeinerte 5%-Quantile des Schätzers m' für Horizontversuche bei verschiedenen Horizontspreizungen z' , mit $z' = 0,5$ in a), $z' = 1,0$ in b), $z' = 1,5$ in c), $z' = 2,0$ in d), $z' = 3,0$ in e) sowie $z' = 4,0$ in f). Zum Vergleich sind jeweils die Simulationsergebnisse aus den Treppenstufenversuchen nach HÜCK mit Anschnittsdefinition nach HÜCK für $d' = 1,036$ eingetragen.

4 Vergleich der Simulationsergebnisse

Für die Abschätzung der Quantile des Schätzers m' aus dem verallgemeinerten Standardfehler bei Horizontversuchen ergibt sich ein ähnliches Bild. Wie in der Abbildung 4.14 dargestellt, können die verallgemeinerten Quantile recht gut nachgebildet werden, wenn der Standardfehler mit dem entsprechenden Quantil der t-Verteilung multipliziert wird.⁵ Unberührt davon bleibt natürlich die Frage, ob Schätzungen des Erwartungswertes mit einer Konfidenzintervallbreite in der Größenordnung der Standardabweichung der Grundgesamtheit ein für den Experimentator zufriedenstellendes Ergebnis darstellen. Um ein möglichst kleines Konfidenzintervall zu erreichen, müssen möglichst viele Proben geprüft werden. Für Versuchsreihen mit mehr als 30 Einzelversuchen ist der Unterschied zwischen den Quantilen der t-Verteilung und der Normalverteilung vernachlässigbar klein.

⁵Eigentlich werden Quantile der t-Verteilung dazu genutzt, den Vertrauensbereich der Erwartungswertschätzung bei unbekannter Standardabweichung einer Zufallsgröße mit metrischer Merkmalsausprägung zu beschreiben. Ohne das dies weiter begründbar ist, zeigen die Ergebnisse, dass die ermittelten verallgemeinerten Quantile aus den oben berechneten verallgemeinerten Standardfehler recht gut abgeschätzt werden.

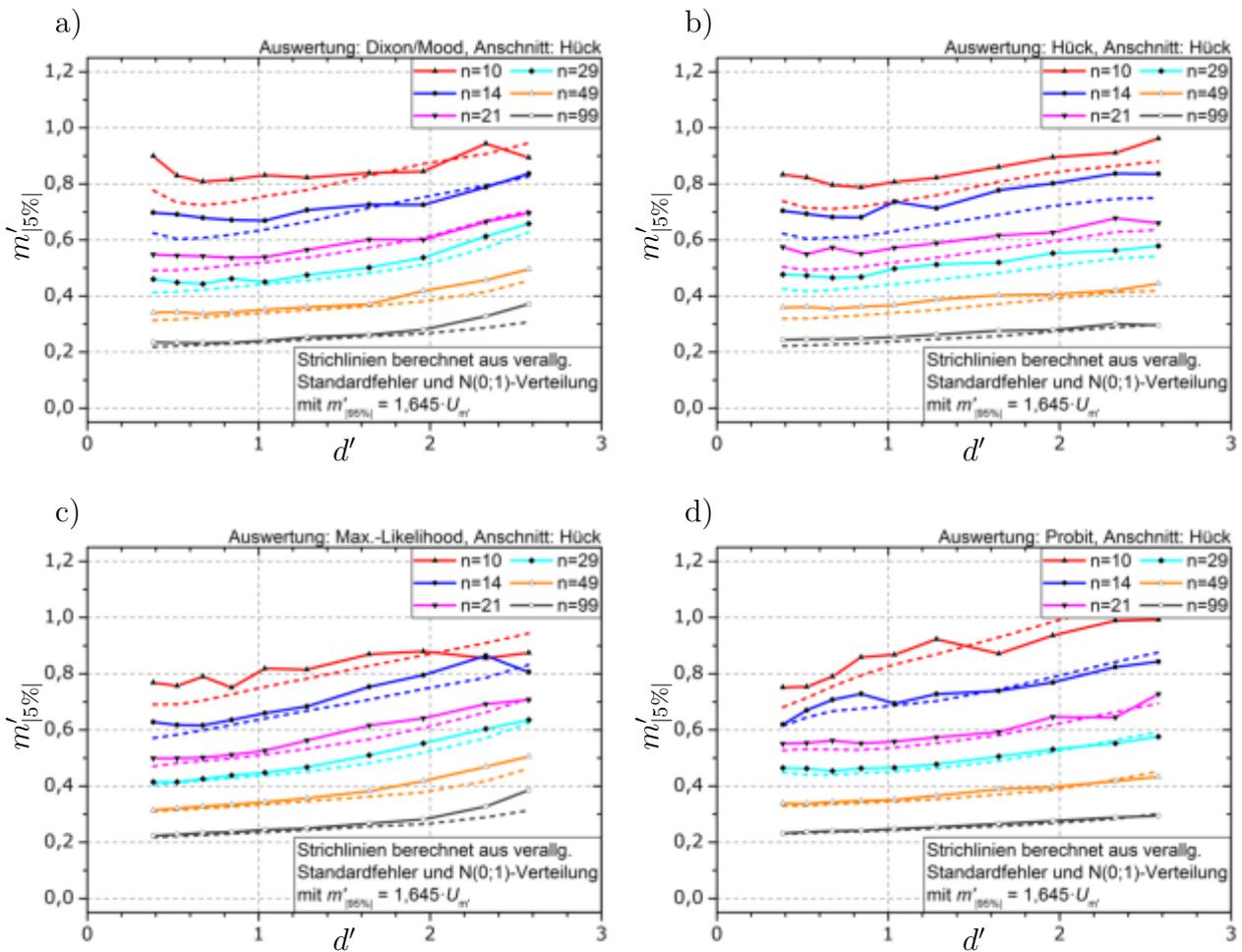


Abb. 4.12: Vergleich rechnerischer und verallgemeinerter 5 %-Quantile des Schätzers m' für Treppenstufenversuche für verschiedene Auswerteverfahren, jeweils mit Anschnittsdefinition nach HÜCK. Für die Auswertemethode nach DIXON/MOOD in a), HÜCK bzw. DEUBELBEISS in b), MAX.-LIKELIHOOD-METHODE in c) sowie PROBIT-METHODE in d).

4.2 Schätzung des Erwartungswertes bei bekannter Standardabweichung

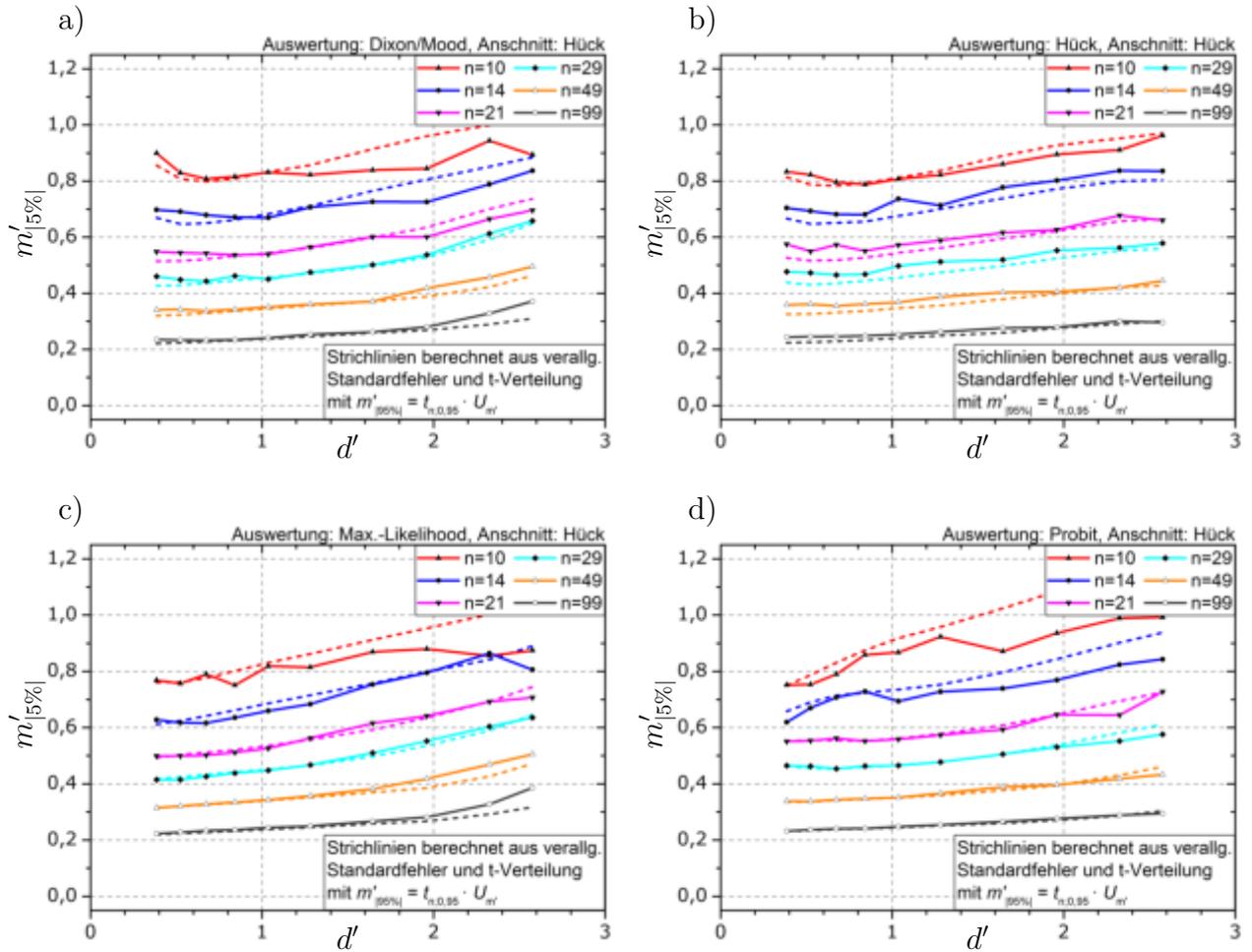


Abb. 4.13: Vergleich rechnerischer und verallgemeinerter 5%-Quantile des Schätzers m' für Treppenstufenversuche für verschiedene Auswerteverfahren, jeweils mit Anschnittsdefinition nach HÜCK. Für die Auswertemethode nach DIXON/MOOD in a), HÜCK bzw. DEUBELBEISS in b), MAX.-LIKELIHOOD-METHODE in c) sowie PROBIT-METHODE in d).

4 Vergleich der Simulationsergebnisse

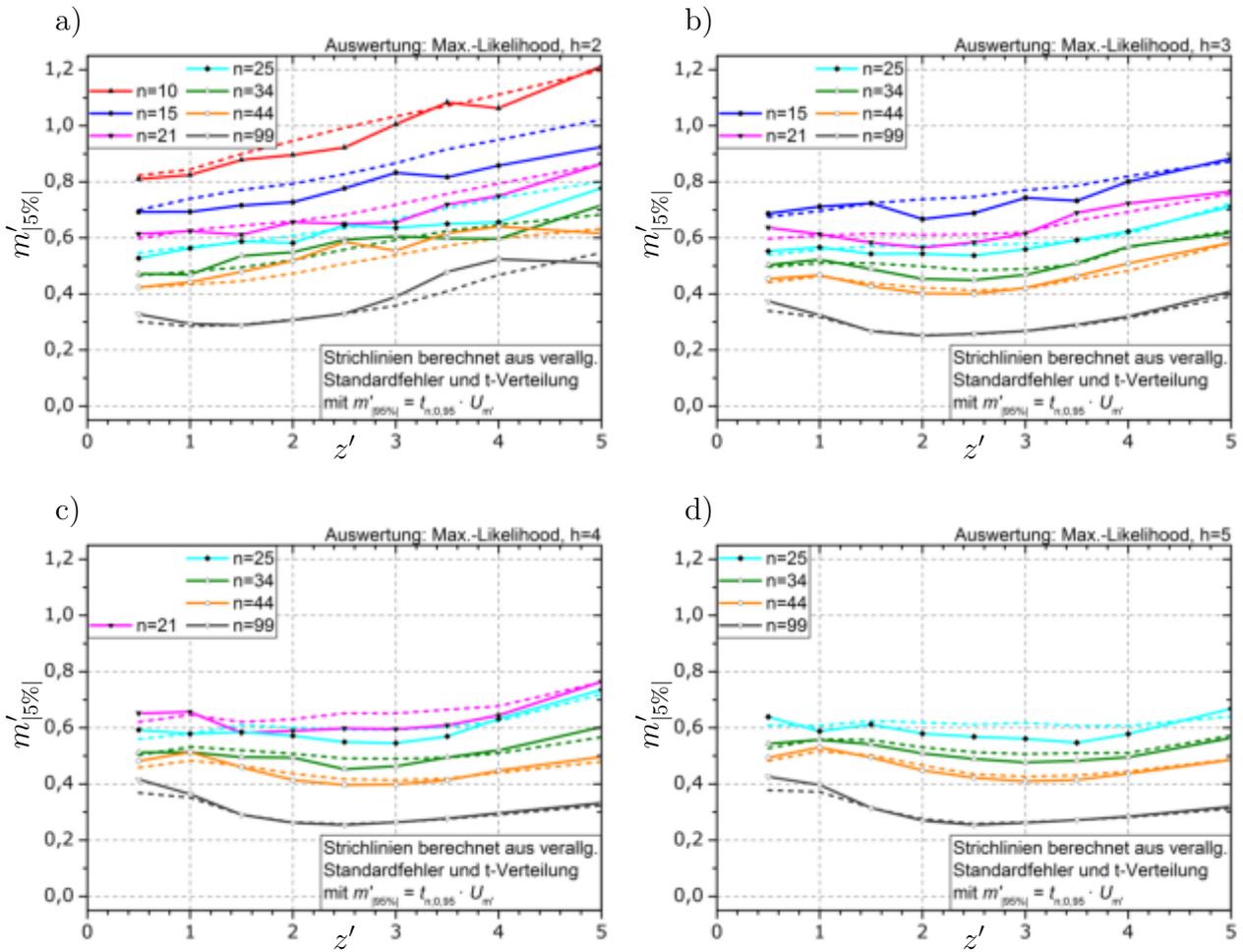


Abb. 4.14: Vergleich rechnerischer und verallgemeinerter 5%-Quantile des Schätzers m' für nach der MAX-LIKELIHOOD-METHODE ausgewertete Horizontversuche. Für Versuche auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

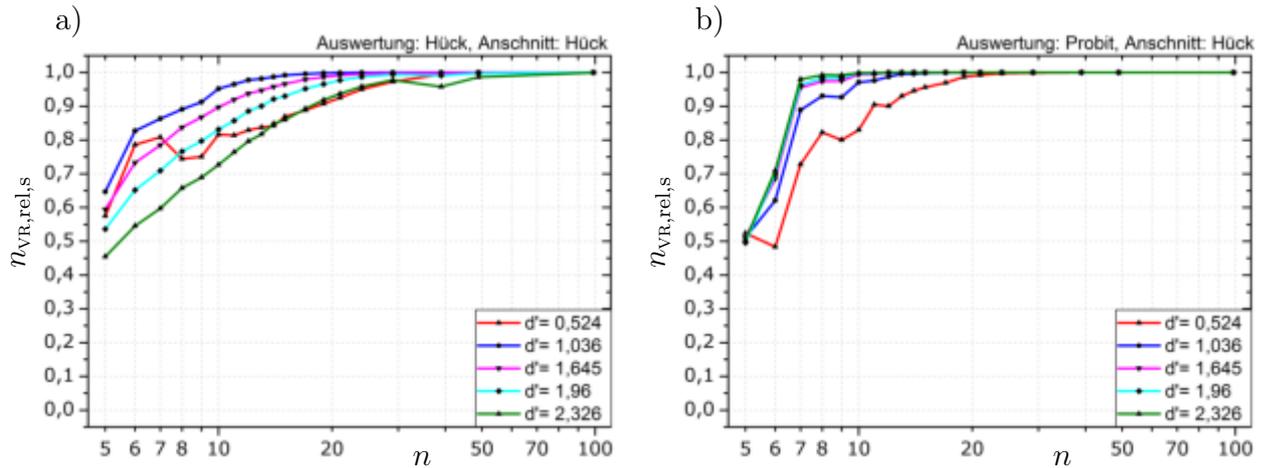


Abb. 4.15: Anteil auswertbarer Versuchsreihen für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, bei verschiedenen Stufenabständen und mit Auswertungen nach HÜCK in a) und PROBIT-METHODE in b).

4.3 Schätzung der Standardabweichung

4.3.1 Anteil auswertbarer Versuchsreihen

Die Auswertungen der Treppenstufenversuche nach DIXON/MOOD⁶, DEUBELBEISS oder der MAX.-LIKELIHOOD-METHODE liefern für jede Versuchsreihe stets einen Schätzwert für die Standardabweichungen. Bei den Auswertungen nach HÜCK oder der PROBIT-METHODE ist dies nicht der Fall. Die relative Anzahl der auswertbaren Versuchsreihen ist für beide Verfahren in Abbildung 4.15 dargestellt. Werden Horizontversuche mit der PROBIT-METHODE oder der MAX.-LIKELIHOOD-METHODE ausgewertet, liefern ebenfalls einige Versuchsreihen nicht sinnvolle (negative) Schätzwerte für die Standardabweichung. Im Vergleich mit den Treppenstufenversuchen bei einem Stufenabstand von $d' \approx 1$ besteht vor allem bei kleinen Horizontspreizungen ein deutliches Risiko, trotz $n \geq 20$ Einzelversuchen, eine nicht auswertbare Versuchsreihe zu erhalten, vgl. Abbildung 4.16. Berücksichtigt werden muss dabei allerdings, dass sich der hier dargestellte Anteil auswertbarer Versuchsreihen immer auf Versuchsreihen bezieht, die bei unbegrenzt zur Verfügung stehender Anzahl an Proben die erforderliche Anzahl an gültigen Einzelversuchen erreicht haben. Wird die Anzahl der zur Verfügung stehenden Proben von vornherein begrenzt, kann der Anteil nicht auswertbarer Versuchsreihen höher sein (siehe Abschnitt 4.1).

4.3.2 Mittlerer zufälliger und systematischer Schätzfehler

Das Verhältnis des Schätzers s zur tatsächlichen Standardabweichung σ_G wird als logarithmisch normalverteilt vorausgesetzt, d. h. der Logarithmus des Verhältnisses wird als normalverteilt angenommen und als solches ausgewertet. Als systematischer Schätzfehler wird dann der Erwartungswert und als zufälliger Schätzfehler die Standardabweichung des

⁶Versuchsreihen mit $k_{DM} < 0,3$ werden in den Simulationen mit der MAX.-LIKELIHOOD-METHODE ausgewertet, vgl. Abschnitt 2.2.2.1 S. 46.

4 Vergleich der Simulationsergebnisse

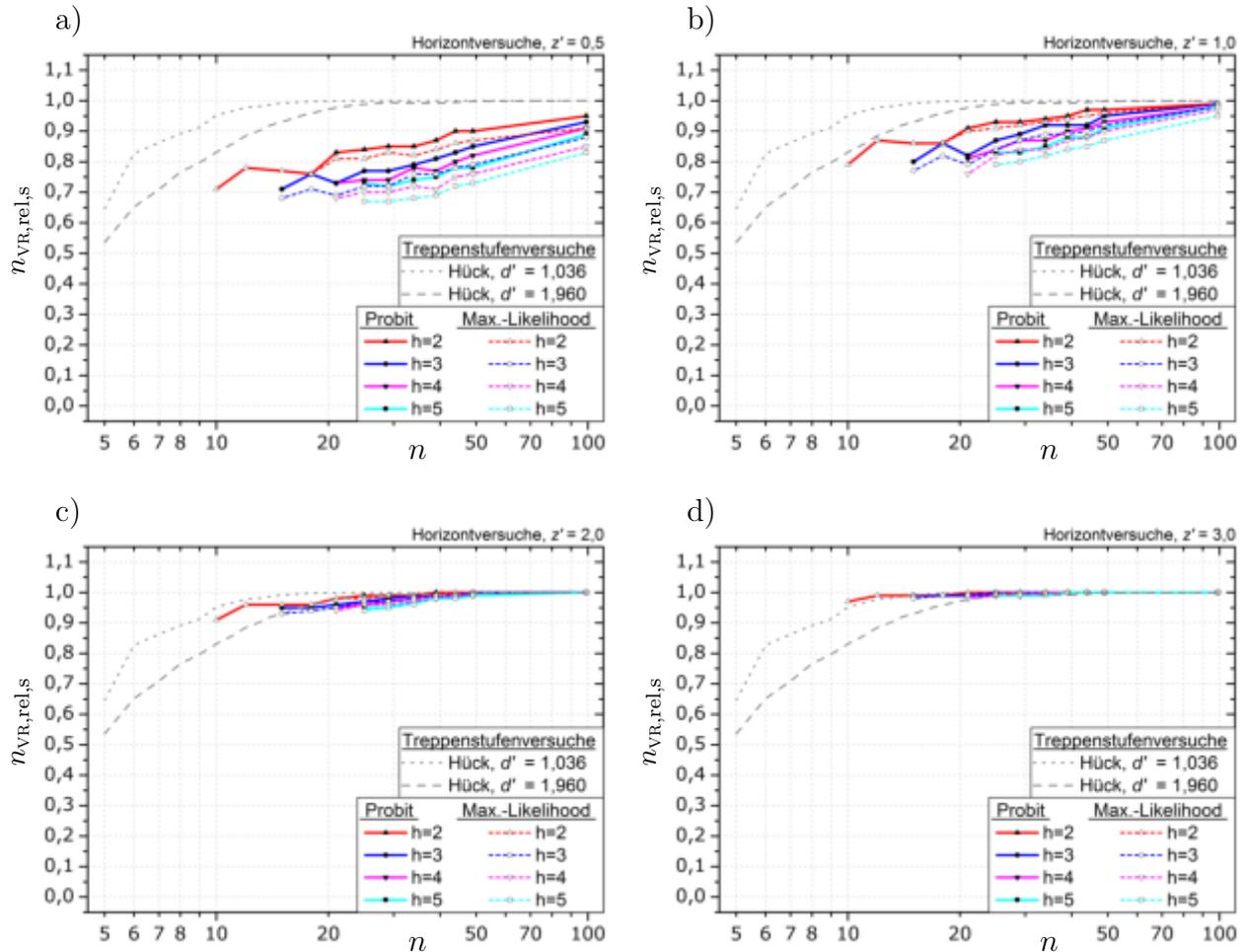


Abb. 4.16: Relative Anzahl auswertbarer Versuchsreihen für Horizontversuche bei verschiedenen Horizontspreizungen z' sowie zum Vergleich die Ergebnisse aus Treppenstufenversuchen nach HÜCK für $d' = 1,036$ und $d' = 1,96$. Ergebnisse für Horizontspreizungen von $z' = 0,5$ in a), $z' = 1,0$ in b), $z' = 2,0$ in c) sowie $z' = 3,0$ in d).

logarithmierten Verhältnisses verstanden. Dabei muss beachtet werden, dass ein etwaiger Erwartungswert von $\mathbf{E}[\lg(\frac{s}{\sigma_G})] = \mathbf{E}[s'_{lg}] = 0$ nicht gleichbedeutend mit einer erwartungstreuen Schätzung der Standardabweichung ist, da sich der tatsächliche Erwartungswert einer logarithmisch normalverteilten Zufallsgröße nach Gleichung 2.19 auf S. 32 bestimmt. Ein Erwartungswert von $\mathbf{E}[s'_{lg}] = 0$ bedeutet vielmehr eine mediantreue Schätzung der Standardabweichung, bei der gleich viele Über- wie Unterschätzungen der tatsächlichen Standardabweichung auftreten.

Die im Abschnitt 3.4.3.2 ab S. 77 aus den Simulationsergebnissen berechneten mittleren Verzerrungen des Schätzers s'_{lg} sind in Abbildung 4.17 für die verschiedenen Auswertungen der **Treppenstufenversuche** dargestellt. Demnach zeigen die Auswertemethoden nach DIXON/MOOD und MAX.-LIKELIHOOD-METHODE stets negative Erwartungswerte. Sie schätzen also in der Mehrzahl der simulierten Versuchsreihen zu geringe Standardabweichungen. Mit zunehmender Versuchsanzahl verbessern sich die Schätzungen. Beide Methoden zeigen ein, bei kleinen Stufenabständen deutlich stärker als bei großen Stufenabständen ausgeprägtes, asymptotisch mediantreues Verhalten. Die Schätzungen der Standardabweichungen mittels PROBIT-METHODE oder nach HÜCK zeigen im Sinne einer mediantreuen Schätzung die besten

Ergebnisse. Durch die von HÜCK eingeführten Korrekturfunktionen (siehe Abschnitt 2.2.2.2, S. 48) ist es demnach möglich, über einen weiten Bereich der Stufenabstände mediantreue Schätzer zu erhalten. Lediglich bei Versuchsanzahlen $n > 50$ scheint die Korrekturfunktion selbst korrekturbedürftig zu sein. Aus Abbildung 4.17 wird ebenfalls deutlich, dass der untersuchte Einfluss der Anschnittsdefinition auf den Erwartungswert vernachlässigbar ist.

Die Auswertemethode nach DEUBELBEISS verdient besondere Aufmerksamkeit. Wie im Abschnitt 2.2.2.3, S. 50 beschrieben, ist es nicht zu erwarten, mit dieser Auswertemethode erwartungstreue Schätzwerte der Standardabweichung zu ermitteln. Die Größe der systematischen Verschätzung ist direkt und in viel stärkerem Maß als bei allen anderen Auswerteverfahren vom gewählten Stufenabstand abhängig, wie aus Abbildungen 4.17 zu entnehmen ist. Gleichwohl ist es bei einem Stufenabstand von $1,2 \leq d' \leq 1,4$ möglich, nahezu mediantreue Schätzungen zu erhalten, vgl. Abbildung A.32 a), S. 209. Um aber einen solchen Stufenabstand gezielt einzustellen, müsste die Standardabweichung schon bekannt sein und muss nicht mehr ermittelt werden. Wird viel Wert auf eine möglichst mediantreue Schätzung der Standardabweichung der Grundgesamtheit gelegt, was dann als nötig erscheint, wenn Ergebnisse verschiedener Versuchsreihen berücksichtigt werden, zum Beispiel in Ringversuchen oder bei der Erstellung von Werkstoffkatalogen, ist die Auswertemethode nach DEUBELBEISS ungeeignet.

Für die **Horizontversuche** zeigen die Schätzer s'_{lg} bei der Auswertung mittels MAX.-LIKELIHOOD-METHODE, wie schon bei den Treppenstufenauswertungen, negative mittlere Erwartungswerte, vgl. Abbildung 4.18. Es werden also die Mehrzahl der Schätzungen der Standardabweichung zu gering ausfallen. Diese Abweichungen verringern sich mit zunehmender Versuchsanzahl, vergrößern sich jedoch mit zunehmender Horizontspreizung. Für eine möglichst mediantreue Schätzung der Standardabweichung ist es prinzipiell zu empfehlen, Versuchsreihen mit möglichst vielen Horizonten ($h > 3$) zu belegen. Werden die Horizontversuche mit der PROBIT-METHODE ausgewertet, kommt es bei kleinen Horizontspreizungen für eine Mehrzahl der Versuchsreihen zu einer Unterschätzung und bei großen Stufenabständen zu einer systematischen Überschätzung der Standardabweichung. Wird die Horizontspreizung günstig mit $z \approx 2$ gewählt, treten unabhängig von der Anzahl der belegten Horizonte, nahezu mediantreue Schätzungen auf. In jedem Fall liegt der Median der Standardabweichungsschätzungen mit der PROBIT-METHODE näher am Wert der wahren Standardabweichung als bei Auswertungen mit der MAX.-LIKELIHOOD-METHODE. Das erlaubt oft eine im Mittel konservativere Schätzung der Standardabweichung.

Die Betrachtung des mittleren Standardfehlers $C_{s',lg}$ der **Treppenstufenversuche** in Abbildung 4.19 zeigt, dass bei kleinen Stufenabständen die Auswertungen nach HÜCK und mit der PROBIT-METHODE eher ungeeignet sind und die Auswertungen nach DIXON/MOOD bzw. mit der MAX.-LIKELIHOOD-METHODE bevorzugt werden sollten. Bei Stufenabständen von $d' \approx 1$ oder größer zeigen hingegen die Auswertungen nach HÜCK und mit PROBIT-METHODE geringere zufällige Schätzfehler als die Auswertungen nach DIXON/MOOD oder der MAX.-LIKELIHOOD-METHODE. Eine gesonderte Betrachtung verdient wieder die Auswertung nach DEUBELBEISS, da hiermit durchgängig die kleinsten mittleren Standardfehler zu erreichen sind.

Für **Horizontversuche** ist der mittlere Standardfehler des Schätzers s'_{lg} in Abbildung 4.20 dargestellt. Zum Vergleich ist auch der mittlere Standardfehler für Treppenstufenversuche und die Auswertung nach der MAX.-LIKELIHOOD-METHODE bei $d' \approx 1$ eingezeichnet. Demnach

4 Vergleich der Simulationsergebnisse

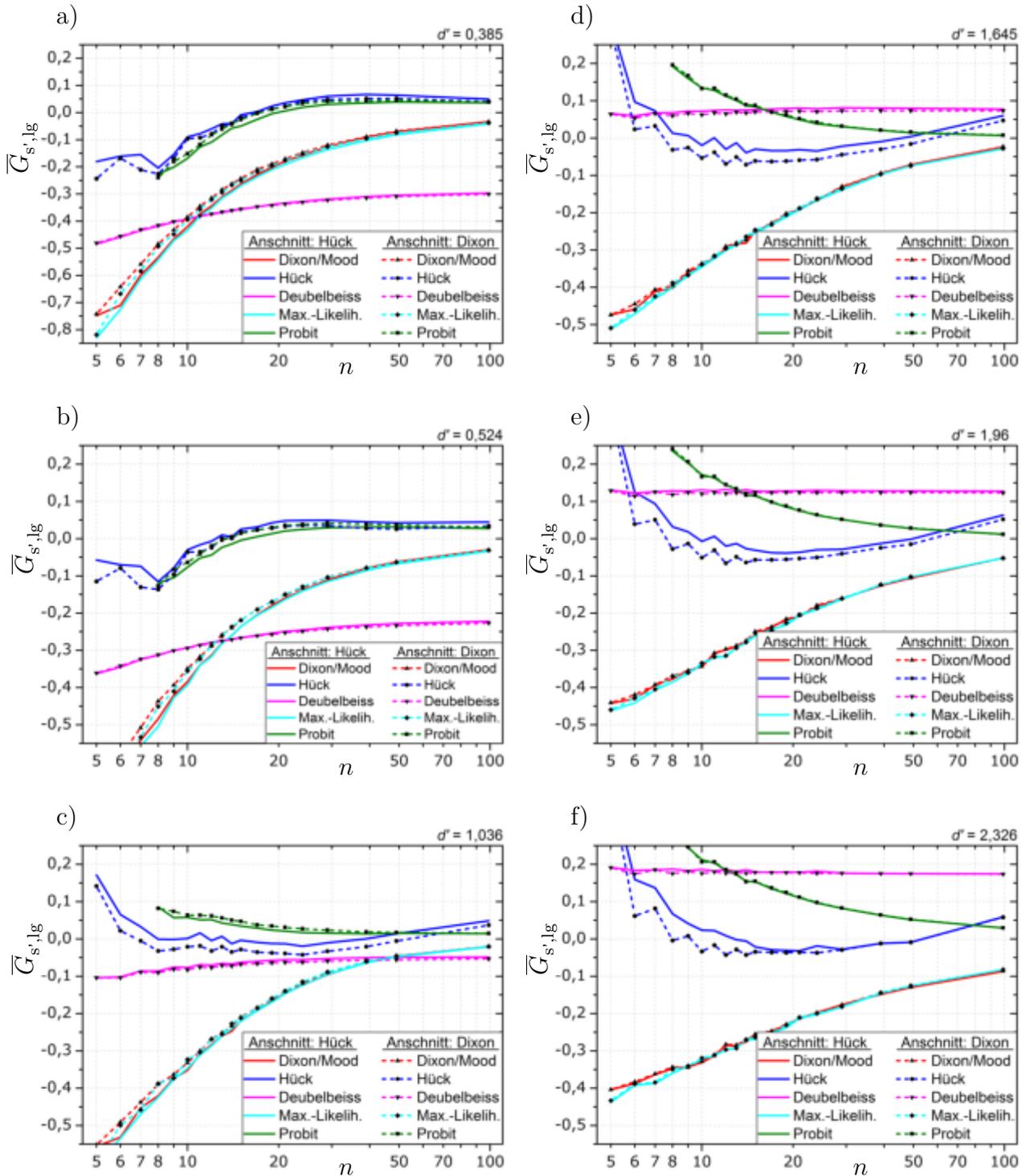


Abb. 4.17: Mittlerer systematischer Fehler des Schätzers s'_{lg} für verschiedene Treppenstufenauswertungen. Simulationsergebnisse bei einem vorgegebenen Stufenabstand von $d' = 0,385$ in a), $d' = 0,524$ in b), $d' = 1,036$ in c), $d' = 1,645$ in d), $d' = 1,960$ in e) sowie $d' = 2,326$ in f).

führt eine zu klein gewählte Horizontspreizung zu einem größeren Standardfehler als bei Treppenstufenversuchen. Erst bei einer Horizontspreizung von $z' \geq 2$ ist es möglich, Schätzer mit geringeren oder gleichen Streuungen zu erhalten. Für die Auswertungen nach der MAX.-LIKELIHOOD-METHODE ist dies nur möglich, wenn die Versuche auf möglichst viele Horizonte

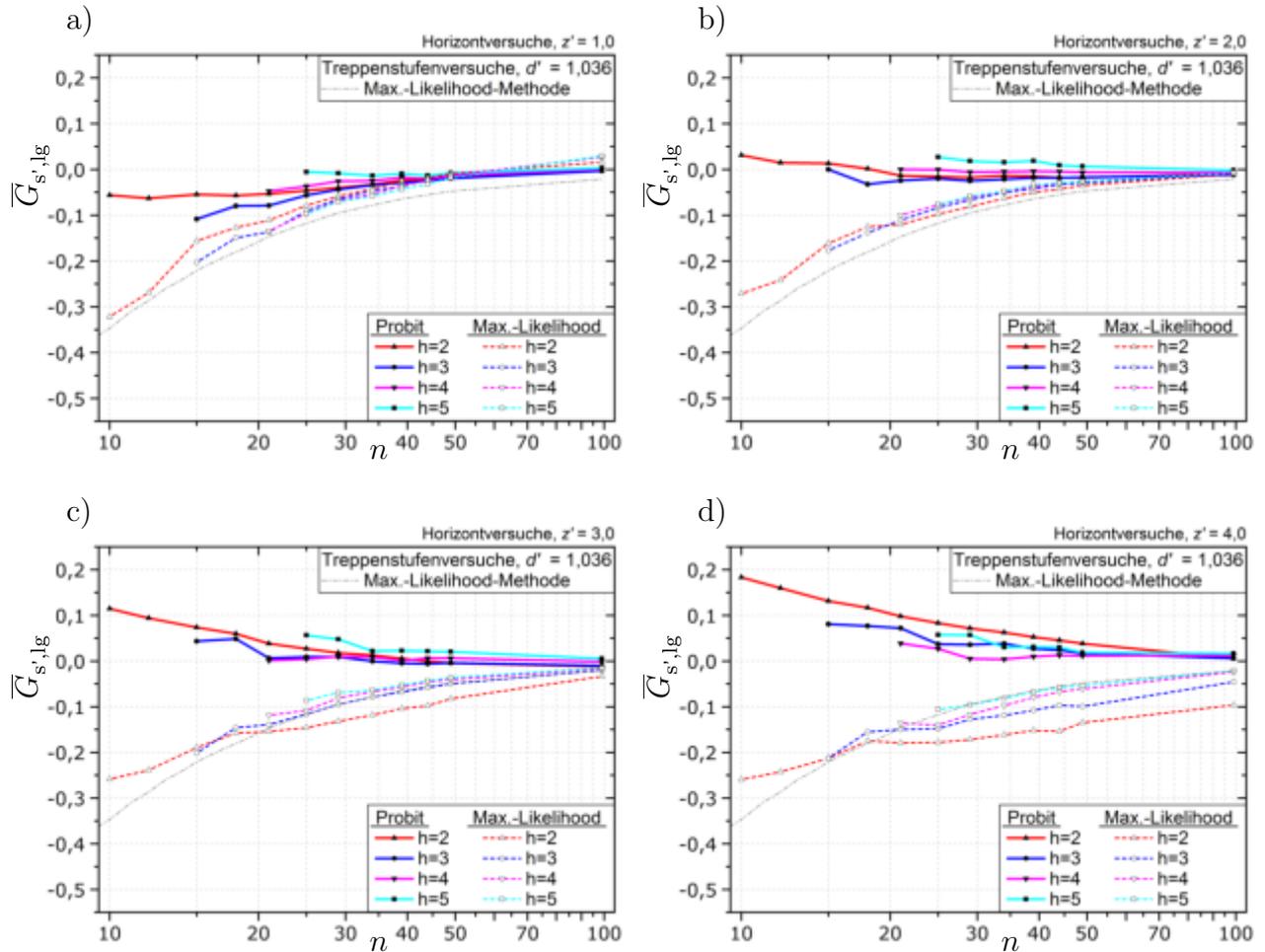


Abb. 4.18: Mittlerer systematischer Fehler des Schätzers s'_{lg} für Horizontversuche bei verschiedenen Horizontspreizungen von $z' = 1$ in a), $z' = 2$ in b), $z' = 3$ in c) sowie $z' = 4$ in d). Zum Vergleich jeweils die Ergebnisse aus Treppenstufenversuchen und der Auswertung mit der MAX.-LIKELIHOOD-METHODE für $d' = 1,036$.

aufgeteilt werden. Im Gegensatz dazu werden bei der PROBIT-METHODE dann bessere Ergebnisse erzielt, wenn die Einzelversuche auf möglichst wenige Horizonte verteilt werden. Auffällig ist hierbei der geringe Standardfehler der PROBIT-METHODE bei Auswertungen von Versuchsreihen mit zwei Horizonten und großen Horizontspreizungen. Der Grund hierfür ist in der Behandlung von Horizonten mit nur Brüchen oder nur Durchläufern zu suchen, die bei Versuchsreihen mit großen Horizontspreizungen vermehrt auftreten. Diesen Horizonten wird bei der Auswertung eine Positionswahrscheinlichkeit in ausschließlicher Abhängigkeit der Anzahl der Versuche auf dem jeweiligen Horizont zugeordnet (siehe Ungleichungen 2.29 und 2.30 auf S. 40). Bei gleicher Lage der Horizonte und gleicher Anzahl an Versuchen je Horizont ergeben sich dadurch meist zu große, aber identische Schätzungen der Standardabweichungen.

4 Vergleich der Simulationsergebnisse

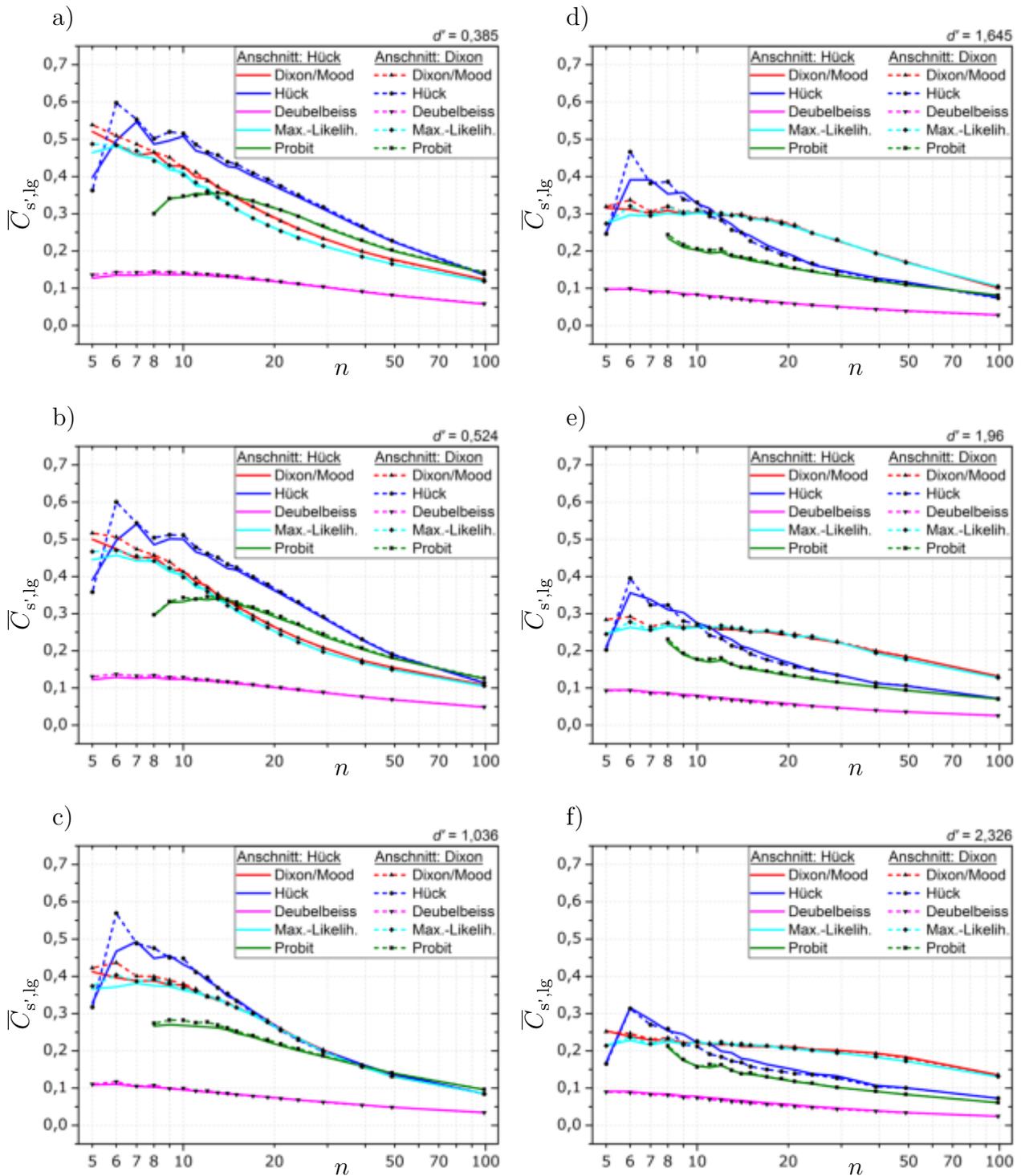


Abb. 4.19: Mittlerer zufälliger Fehler des Schätzers s'_{lg} für Treppenstufenversuche mit verschiedenen Anschnittsdefinitionen. Simulationsergebnisse bei einem vorgegebenen Stufenabstand von $d' = 0,385$ in a), $d' = 0,524$ in b), $d' = 1,036$ in c), $d' = 1,645$ in d), $d' = 1,96$ in e) sowie $d' = 2,326$ in f).

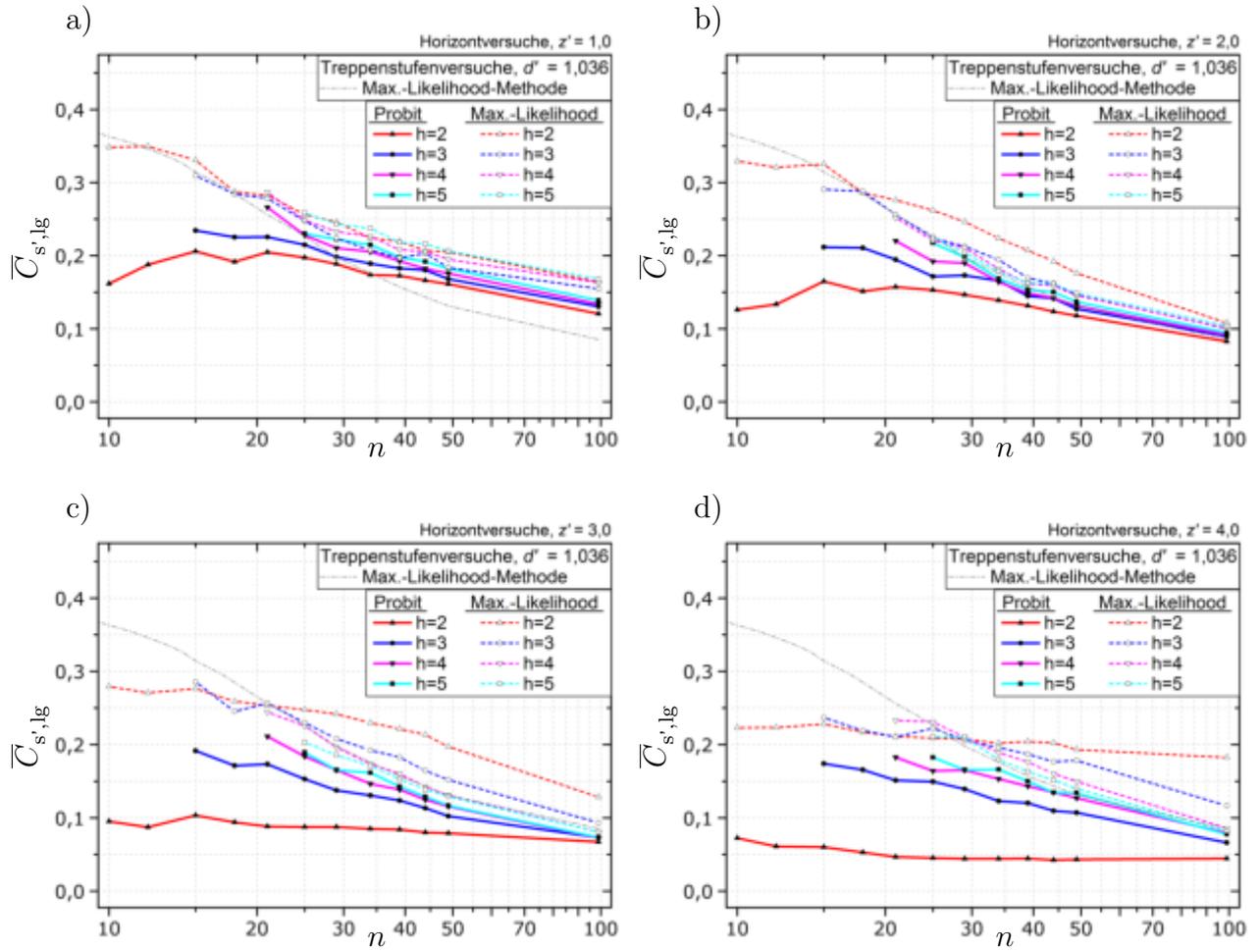


Abb. 4.20: Mittlerer Standardfehler des Schätzers s'_{lg} für Horizontversuche bei verschiedenen Horizontspreizungen von $z' = 1$ in a), $z' = 2$ in b), $z' = 3$ in c) sowie $z' = 4$ in d). Zum Vergleich jeweils die Ergebnisse aus Treppenstufenversuchen und der Auswertung mit der MAX.-LIKELIHOOD-METHODE für $d' = 1,036$.

4.3.3 RMS-Fehler

Wie im vorhergehenden Abschnitt beschrieben, ergeben sich für die verschiedenen Methoden der Versuchsführung und -auswertung je nach Versuchsanzahl und Stufenabstand bzw. Horizontspreizung und je nachdem, ob der systematische oder der zufällige Anteil des Schätzfehlers betrachtet wird, unterschiedliche Schlüsse bzgl. der Treffsicherheit. Eine übersichtlichere Darstellung der Ergebnisse kann durch die Betrachtung des mittleren quadratischen Schätzfehlers bzw. durch dessen als mittleren RMS-Fehler bezeichnete Quadratwurzel, vgl. Gleichung 2.5, erreicht werden, da hier sowohl der systematische als auch der zufällige Fehler zusammen betrachtet werden. Damit geht natürlich die Information über eine systematische mögliche Verschätzung verloren. Werden die Ergebnisse einer Schätzung mit kleinem RMS-Fehler vieler gleichartiger Versuchsreihen gemittelt, kann das Ergebnis unter Umständen eine schlechtere Schätzung als das gemittelte Ergebnis einer Schätzmethode mit größerem RMS-Fehler ergeben. In den meisten Fällen erfolgt die Bestimmung der Festigkeit eines Werkstoffs oder Bauteils aber einmalig, d. h. es liegen keine weiteren Ergebnisse vor, mit denen durch Mittlung der zufällige Einfluss reduziert werden kann. Es ist dann entscheidend, mit dieser einen Schätzung möglichst nah an den wahren Wert heran zu kommen. Dementsprechend ist ein verzerrter Schätzer einem unverzerrten (erwartungstreuen) Schätzer dann vorzuziehen, wenn er einen kleineren RMS Fehler besitzt, also im Mittel trotz systematischen Fehlers näher am wahren Wert liegt.

Die bei der Auswertung von **Treppenstufenversuchen** zu erwartenden mittleren RMS-Fehler sind in Abbildung 4.21 dargestellt.⁷ Gut zu erkennen ist hier, dass die Auswertungen nach der MAX.-LIKELIHOOD-METHODE und nach DIXON/MOOD zu gleichen Schätzfehlern führen. Beide Auswertemethoden schätzen die Standardabweichung nur dann genauer als die anderen Verfahren, wenn kleine Stufenabstände und Versuchsanzahlen $n > 20$ vorliegen. In allen anderen Fällen sind die Auswertemethode nach HÜCK oder die PROBIT-METHODE zu bevorzugen.

Wie schon im vorhergehenden Abschnitt beschrieben, verdient die Auswertung nach DEUBELBEISS eine gesonderte Betrachtung. Bei kleinen sowie sehr großen Stufenabständen führt die nicht mediantreue Schätzung der Standardabweichung zu großen RMS-Fehlern. Auch eine Erhöhung der Probenanzahl führt dann nicht zu einer wesentlichen Verringerung des Schätzfehlers. Gleichwohl ist es für eine einzelne Versuchsreihe mit etwa $n \leq 20$ gültigen Einzelversuchen in einem weiten Bereich des Stufenabstandes von etwa $0,5 \leq d' \leq 2$ möglich, dem wahren Wert der Standardabweichung näher zu kommen als mit allen anderen untersuchten Treppenstufenauswertemethoden. Für Treppenstufenversuche mit $n < 10$ gültigen Einzelversuchen liegen die Schätzer sogar nahezu unabhängig vom tatsächlichen Stufenabstand immer näher am tatsächlichen Wert der Standardabweichung.

Wird der RMS-Fehler der Standardabweichungsschätzung als Entscheidungskriterium zur Wahl der Auswertemethode herangezogen, dann ist für Versuchsreihen mit geringem Stichprobenumfang die Methode nach DEUBELBEISS zur Auswertung von Treppenstufenversuchen zu wählen.

⁷Auf die Darstellung der Ergebnisse für die Anschnittsdefinition nach DIXON wird verzichtet, da diese Unterscheidung keinen Einfluss auf den mittleren RMS-Fehler des Schätzers s'_{lg} besitzt.

Die Ergebnisse der **Horizontversuche** sind in Abbildung 4.22 dargestellt. Zum Vergleich sind hier auch die Ergebnisse von mittels PROBIT-METHODE oder MAX.-LIKELIHOOD-METHODE ausgewerteten Treppenstufenversuchen eingetragen. Mit der MAX.-LIKELIHOOD-METHODE können demnach keine nennenswert besseren Ergebnisse als bei mit der gleichen Methode ausgewerteten Treppenstufenversuchen erreicht werden. Auch mit der PROBIT-METHODE lassen sich keine nennenswert kleineren RMS-Fehler erreichen als mit den zum Vergleich gewählten und ebenfalls mit der PROBIT-METHODE ausgewerteten Treppenstufenversuchen. Erst bei $n > 50$ Einzelversuchen und auch nur dann, wenn die Horizontspreizung gut gewählt wird ($z' \geq 3,0$), besitzen die Standardabweichungsschätzungen von mit der PROBIT-METHODE ausgewerteten Horizontversuchen kleinere RMS-Fehler. Auffallend ist der durchweg geringe RMS-Fehler des Abgrenzungsverfahrens. Wie schon bei der Betrachtung des Standardfehlers des Schätzers s'_{lg} ist dies auf die günstige Bewertung von reinen Bruch- bzw. Durchläuferhorizonten zurückzuführen.

4 Vergleich der Simulationsergebnisse

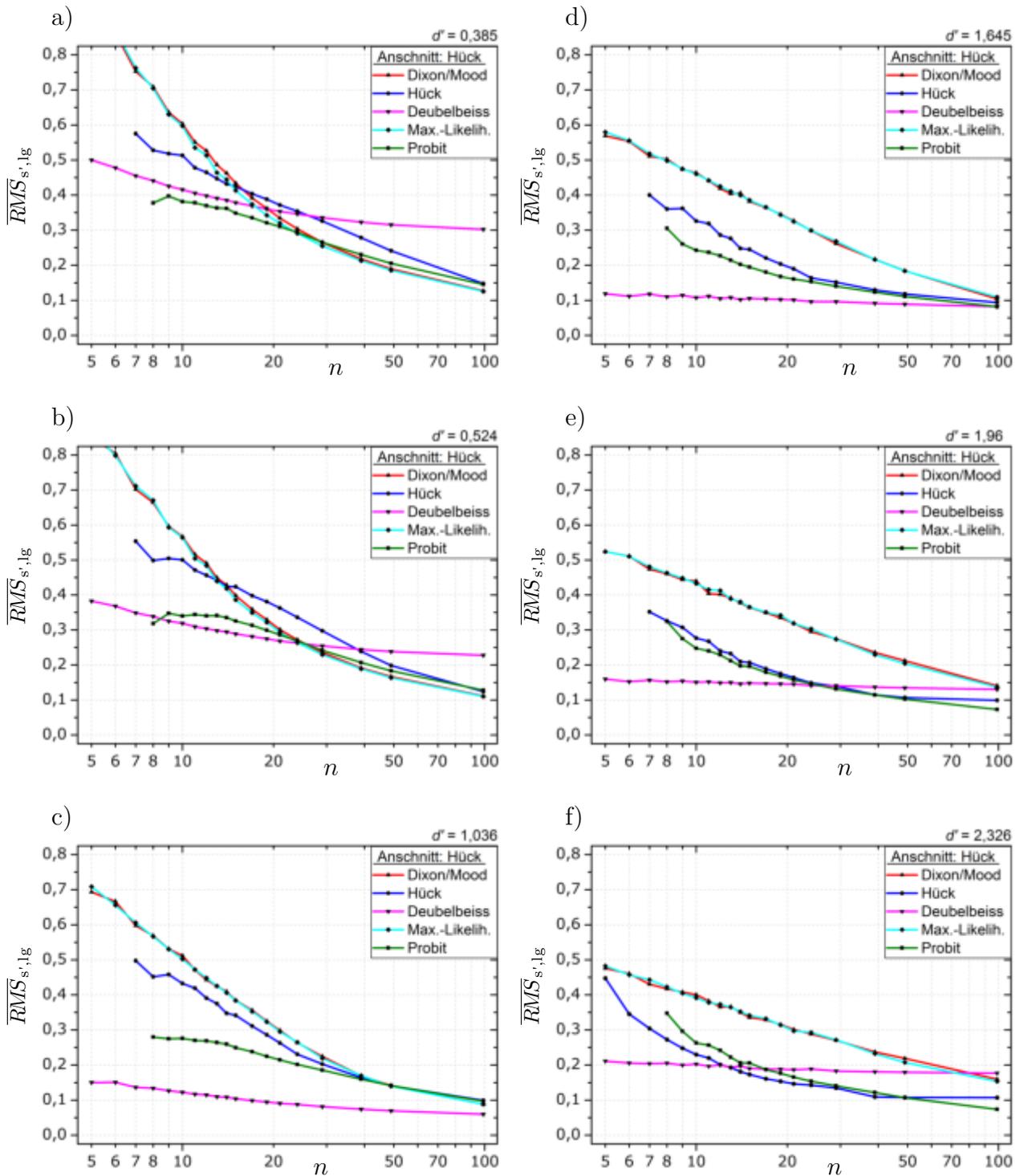


Abb. 4.21: Mittlerer RMS-Fehler des Schätzers s'_{lg} für verschiedene Treppenstufenauswertungen. Simulationsergebnisse bei einem vorgegebenen Stufenabstand von $d' = 0,385$ in a), $d' = 0,524$ in b), $d' = 1,036$ in c), $d' = 1,645$ in d), $d' = 1,96$ in e) sowie $d' = 2,326$ in f).

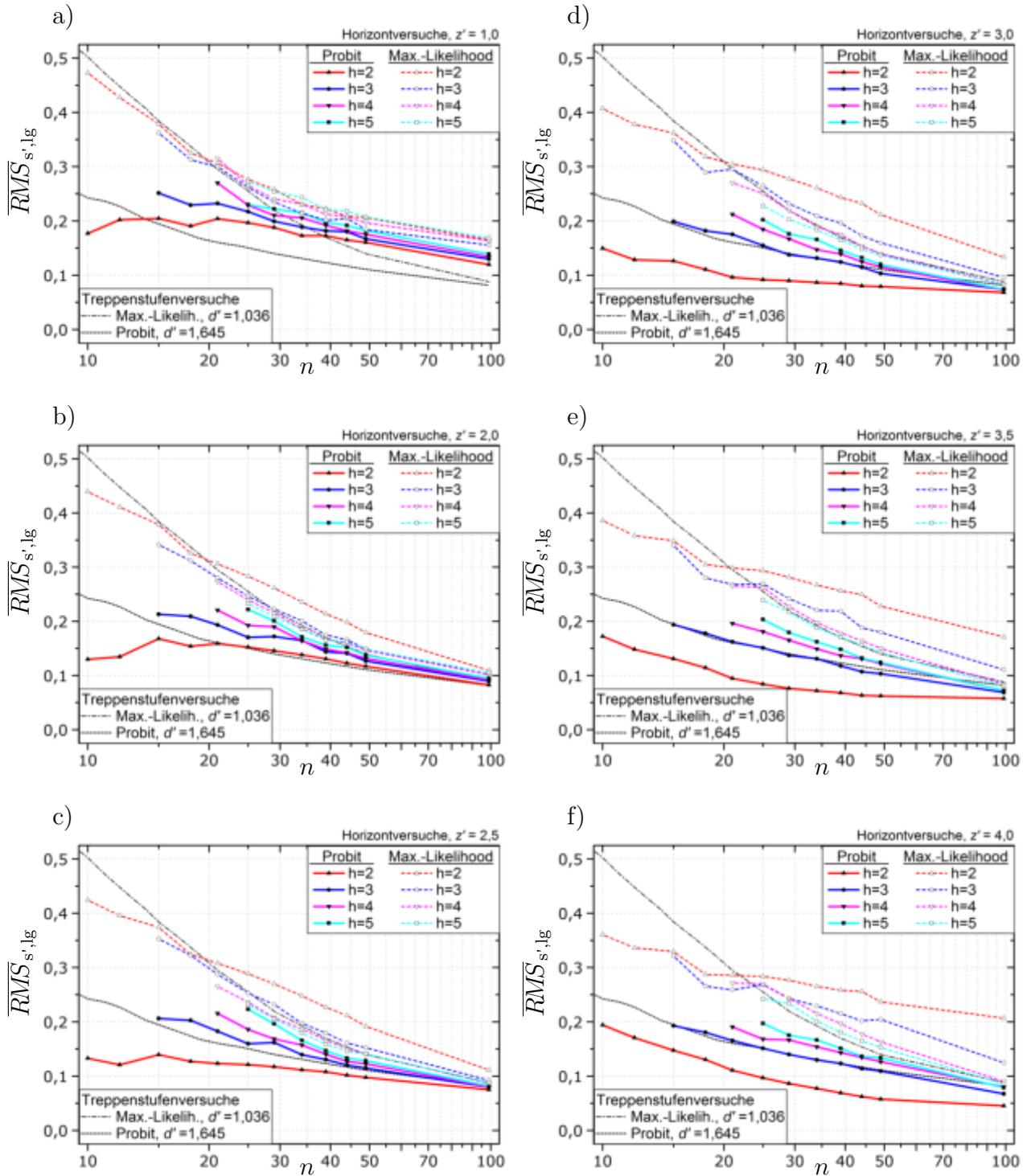


Abb. 4.22: Mittlerer Standardfehler des Schätzers s'_{lg} für Horizontversuche bei verschiedenen Horizontspreizungen von $z' = 1,0$ in a), $z' = 2,0$ in b), $z' = 2,5$ in c), $z' = 3,0$ in d), $z' = 3,5$ in e) sowie $z' = 4,0$ in f). Zum Vergleich die Ergebnisse aus Treppenstufenversuchen und der Auswertung mittels MAX.-LIKELIHOOD-METHODE für $d' = 1,036$ sowie mittels PROBIT-METHODE für $d' = 1,645$.

4.3.4 Empirischer Vertrauensbereich der geschätzten Standardabweichung

Neben der bisher durchgeführten Betrachtung zu den Punktschätzungen der Standardabweichung ist es für den Experimentator auch von Interesse, ein Vertrauensintervall für die geschätzte Standardabweichung angeben zu können. Von praktischer Bedeutung sind dabei die linksseitigen Quantile des bezogenen Schätzers der Standardabweichung, da diese die von nur wenigen Versuchsreihen erzeugten größten Unterschätzungen der Standardabweichung darstellen. Mit ihnen ist es möglich, für den wahren Wert der Standardabweichung eine obere Grenze anzugeben, die mit einer vorgegebenen Irrtumswahrscheinlichkeit vom tatsächlichen Wert der Standardabweichung nicht überschritten wird. Wie im Abschnitt 3.4.3.4 beschrieben, können für den Schätzer s'_{lg} mittlere 5 %-Quantile aus den Simulationsdaten abgeleitet werden, mit denen sich die Quantile $s_{5\%}$ der geschätzten Standardabweichung durch

$$s_{5\%} = \sigma_G \cdot 10^{\overline{s'}_{lg,5\%}}$$

abschätzen lassen.⁸ Für den wahren Wert der Standardabweichung der Grundgesamtheit lässt sich dann zu einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$ die obere Grenze

$$\sigma_G \leq s \cdot 10^{-\overline{s'}_{lg,5\%}}$$

des Vertrauensbereichs angeben. Die berechneten mittleren 5 %-Quantile des Schätzers s'_{lg} sind in den Abbildungen 4.23 für Treppenstufenversuche und 4.24 sowie 4.25 für Horizontversuche angegeben.

Zusätzlich werden dort den mittleren empirischen Quantilen die rechnerischen Quantile, jeweils ermittelt aus den mittleren systematischen und zufälligen Schätzfehlern und der Normalverteilungsannahme, gegenübergestellt. Für die ausgewerteten **Treppenstufenversuche** zeigt sich, dass die empirischen Quantile durch die rechnerischen Quantile nur bei Versuchszahlen $n > 20$ zufriedenstellend abgebildet werden können. Wird diese Bedingung eingehalten, dann liefern die rechnerischen Quantile in den allermeisten Fällen eine konservative Abschätzung der ermittelten empirischen Quantile.⁹ Bei den **Horizontversuchen** lassen sich jedoch sowohl bei den Auswertungen nach der PROBIT-METHODE als auch nach der MAX.-LIKELIHOOD-METHODE die ermittelten empirischen Quantile nicht zufriedenstellend aus den berechneten Verteilungsparametern bestimmen. Lediglich bei der PROBIT-METHODE

⁸Werden die Schätzungen also nur oft genug wiederholt, so werden nur 5 % der Schätzungen kleiner als $s_{5\%}$ ausfallen. Im Gegenzug liegt die tatsächliche Standardabweichung natürlich auch nur in 5 % aller Schätzungen um mehr als den Faktor $10^{-\overline{s'}_{lg,5\%}}$ von der geschätzten Standardabweichung entfernt.

⁹Unter der Bezeichnung konservativ wird hier eine zu große Unterschätzung der 5 %-Quantile der Standardabweichungsschätzung verstanden, wenn diese aus dem Standardfehler $\overline{C}_{s',lg}$ berechnet werden. Diese fällt bei den Auswertungen nach HÜCK, DIXON/MOOD und mit der MAX.-LIKELIHOOD-METHODE teils recht deutlich aus. Vermutlich liegt das an den bei größeren Stufenabständen häufiger auftretenden Versuchsreihen mit lediglich zwei oder drei Horizonten. Die Form der Versuchsreihen, also die Anzahl der Brüche und Durchläufer je Horizont, und damit die Ergebnisse der Standardabweichungsschätzung sind dann entsprechend oft gleich. Besonders kleine Werte, wie sie mit dem Modell der Normalverteilung vorhergesagt werden, sind in der Simulation dadurch schlichtweg nicht vorhanden. Für die Auswertungen nach DIXON/MOOD und mit der MAX.-LIKELIHOOD-METHODE ist die Konservativität nicht über alle Stufenabstände gegeben. Der Grund dafür ist dem Autor nicht bekannt. Bemerkenswert erscheint aber, das stets nahezu gleiche Verhalten der beiden Auswertungsverfahren. Ein fehlerhafte Rechnung bzw. Schätzung durch einen der beiden Berechnungsalgorithmen erscheint daher unwahrscheinlich.

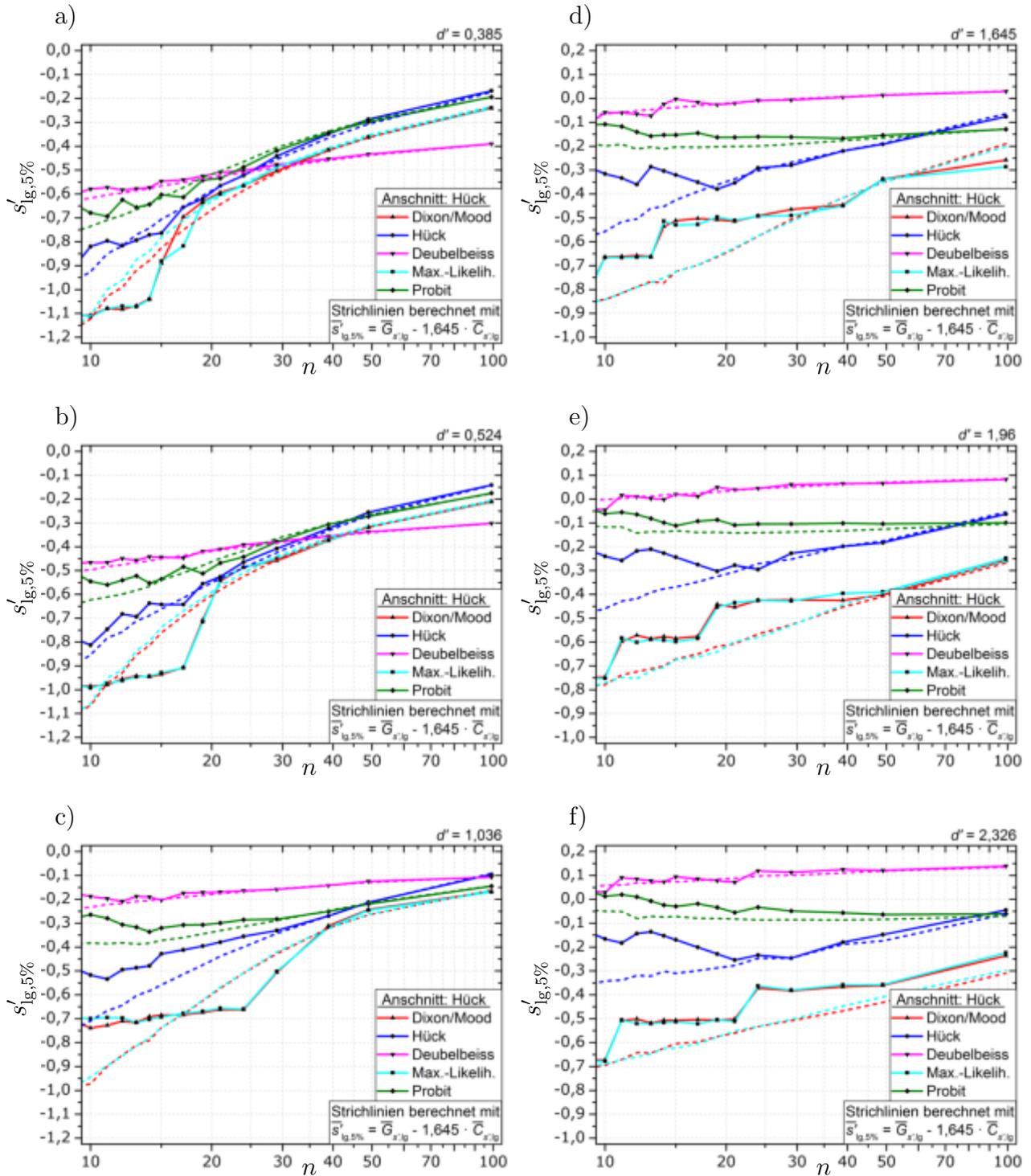


Abb. 4.23: Mittlere 5%-Quantile des Schätzers s'_{lg} für verschiedene Treppenstufenauswertungen bei verschiedenen Stufenabständen.

auf zwei Horizonten bzw. dem Abgrenzungsverfahren lassen sich zumindest gleichmäßige, stark konservative Abschätzungen für die obere Grenze eines Vertrauensbereichs ableiten, siehe Abbildung 4.24. Für alle anderen Horizontversuche sind Aussagen zur Größe des Vertrauensbereichs, zumindest mit der hier gewählten Art und Weise der Auswertung der Simulationsergebnisse, erst für $n = 50$ oder mehr Einzelversuchen möglich. Dieser Bereich

4 Vergleich der Simulationsergebnisse

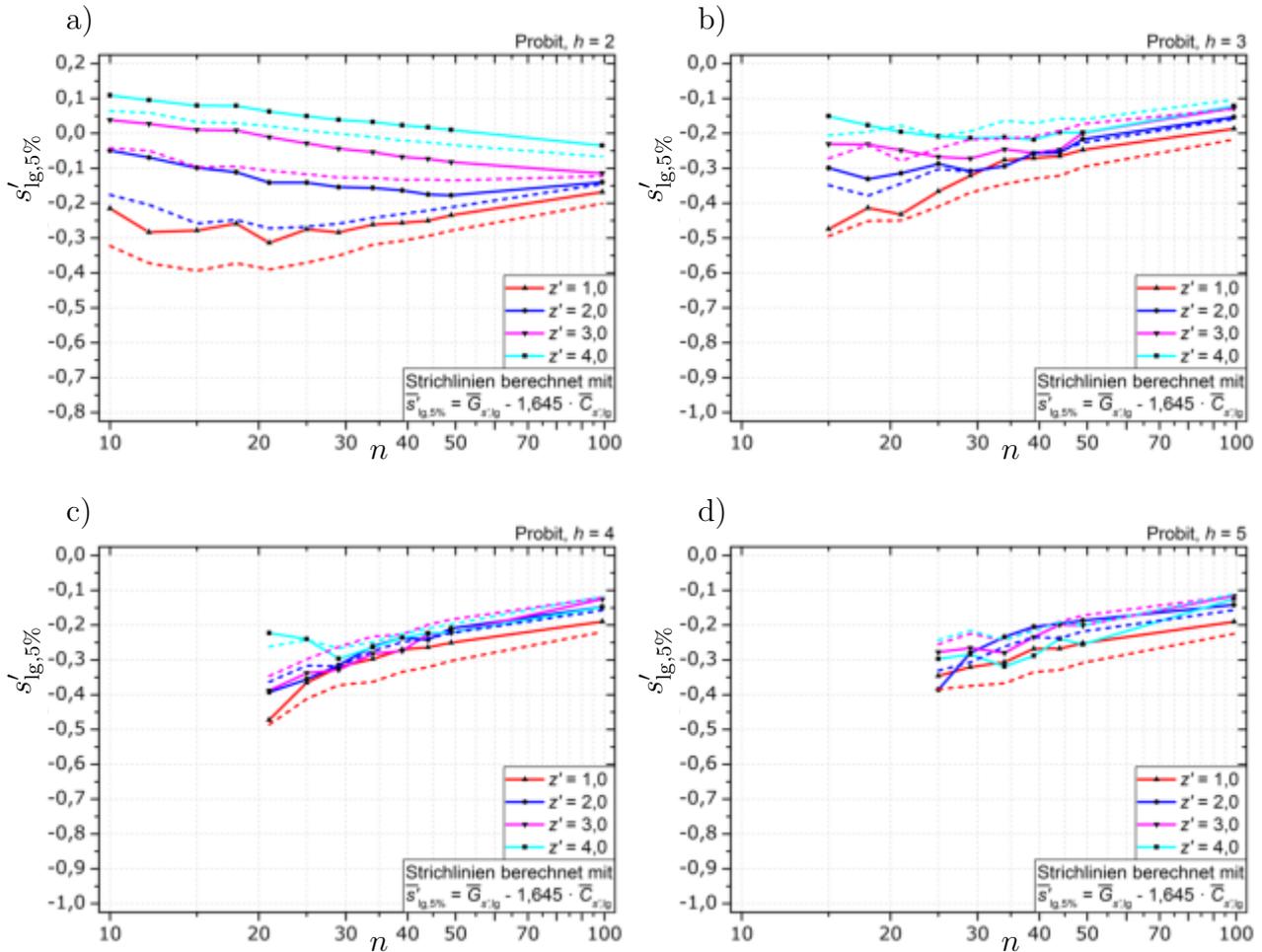


Abb. 4.24: Mittlere 5%-Quantile des Schätzers s'_{lg} für mittels PROBIT-METHODE ausgewertete Horizontversuche auf 2 Horzonten in a), 3 Horzonten in b), 4 Horzonten in c) sowie 5 Horzonten in d).

wird aber nicht ausreichend durch Simulationsergebnisse belegt und ist auch für die praktische Anwendung uninteressant.

Alle bisher dargestellten Ergebnisse geben die Genauigkeit der Standardabweichungsschätzung in Abhängigkeit der tatsächlichen Standardabweichung an, bei Treppenstufenversuchen durch den Parameter $d' = d/\sigma_G$ und bei Horizontversuchen durch den Parameter $z' = z/\sigma_G$. Es ist für ein einzelnes Versuchsergebnis nicht möglich, den tatsächlichen Stufenabstand bzw. die tatsächliche Horizontspreizung anzugeben, da dafür ja die tatsächliche Standardabweichung schon bekannt sein müsste. Wenn sich die Parameter d' und z' aber nicht mit Sicherheit angeben lassen, ist auch die Angabe der Vertrauensbereichsgrenzen unsicher. Im Sinne einer praktikablen Auswertung scheint es daher sinnvoll, die Größe der beiden Parameter jeweils nur in die Gruppen „klein“, „mittel“ und „groß“ einzuteilen und für die Bewertung jeweils Quantile mit $d' = 0,385 \approx 0,4$ bzw. $z' = 1,0$ für „kleine“, $d' = 1,036 \approx 1$ bzw. $z' = 2,5$ für „mittlere“ oder $d' = 2,326 \approx 2,3$ bzw. $z' = 4,0$ für „große“ Stufenabstände bzw. Horizontspreizungen anzugeben. Um dabei direkt die obere Grenze des Vertrauensintervalls der Standardabweichung berechnen zu können, wird aus den Quantilen der Faktor $K_{s,5\%} = 10^{-s'_{lg,5\%}}$ ermittelt. Damit kann die obere Grenze des einseitigen Vertrauensintervalls zu einer Irrtumswahrscheinlichkeit

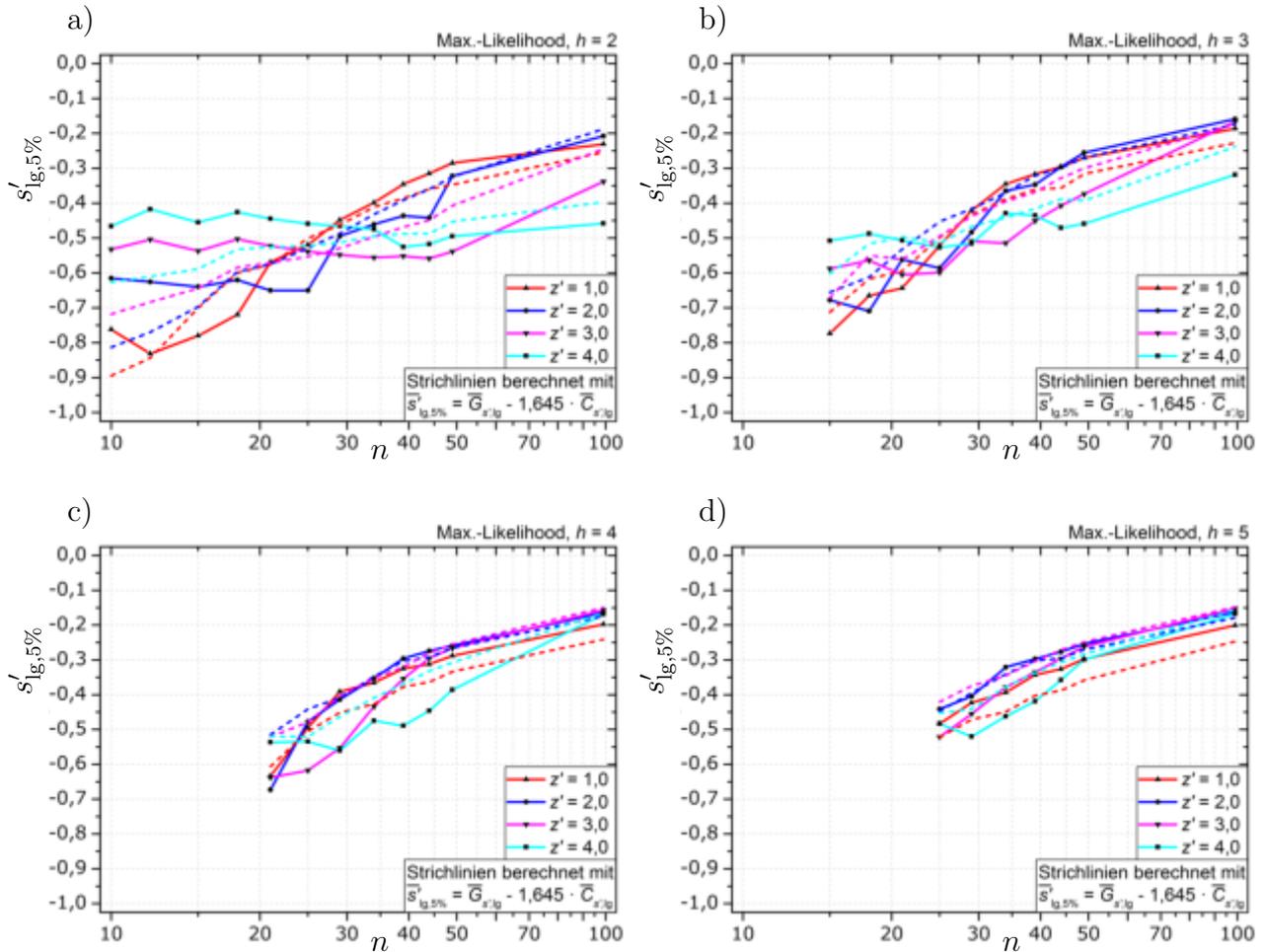


Abb. 4.25: Mittlere 5%-Quantile des Schätzers s'_{lg} für mittels MAX.-LIKELIHOOD-METHODE ausgewertete Horizontversuche auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

von $\alpha = 5\%$ durch

$$\sigma_G \leq K_{s,5\%} \cdot s$$

angegeben werden. Die so ermittelten Faktoren sind in Abbildung 4.26 für Treppenstufenversuche und in Abbildung 4.27 für Horizontversuche dargestellt. Für die Treppenstufenversuche werden hier die rechnerischen Quantile als konservative Abschätzung verwendet, während für die Horizontversuche die empirischen mittleren Quantile dargestellt sind. Zusätzlich ist in beiden Abbildungen noch der Faktor $K_{s,50\%}$ angegeben, mit dem der Schätzer der Standardabweichung korrigiert werden muss, um eine mediantreue Schätzung zu erhalten. Gut zu erkennen ist in den Abbildungen, wie groß die Unterschätzungen der Standardabweichung bei der Auswertung von **Treppenstufenversuchen** ausfällt, wenn diese nach DIXON/MOOD oder der MAX.-LIKELIHOOD-METHODE ausgewertet werden. Erfolgt eine solche Schätzung anhand einer Versuchsreihe mit nur 20 Einzelversuchen, dann ist das Ergebnis mindestens mit einem Faktor von $K_{s,50\%} = 1,4$ zu multiplizieren, um eine mediantreue Schätzung zu erhalten. Für die Angabe einer einseitigen oberen Vertrauensgrenze (zu $\alpha = 0,05$) muss das Ergebnis mit einem Faktor von ca. $K_{s,5\%} = 4$ multipliziert werden. Demgegenüber ist es mit der Auswertung nach HÜCK möglich, nahezu mediantreue Ergebnisse zu erhalten. Für das Ergebnis einer Versuchsreihe mit 20 Einzelversuchen und einem mutmaßlich mittleren Stufenabstand lassen

4 Vergleich der Simulationsergebnisse

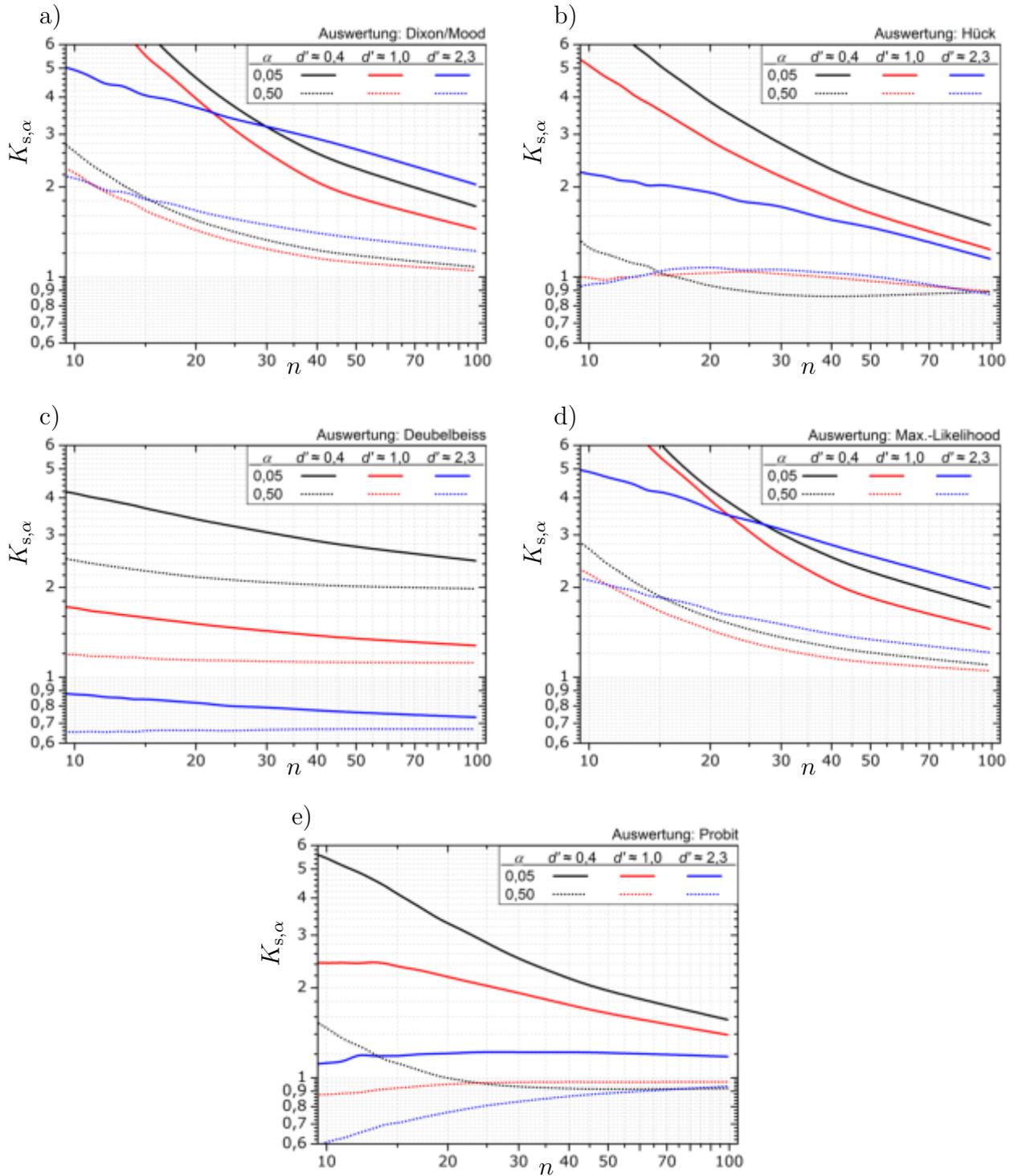


Abb. 4.26: Faktoren zur Bestimmung des einseitigen Vertrauensbereichs der Standardabweichung für Treppenstufenversuche bei kleinen, mittleren und großen Stufenabständen. Für die Auswertung nach DIXON/MOOD in a), HÜCK in b), DEUBELBEISS in c), mit MAX.-LIKELIHOOD-METHODE in d) sowie PROBIT-METHODE in e).

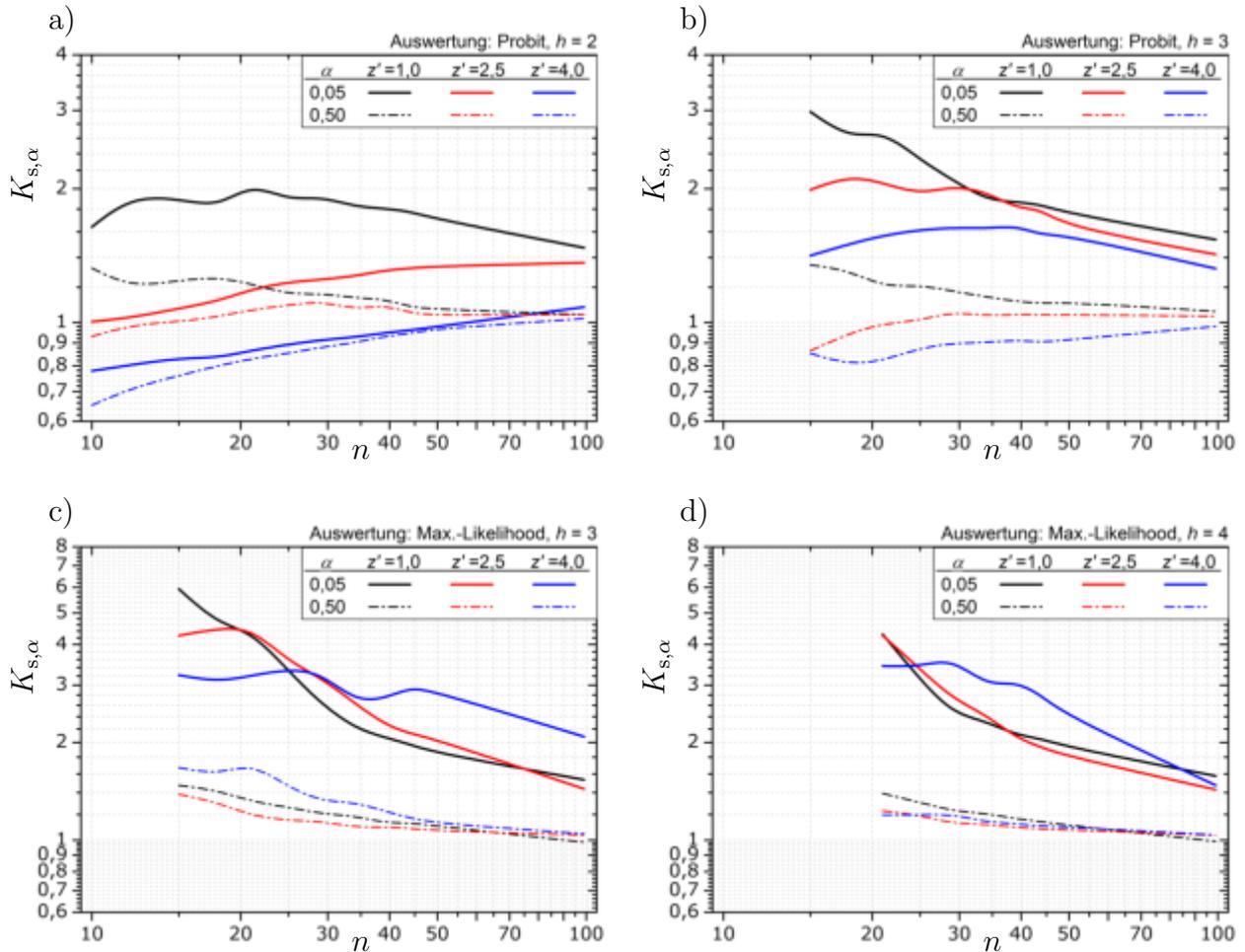


Abb. 4.27: Faktoren zur Bestimmung des einseitigen Vertrauensbereichs der Standardabweichung für Horizontversuche. Auswertung mittels PROBIT-METHODE von Versuchsreihen mit 2 Horizonten in a) und 3 Horizonten in b) sowie Auswertung mittels MAX.-LIKELIHOOD-METHODE von Versuchsreihen mit 3 Horizonten in c) und 4 Horizonten in d).

sich aus der Abbildung 4.26 b) die Faktoren $K_{s,5\%} = 4$ und $K_{s,5\%} = 2,8$ ablesen. Das bedeutet jedoch nicht, dass die Auswertungen nach HÜCK deutlich treffsicherere Ergebnisse liefern. Für die obere Vertrauensgrenze der tatsächlichen Standardabweichung würde das bereits um den systematischen Fehler korrigierte Schätzergebnis für die Auswertungen nach DIXON/MOOD mit dem Faktor $K = 4/1,4 = 2,86$ multipliziert. Für die Auswertung nach HÜCK wäre der Faktor weiterhin mit etwa $K = 2,8$ anzusetzen. Die Ergebnisse **Horizontversuche** zeigen, dass sich mit dieser Versuchsführung und ausreichend großem Stichprobenumfang die Standardabweichungen mediantreu schätzen lassen. Deutlich kleinere Vertrauensgrenzen können sich aber im Vergleich zu den Treppenstufenversuchen nicht angeben lassen.

4.4 Schätzung des 2,5 %-Quantils bei bekannter Standardabweichung

4.4.1 Mittlerer zufälliger und systematischer Schätzfehler

Wenn die Verteilungsparameter der Grundgesamtheit bekannt sind, dann berechnet sich das 2,5 %-Quantil einer normalverteilten Zufallsgröße durch

$$x_{2,5\%} = \mu_G - 1,96 \cdot \sigma_G$$

aus den Verteilungsparametern μ_G und σ_G^2 . Liegt für einen oder beide Verteilungsparameter der Grundgesamtheit jeweils nur ein Schätzwert vor, ist auch das berechnete 2,5 %-Quantil nur eine Schätzung. Im Falle eines geschätzten Erwartungswertes und einer bekannten Standardabweichung ist die Unsicherheit dieser Schätzung ebenso groß wie die bereits untersuchte Unsicherheit des geschätzten Erwartungswertes. Im Falle eines geschätzten Erwartungswertes und einer geschätzten Standardabweichung fällt die Unsicherheit für das geschätzte 2,5 %-Quantil deutlich größer aus. Für einen objektiven Vergleich der einzelnen Auswertemethoden wird der Schätzfehler des 2,5 %-Quantils auf die tatsächliche Standardabweichung normiert, d. h. der Schätzer x' gebildet, vgl. Abschnitt 3.4.5 S. 85. Es ist dann möglich, die einzelnen Auswerteverfahren bzgl. der Schätzfehler untereinander zu vergleichen. Für die praktische Anwendung sind die angegebenen Fehler uninteressant, da die tatsächliche Standardabweichung in den meisten Fällen nicht bekannt sein dürfte und andernfalls das 2,5 %-Quantil nicht nach Gleichung 3.20 ermittelt werden würde.

Die berechneten systematischen Schätzfehler bzw. Verzerrungen des Schätzers x' sind für die **Treppenstufenversuche** in Abbildung 4.28 dargestellt. Die Auswertungen nach DIXON/MOOD und der MAX.-LIKELIHOOD-METHODE schätzen das 2,5 %-Quantil nahezu identisch und weisen asymptotisch erwartungstreues Verhalten auf. Eine Vergrößerung des Stufenabstandes führt bei geringem Stichprobenumfang zu einer kleinen Verringerung des systematischen Schätzfehlers, bei großem Stichprobenumfang vergrößert sich hingegen die Verzerrung des Schätzers x' . Die Auswertung nach DEUBELBEISS zeigt kein erwartungstreues oder asymptotisch erwartungstreues Verhalten. Der systematische Schätzfehler ist stark von der Größe des Stufenabstandes abhängig. Dies führt bei großen Stufenabständen zu einer deutlichen systematischen Unterschätzung des 2,5 %-Quantils, die auch durch das Hinzufügen weiterer Versuche nicht verbessert werden kann. Die Auswertungen nach HÜCK schätzen das 2,5 %-Quantil bei kleinen Stufenabständen im Mittel zu klein, sind also hier stets konservative Abschätzungen. Bei mittleren und großen Stufenabständen sind nahezu erwartungstreue Schätzungen möglich. Auffallend ist die Vergrößerung des systematischen Schätzfehlers bei Versuchsreihen mit $n > 50$ Einzelversuchen. Es ist zu vermuten, dass dieser Umstand der nicht ausreichenden Korrektur der Varianzkennzahl k_h und der damit verbundenen schlechteren Schätzung der Standardabweichung geschuldet ist, vgl. Abschnitt 4.3.2. Mit der Auswertung nach der PROBIT-METHODE erfolgen die Schätzungen immer stark konservativ, aber asymptotisch erwartungstreu. Der Betrag des systematischen Schätzfehlers ist nahezu immer größer als bei Schätzungen mit der PROBIT-METHODE oder nach der MAX.-LIKELIHOOD-METHODE. Die Wahl der Anschnittsdefinition hat hingegen keine merklichen Auswirkungen auf die Größe des systematischen Schätzfehlers. Richtet sich die

Wahl der Auswertemethode vor allem nach der Erwartungstreue der Schätzung, dann ist unter den Treppenstufenauswertungen die Auswertung nach HÜCK zu bevorzugen. Mit ihr ist es möglich, das 2,5 %-Quantil nahezu erwartungstreu und stets konservativ zu schätzen, wenn der Stufenabstand nicht zu klein gewählt wird. Die Auswertungen nach DEUBELBEISS lassen zwar prinzipiell auch erwartungstreue Schätzungen zu, dafür muss aber der Stufenabstand möglichst exakt festgelegt werden. Ist dies nicht der Fall, treten teils sehr große systematische, nicht konservative Schätzfehler auf, die auch durch Erhöhung des Versuchsumfanges nicht reduziert werden können. Bei einer Schätzung des 2,5 %-Quantils mittels **Horizontversuchen**, vgl. Abbildung 4.29, fallen die Verzerrungen des Schätzers allgemein geringer aus. Werden Versuche mit mittleren Horizontspreizungen ($z' \approx 2$) durchgeführt, sind die Schätzungen nahezu erwartungstreu. Ausgenommen davon sind Auswertungen mittels **PROBIT-METHODE** für Versuchsreihen mit mehr als 3 Horizonten. Hier zeigt sich, dass es bei dieser Art der Auswertung vorteilhaft ist, möglichst viele Einzelversuche auf möglichst wenigen Horizonten anzuordnen, um kleine Verzerrungen zu erhalten. Genau umgekehrt verhält es sich bei den Auswertungen nach der **MAX.-LIKELIHOOD-METHODE**. Hier verringern sich die Verzerrungen dann, wenn möglichst viele Horizonte mit Versuchen belegt werden. Bemerkenswert ist außerdem, dass Schätzungen nach der **PROBIT-METHODE** im Mittel konservativ ausfallen, während Schätzungen mit der **MAX.-LIKELIHOOD-METHODE** wie schon bei den Treppenstufenversuchen im Mittel zu nichtkonservativen Ergebnissen führen. Wenn eine erwartungstreue Schätzung des 2,5 %-Quantils gefordert wird, ist daher eine Auswertung nach der **PROBIT-METHODE** zu bevorzugen. Die ermittelten zufälligen Schätzfehler bzw. die Standardfehler des Schätzers x' sind für **Treppenstufenversuche** in der Abbildung 4.30 zusammengefasst. Unter allen betrachteten Auswerteverfahren für Treppenstufenversuche zeigen die Auswertungen nach DEUBELBEISS stets die geringsten Standardfehler. Ebenfalls recht geringe zufällige Fehler zeigen die Auswertung nach DIXON/MOOD und der **MAX.-LIKELIHOOD-METHODE**. Die Ergebnisse beider Methoden unterscheiden sich nur bei kleinen Stufenabständen geringfügig. Bei mittleren und großen Stufenabständen liefern beide Methoden das gleiche Ergebnis. Die Auswertungen nach der **PROBIT-METHODE** und nach HÜCK liefern meist die Schätzer des 2,5 %-Quantils mit den größten Standardfehlern. Lediglich bei großen Stufenabständen und Versuchsreihen mit vielen Einzelversuchen ist es möglich, kleinere Standardfehler als mit der Auswertung nach DIXON/MOOD oder der **MAX.-LIKELIHOOD-METHODE** zu erhalten. Die Ergebnisse der **Horizontauswertungen**, vgl. Abbildung 4.31, zeigen bei kleinen Horizontspreizungen sehr variable aber vor allem deutlich größere Standardfehler als bei den Auswertungen der Treppenstufenversuche. Das ist erstaunlich, da die ermittelten Standardfehler der Erwartungswertschätzung (S. 103) und der Standardabweichungsschätzung (S. 121) solche großen Unterschiede nicht erwarten lassen. Eine Begründung dafür ist vermutlich, dass in beiden Auswerteverfahren die Verteilungsparameter nicht unabhängig voneinander ermittelt werden. Bei kleinen Horizontspreizungen führt eine zufällige ungünstige Lage und Belegung der Horizonte zu gleichzeitig sehr schlechten Schätzergebnissen für den Erwartungswert und die Standardabweichung der Grundgesamtheit. Erst bei einer ausreichend großen Horizontspreizung von $z' \geq 2$ verringert sich dieser Effekt. Die Schätzer des 2,5 %-Quantils besitzen dann die geringsten Standardfehler, wenn Versuchsreihen mit 2 oder 3 Horizonten nach der **PROBIT-METHODE** und Versuchsreihen mit mehr Horizonten nach der **MAX.-LIKELIHOOD-METHODE** ausgewertet werden. Selbst dann ist es aber kaum möglich, für Schätzungen des 2,5 %-Quantils geringere zufällige Fehler zu ermitteln, als dies mit den Treppenstufenauswertungen möglich ist.

4 Vergleich der Simulationsergebnisse

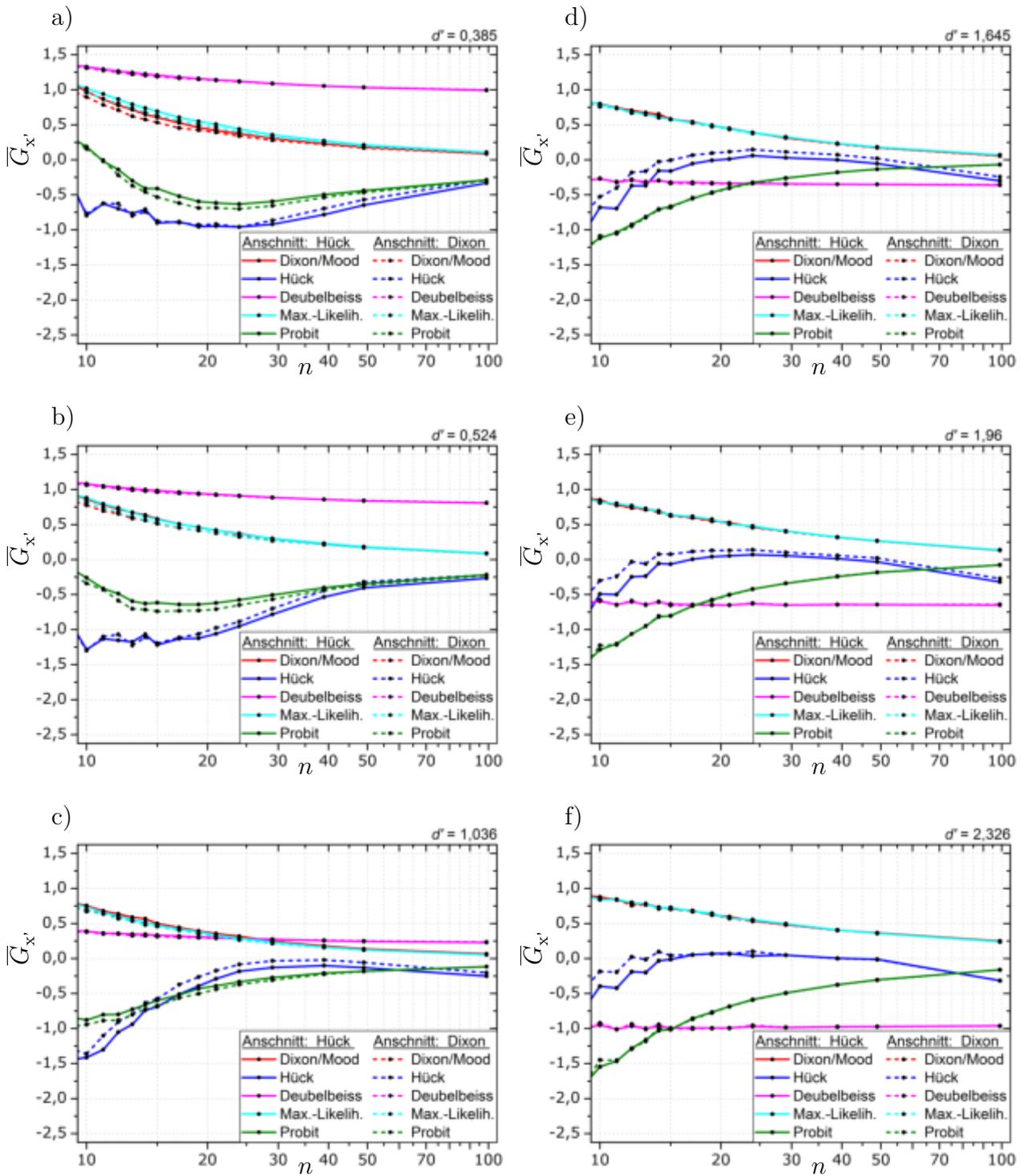


Abb. 4.28: Mittlere Verzerrungen \bar{G}_x des Schätzers x' aus Treppenstufenversuchen für verschiedene bezogene Stufenabstände d' .

4.4 Schätzung des 2,5%-Quantils bei bekannter Standardabweichung

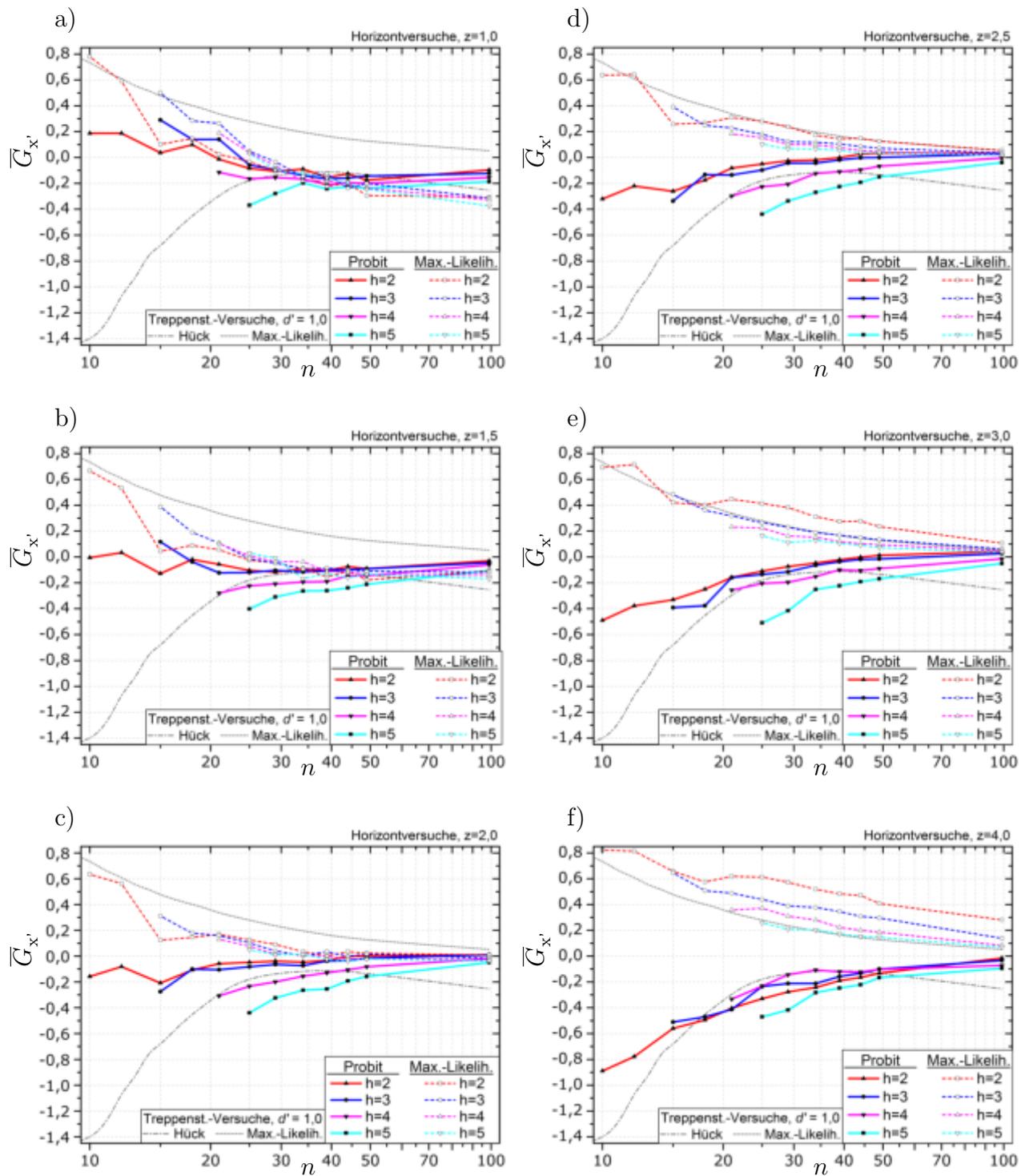


Abb. 4.29: Mittlere Verzerrungen \bar{G}_x des Schätzers x' aus Horizontversuchen für verschiedene bezogene Horizontspreizungen z' . Zum Vergleich mit eingetragen sind die Ergebnisse aus Treppenstufenauswertungen nach HÜCK und der MAX.-LIKELIHOOD-METHODE für $d' \approx 1$.

4 Vergleich der Simulationsergebnisse

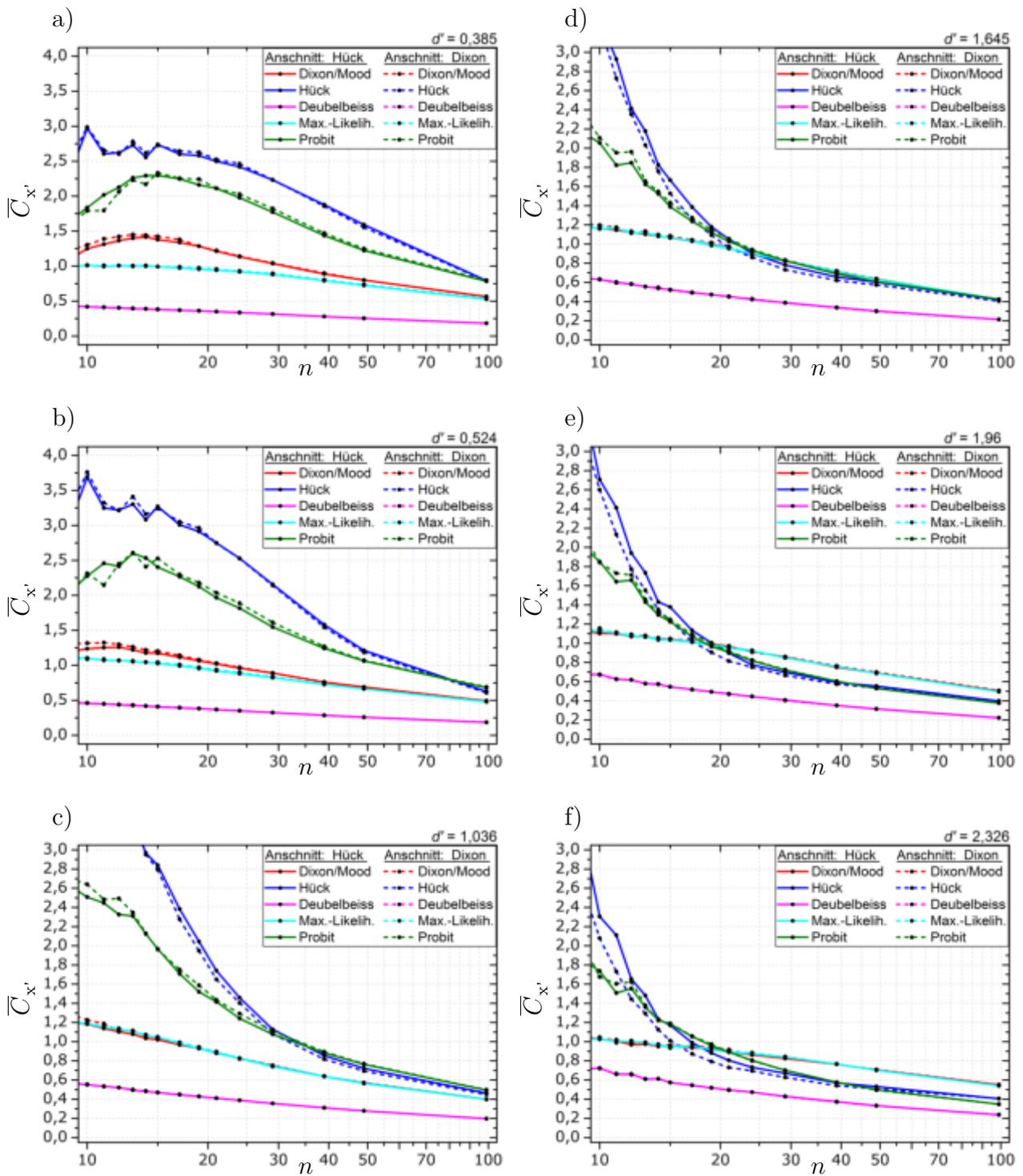


Abb. 4.30: Mittlere Standardfehler \bar{C}_x' des Schätzers x' aus Treppenstufenversuchen für verschiedene bezogene Stufenabstände d' .

4.4 Schätzung des 2,5%-Quantils bei bekannter Standardabweichung

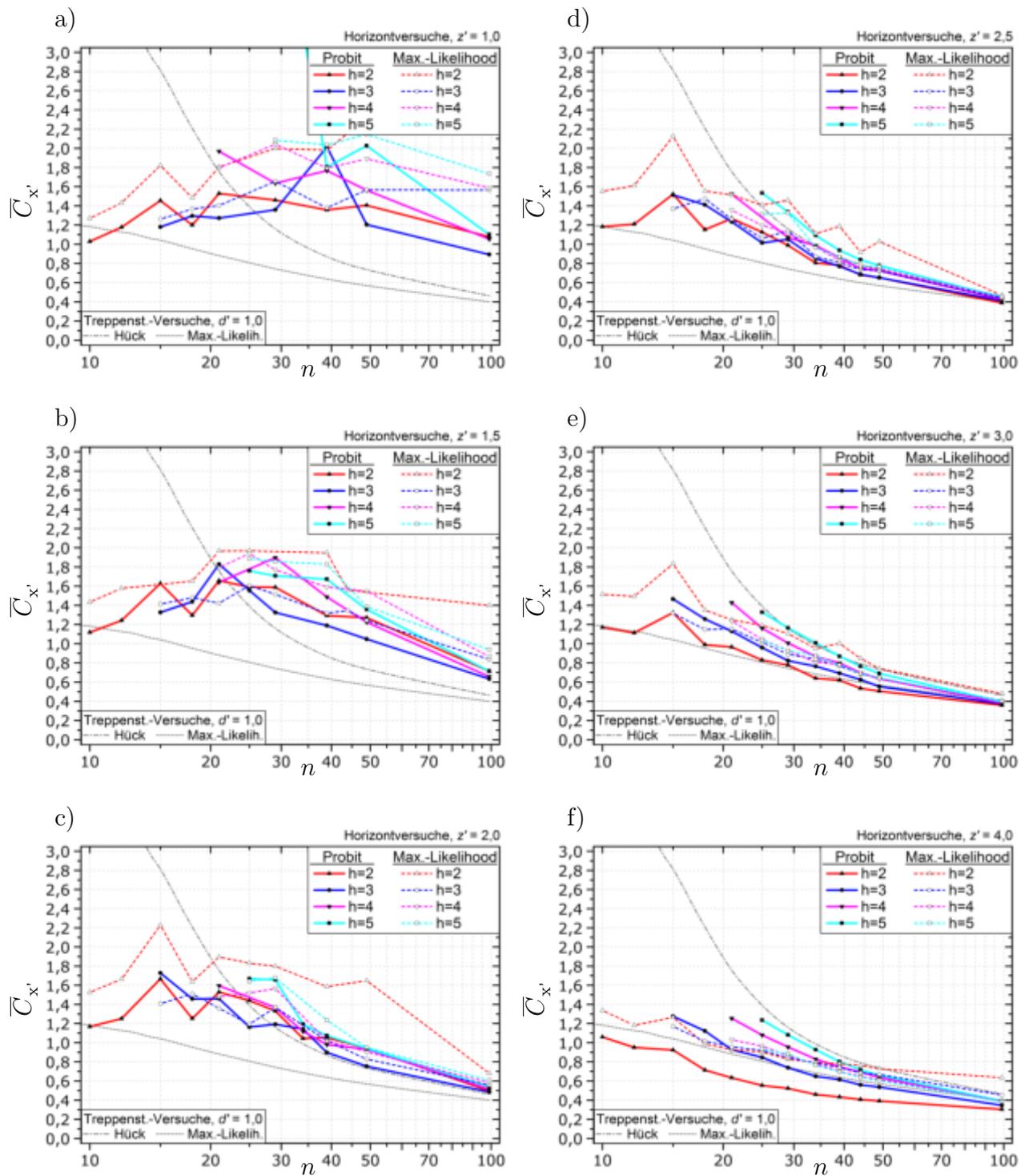


Abb. 4.31: Mittlere Standardfehler $\bar{C}_{x'}$ des Schätzers x' für verschiedene bezogene Horizontspreizungen z' . Zum Vergleich mit eingetragen sind die Ergebnisse der Treppenstufenauswertungen nach HÜCK und der MAX.-LIKELIHOOD-METHODE für $d' \approx 1$.

4.4.2 RMS-Fehler

Wie aus dem systematischen und zufälligen Schätzfehler der mittlere RMS-Fehler des 2,5 %-Quantils ermittelt wird, ist in Abschnitt 3.4.5.1 auf S. 85 beschrieben. Der RMS-Fehler ist geeignet, den Fehler der Schätzung einer einzelnen Versuchsreihe zu beschreiben. Eine Unterscheidung in konservative oder weniger konservative Schätzungen ist nicht möglich. Auch führt eine Mittlung der Ergebnisse mehrerer Versuchsreihen nur bei jeweils erwartungstreuen Schätzmethode auch zu einer Verringerung des RMS-Fehlers.

Die Ergebnisse für **Treppenstufenversuche** sind in der Abbildung 4.32 dargestellt. Bis auf die Auswertung nach DEUBELBEISS unterscheiden sich alle Ergebnisse nur geringfügig vom im vorherigen Abschnitt beschriebenen Standardfehler. Das ist nicht besonders verwunderlich, da die Verzerrung des Schätzers x' meist deutlich kleiner als der Standardfehler ausfällt und damit der RMS-Fehler nahezu vollständig durch den Standardfehler beschrieben wird. Die Auswertung nach DEUBELBEISS zeigt deutlich, dass die bei kleinen oder großen Treppenstufenabständen vorhandene starke Verzerrung der Schätzung zu einem großen RMS-Fehler führt. Dieser liegt dann in Bereichen, wie sie bei kleinen Stufenabständen auch nach der MAX.-LIKELIHOOD-METHODE und bei großen Stufenabständen durch die Auswertung nach HÜCK erreicht werden. Insgesamt bietet die MAX.-LIKELIHOOD-METHODE bei kleinen Stufenabständen die Möglichkeit, das 2,5 %-Quantil mit dem geringsten RMS-Fehler zu schätzen. Bei großen Stufenabständen zeigt die Auswertung nach HÜCK die geringsten RMS-Fehler. Werden Treppenstufenversuche mit nur wenigen Versuchen ($n \approx 10$) durchgeführt, ist eine einmalige Schätzung des 2,5 %-Quantils nach DEUBELBEISS im Mittel näher am tatsächlichen 2,5 %-Quantil als alle anderen Auswertemethoden der Treppenstufenversuche. Es kann also gerechtfertigt sein, die Auswertungen nach DEUBELBEISS durchzuführen, um einen Schätzwert zu erhalten. Problematisch wird dieses Vorgehen nur dann, wenn die Ergebnisse mehrerer Versuchsreihen gemittelt werden, da so eine deutliche Überschätzung des 2,5 %-Quantils entstehen kann, siehe Abbildung 4.28. Mit allen anderen Auswerteverfahren ist es bei vertretbarem Versuchsaufwand nicht möglich, das 2,5 %-Quantil mit einem RMS-Fehler $\overline{RMS}_{x'} \leq 0,5 \cdot \sigma_G$ zu schätzen. Selbst ein recht großer RMS-Fehler von $1 \cdot \sigma_G$ wird nur von Versuchsreihen mit mehr als 20 Einzelversuchen erreicht.

Die Auswertung von **Horizontversuchen** zur Schätzung des 2,5 %-Quantils sind nur sinnvoll, wenn Horizontspreizungen von $z' \geq 2$ vorliegen. Andernfalls ergeben sich, wie im vorhergehenden Abschnitt 4.4.1 beschrieben, zufällige Schätzfehler in Größenordnungen, die es unmöglich machen, dem Ergebnis irgendeine Vertrauenswürdigkeit zuzuschreiben. Die Ergebnisse solcher Horizontversuche sind in Abbildung 4.33 dargestellt. Prinzipiell ist es demnach möglich, mit Horizontversuchen Schätzer mit kleineren RMS-Fehlern zu erhalten, als mit Treppenstufenversuchen und einer Auswertung nach HÜCK. Wird eine möglichst große Horizontspreizung ($z' \geq 3,0$) gewählt, ergeben sich für die nach der PROBIT-METHODE ausgewerteten Horizontversuche die geringsten RMS-Fehler, wenn die Versuche auf möglichst wenige Horizonte verteilt werden. Im Gegensatz dazu ist es bei einer Auswertung der Versuchsreihen nach der MAX.-LIKELIHOOD-METHODE von Vorteil, die Versuche auf möglichst viele Horizonte zu verteilen.

Da es in der praktischen Anwendung unmöglich ist, den Stufenabstand oder die Horizontspreizung exakt vorzugeben, scheint es zielführend unter allen betrachteten Verfahren jene zu bevorzugen, deren Fehler nicht nur klein, sondern auch möglichst stabil verlaufen, d. h.

4.4 Schätzung des 2,5 %-Quantils bei bekannter Standardabweichung

wenig empfindlich auf die Änderung von Stufenabstand oder Horizontspreizung reagieren. Unter diesem Gesichtspunkt scheinen lediglich die Treppenstufenauswertungen nach der MAX.-LIKELIHOOD-METHODE oder nach DIXON/MOOD zur Bestimmung des 2,5 %-Quantils infrage zu kommen. Allerdings wird hier das 2,5 %-Quantil systematisch überschätzt. Bei der Durchführung von Horizontversuchen lassen sich mit der PROBIT-METHODE auf 2 oder 3 Horizonten die besten Ergebnisse erzielen. Die wirksamste Methode, den RMS-Fehler zu minimieren, ist die Vergrößerung des Versuchsumfangs.

4 Vergleich der Simulationsergebnisse

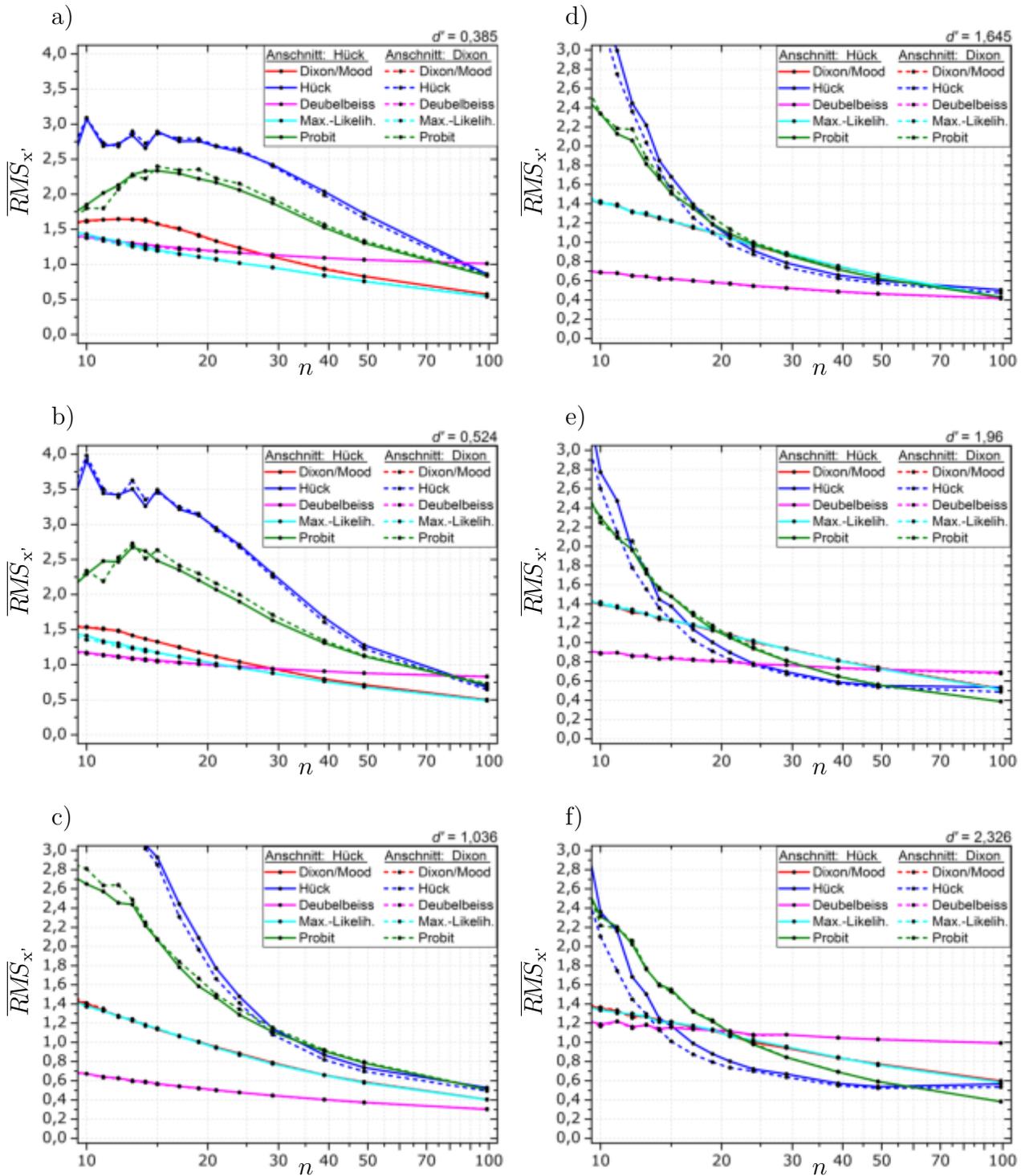


Abb. 4.32: Mittlere RMS-Fehler $\overline{RMS}_{x'}$ des Schätzers x' für verschiedene bezogene Stufenabstände d' .

4.4 Schätzung des 2,5%-Quantils bei bekannter Standardabweichung

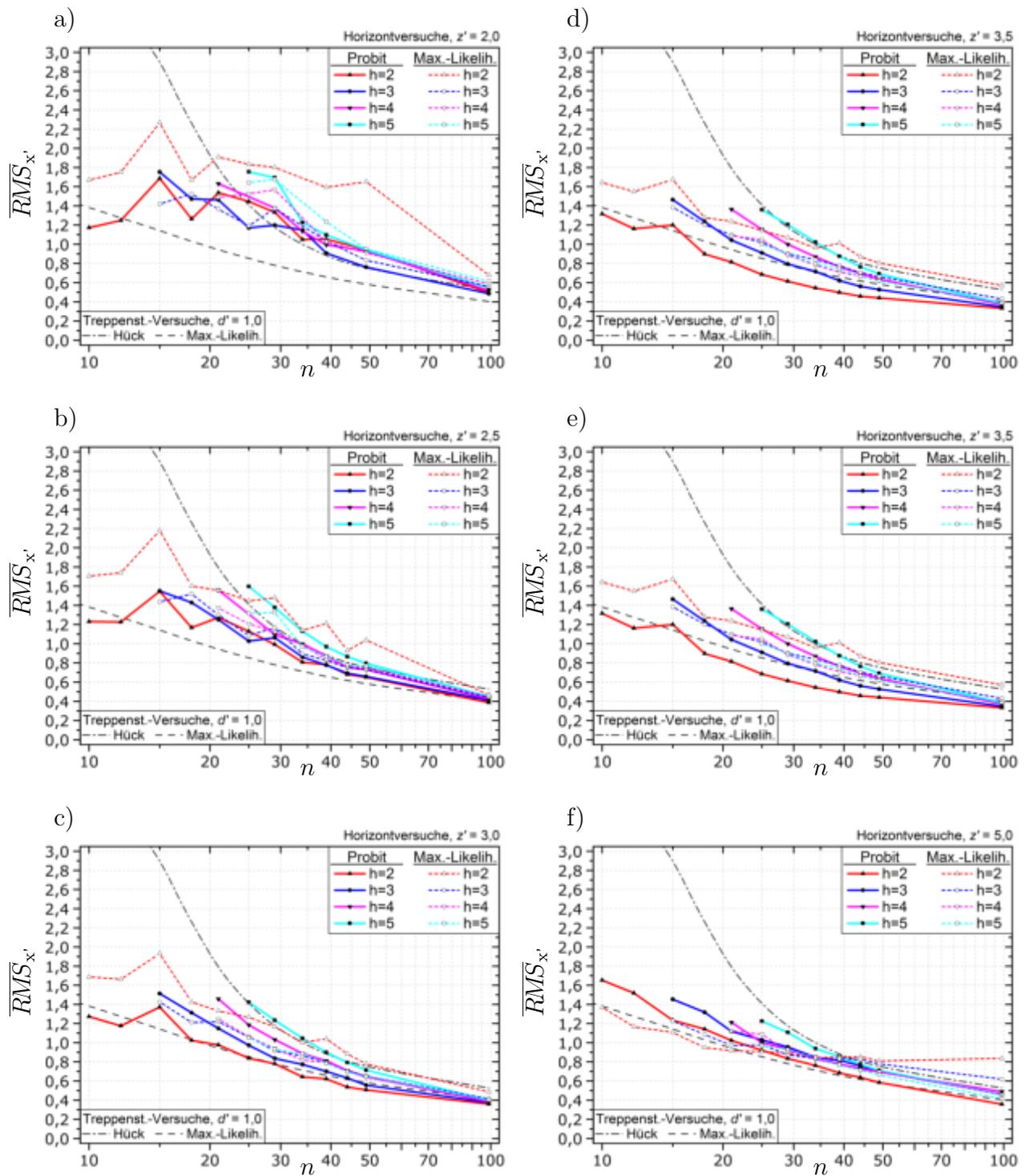


Abb. 4.33: Mittlere RMS-Fehler $\overline{RMS}_{x'}$ des Schätzers x' für verschiedene bezogene Horizontspreizungen z' . Zum Vergleich mit eingetragen sind die Ergebnisse aus Treppenstufenauswertungen nach HÜCK und der MAX.-LIKELIHOOD-METHODE für $d' \approx 1$.

4.4.3 Empirischer Vertrauensbereich des geschätzten 2,5%-Quantils

Die 95 %-Quantile des Schätzers x' geben die größten auftretenden Schätzfehler zu einem Niveau von $(1 - \alpha) = 0,95$ an. Mit ihnen ist es möglich, aus dem geschätzten 2,5 %-Quantil $x_{2,5\%}$ eine untere Vertrauensgrenze für das tatsächliche 2,5 %-Quantil $X_{2,5\%,G}$ zu ermitteln. Dies kann dann mit

$$X_{2,5\%,G} \leq x_{2,5\%} - \overline{x'}_{95\%} \cdot \sigma_G$$

angegeben werden. Hierbei ist natürlich wieder zu beachten, dass die Quantile mithilfe der geschätzten Standardabweichung berechnet werden, die Vertrauensgrenzen aber aus den Quantilen des auf die tatsächliche Standardabweichung bezogenen Schätzers x' folgen. Die ermittelten Vertrauensgrenzen sind daher nicht geeignet, die in praktischen Versuchen auftretenden Fehler zu beurteilen. Die Ergebnisse können aber für einen objektiven Vergleich der Schätzgüte herangezogen werden. Mit Blick auf die in Abbildung 4.34 dargestellten Quantile des Schätzers x' für **Treppenstufenversuche** ist festzustellen, dass es bei den Auswertungen nach DIXON/MOOD sowie mit der MAX.-LIKELIHOOD-METHODE für verschiedenste Stufenabstände die größten maximalen Überschätzungen des Quantils der Festigkeit zu erwarten sind. Mit der Auswertung nach HÜCK sind die maximalen Überschätzungen stets kleiner. Noch geringer fallen sie mit den Auswertungen mittels PROBIT-METHODE oder nach DEUBELBEISS aus. Bei der PROBIT-METHODE liegt das vor allem an der systematischen Unterschätzung des 2,5 %-Quantils der Festigkeit. Sowohl mit der Auswertung nach HÜCK als auch mittels PROBIT-METHODE werden zum Teil deutlich kleinere maximale Verschätzungen ermittelt, als das mit den ermittelten RMS-Fehlern oder den Verzerrungen und Standardfehlern des Schätzers x' zu erwarten ist. Der Grund hierfür liegt in der teils deutlichen Unterschätzung der 2,5 %-Quantils, vgl. Abbildung 3.24 auf S. 89. Die schiefen Häufigkeitsverteilungen des Schätzers dieser Auswertemethoden lassen sich nicht mit der Normalverteilung oder ähnlichen symmetrischen Verteilungen approximieren.

Die Abbildung 4.35 zeigt die 95 %-Quantile des Schätzers x' für **Horizontversuche**. Zum Vergleich sind hier ebenfalls die Ergebnisse für die Treppenstufenauswertungen nach HÜCK sowie mittels MAX.-LIKELIHOOD-METHODE für einen mittleren Stufenabstand von $d' = 1,0$ eingetragen. Wie bei allen bisherigen Ergebnissen auch, zeigen die Ergebnisse der PROBIT-METHODE geringere Fehler als die der MAX.-LIKELIHOOD-METHODE. Ebenso verringern sich die Fehler bei der PROBIT-METHODE mit abnehmender Horizontanzahl, während sie sich bei der MAX.-LIKELIHOOD-METHODE mit abnehmender Horizontanzahl vergrößern. Im Vergleich mit den nach gleicher Methode ausgewerteten Treppenstufenversuchen bei einem recht günstigen mittleren Stufenabstand von $d' = 1,0$, ist es mit Horizontversuchen nicht möglich, kleinere maximale Schätzfehler zu erhalten. Demgegenüber sind die maximalen Fehler bei einer Auswertung mittels PROBIT-METHODE nahezu für alle simulierten Horizontspreizungen kleiner als bei einer Treppenstufenauswertung nach HÜCK.

Ist das Ziel, bei der experimentellen Schätzung des 2,5 %-Quantils der Schwingfestigkeit möglichst die geringsten maximalen Überschätzungen zu erhalten, dann ist eine Auswertung von Horizontversuchen und eine Auswertung mittels PROBIT-METHODE anzustreben. Versuchsreihen mit zwei Horizonten zeigen hier die besten Ergebnisse. Versuchsreihen mit drei Horizonten sind am robustesten gegen Veränderung der Horizontspreizung. Bei entsprechend großer Versuchsanzahl lassen sich auch bei Treppenstufenauswertungen und der Auswertung nach HÜCK vergleichbare Ergebnisse erzielen. Auch die Auswertung von Treppenstufenversuchen

4.4 Schätzung des 2,5 %-Quantils bei bekannter Standardabweichung

mittels der PROBIT-METHODE führt zu geringen minimalen Überschätzungen. Allerdings ist das vor allem auf die systematische Unterschätzung des 2,5 %-Quantils zurückzuführen. Die Auswertung von Treppenstufenversuchen nach DEUBELBEISS scheint aufgrund der starken Abhängigkeit der Schätzfehler vom gewählten Stufenabstand nicht für die Schätzung des 2,5 %-Quantils geeignet.

4 Vergleich der Simulationsergebnisse

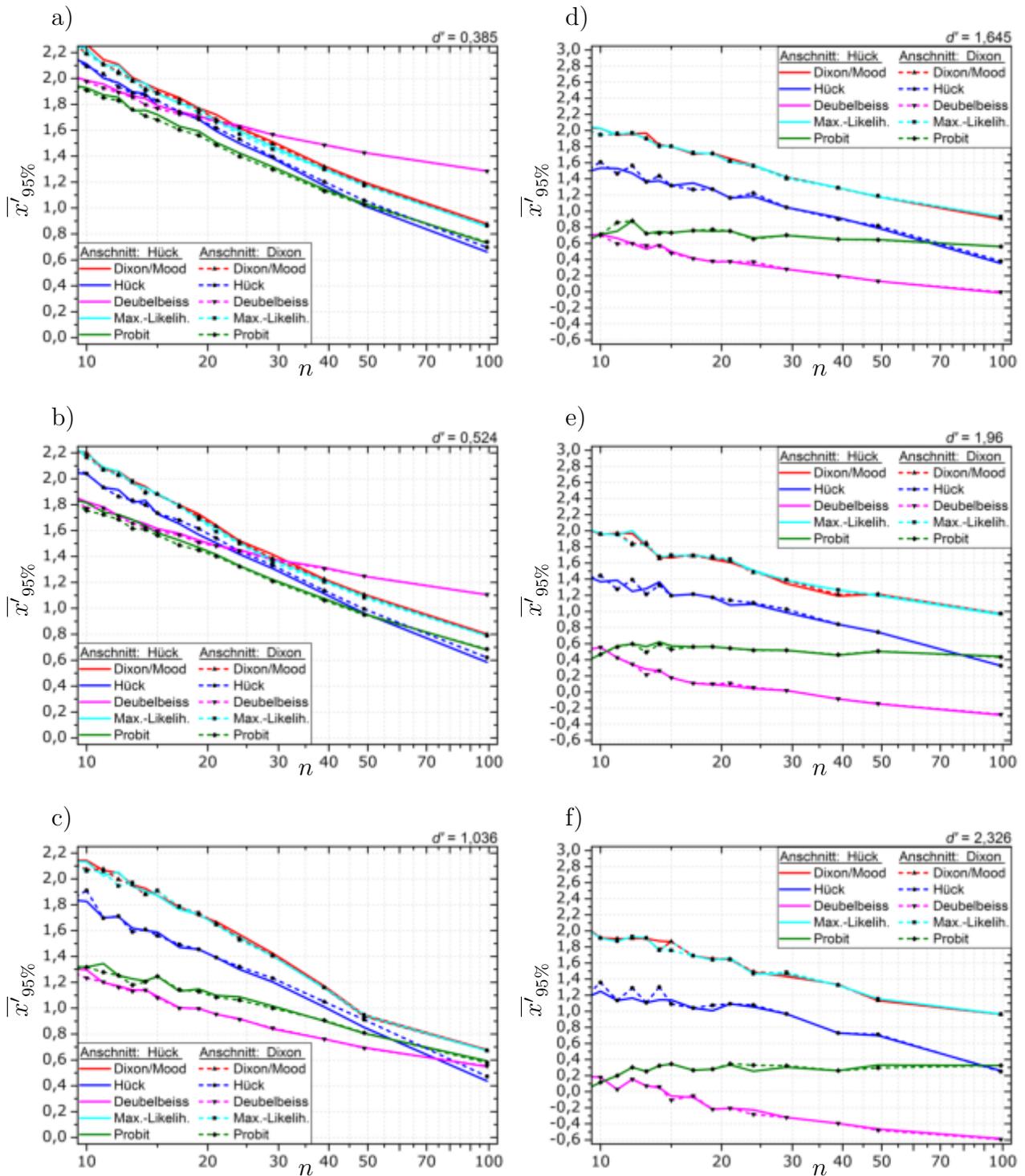


Abb. 4.34: Mittlere 95 %-Quantile $\bar{x}'_{95\%}$ des Schätzers x' für Treppenstufenversuche bei verschiedenen Stufenabständen.

4.4 Schätzung des 2,5 %-Quantils bei bekannter Standardabweichung

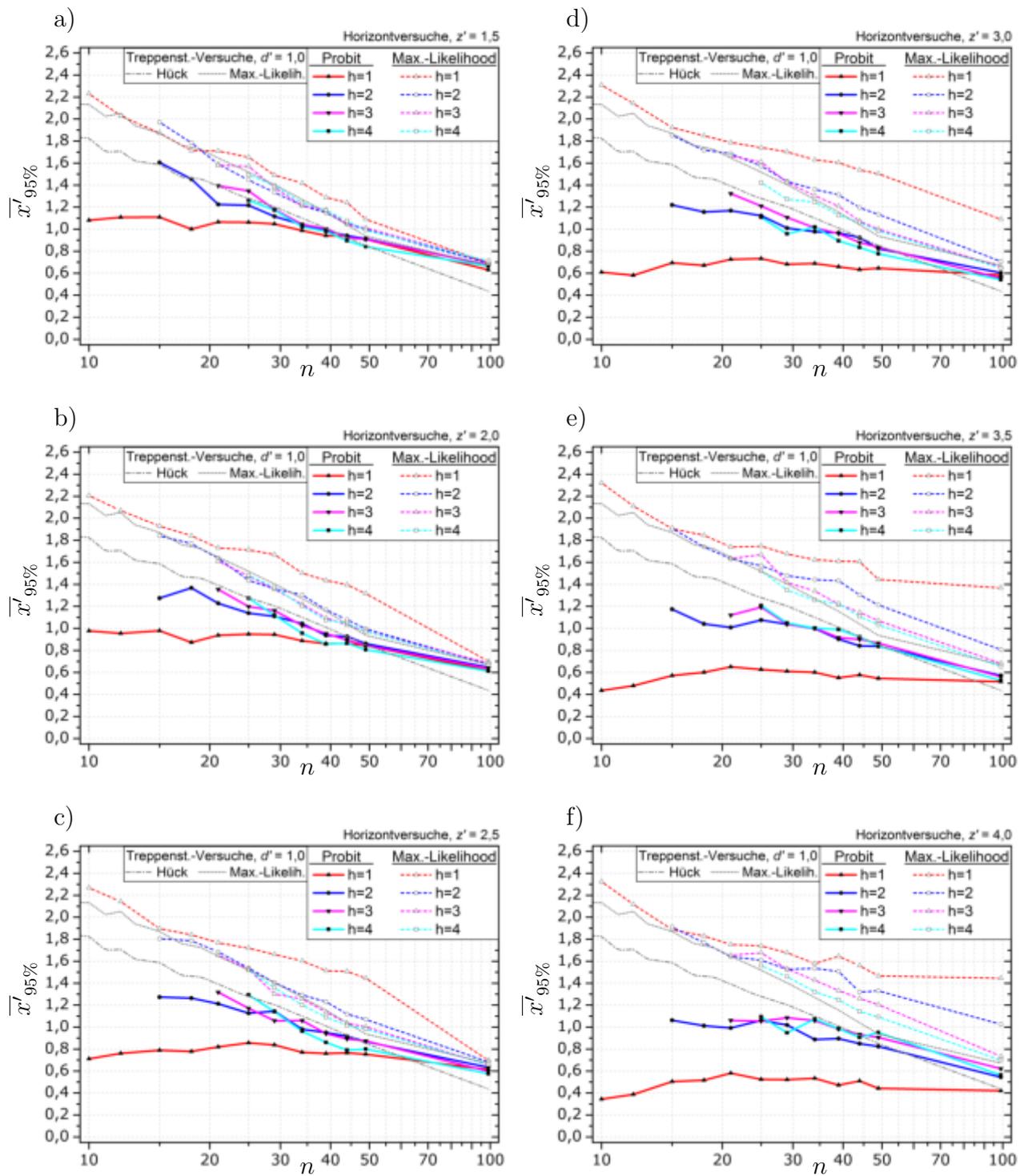


Abb. 4.35: Mittlere 95 %-Quantile $\bar{x}'_{95\%}$ des Schätzers x' für Horizontversuche bei verschiedenen Horizontspreizungen.

4.5 Schätzung des Erwartungswertes bei unbekannter Standardabweichung

4.5.1 Mittlerer zufälliger und systematischer Schätzfehler

Der Schätzer m^* gibt den Schätzfehler des Erwartungswertes in Vielfachen der im selben Versuch geschätzten Standardabweichung s' an, vgl. im Abschnitt 3.4.4.1, S. 81. Mit der mittleren Verzerrung \overline{G}_{m^*} und dem mittleren Standardfehler \overline{C}_{m^*} ist es möglich, die Größe des systematischen und zufälligen Schätzfehlers für den in einer Versuchsreihe ermittelten Schätzwert m anzugeben, wenn gleichzeitig auch die Größe der Standardabweichung geschätzt wird. Das vollständige Messergebnis m_k wird dann durch

$$m_k = m - \overline{G}_{m^*} \cdot s \pm \overline{C}_{m^*} \cdot s$$

berechnet. Da wie oben beschrieben, die Schätzungen der Standardabweichung unterschiedlich stark und vor allem unterschiedlich mediantreu ausfallen, ist ein relativer Vergleich der verschiedenen Auswertemethoden anhand der Faktoren \overline{G}_{m^*} und \overline{C}_{m^*} nicht möglich.¹⁰

Für die Schätzungen aus **Treppenstufenversuchen** sind die ermittelten mittleren Verzerrungen und Standardfehler des Schätzers m^* in Abbildung 4.36 zusammengefasst. Die Schätzfehler sind weiterhin von der Größe des tatsächlichen bezogenen Stufenabstandes abhängig. Dieser ist in praktischen Versuchen unbekannt und kann nur geschätzt werden. Die Ergebnisse sind daher exemplarisch für „kleine“, „mittlere“ und „große“ Stufenabstände dargestellt. Im Vergleich zu den Schätzfehlern bei bekannter Standardabweichung (Abbildung 4.3 und 4.4 ab S. 101) wird deutlich, wie sich durch die Unsicherheit der Standardabweichungsschätzung die Unsicherheit der Erwartungswertschätzung verändert. Im Allgemeinen müssen deutlich größere Unsicherheiten angegeben werden.¹¹ Gleiches gilt für die Ergebnisse der **Horizontversuche**, die in der Abbildung 4.37 dargestellt sind. Auch hier können die Fehler nur für „kleine“, „mittlere“ und „große“ tatsächliche Horizontspreizungen dargestellt werden.

¹⁰Für jede Auswertemethode besitzt der Schätzer m^* einen anderen Bezug. Es ist also nicht zwangsweise davon auszugehen, dass große Verzerrungen oder Standardfehler des Schätzers m^* auch große Verzerrungen oder Standardfehler des Schätzers m bedeuten.

¹¹Das ist auch zu beachten, wenn die Standardfehler nach HÜCK [21] ermittelt werden.

4.5 Schätzung des Erwartungswertes bei unbekannter Standardabweichung

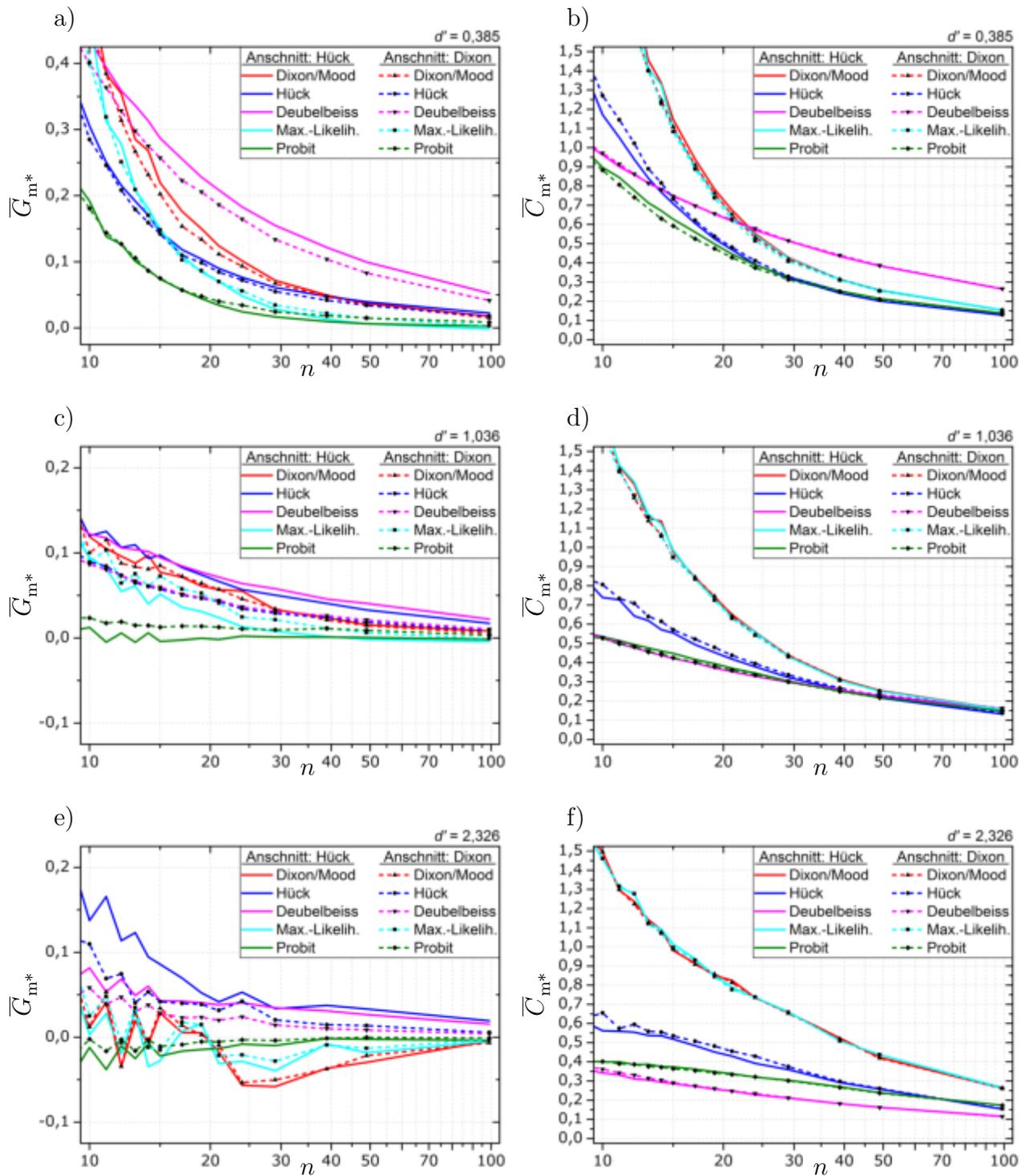


Abb. 4.36: Mittlere Verzerrungen \bar{G}_{m^*} und Standardfehler \bar{C}_{m^*} des Schätzers m^* für verschiedene Treppenstufenauswertungen bei kleinem tatsächlichen Stufenabstand von $d' \approx 0,4$ in a) und b), mittlerem tatsächlichen Stufenabstand $d' \approx 1,0$ in c) und d) sowie großem tatsächlichen Stufenabstand $d' \approx 2,3$ in e) und f).

4 Vergleich der Simulationsergebnisse

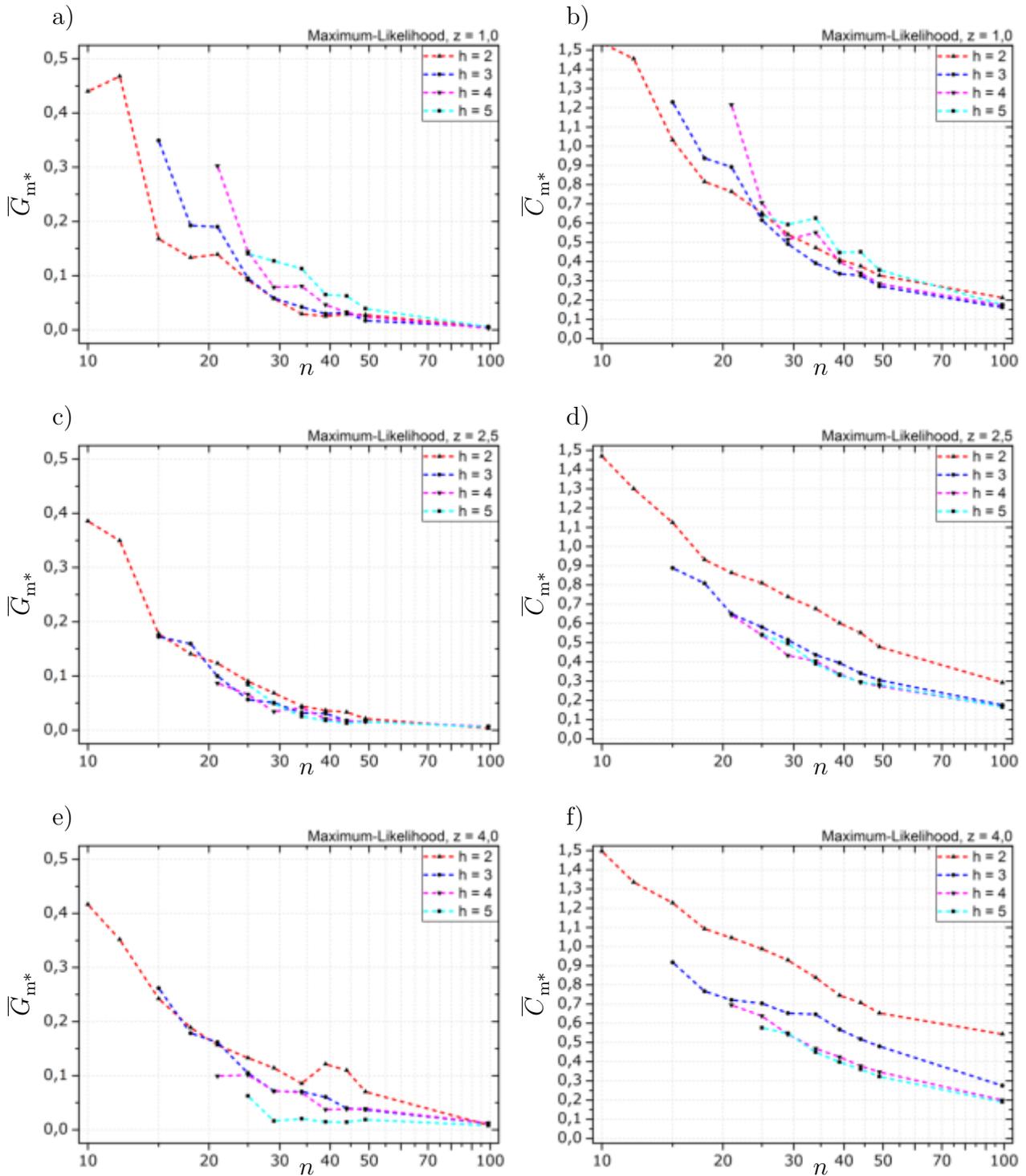


Abb. 4.37: Mittlere Verzerrungen \bar{G}_{m^*} und Standardfehler \bar{C}_{m^*} des Schätzers m^* für mittels MAX.-LIKELIHOOD-METHODE ausgewertete Horizontversuche bei kleiner tatsächlicher Horizontspreizung von $z' = 1,0$ in a) und b), mittlerer tatsächlicher Horizontspreizung $z' = 2,5$ in c) und d) sowie großer tatsächlicher Horizontspreizung $z' = 4,0$ in e) und f).

4.5 Schätzung des Erwartungswertes bei unbekannter Standardabweichung

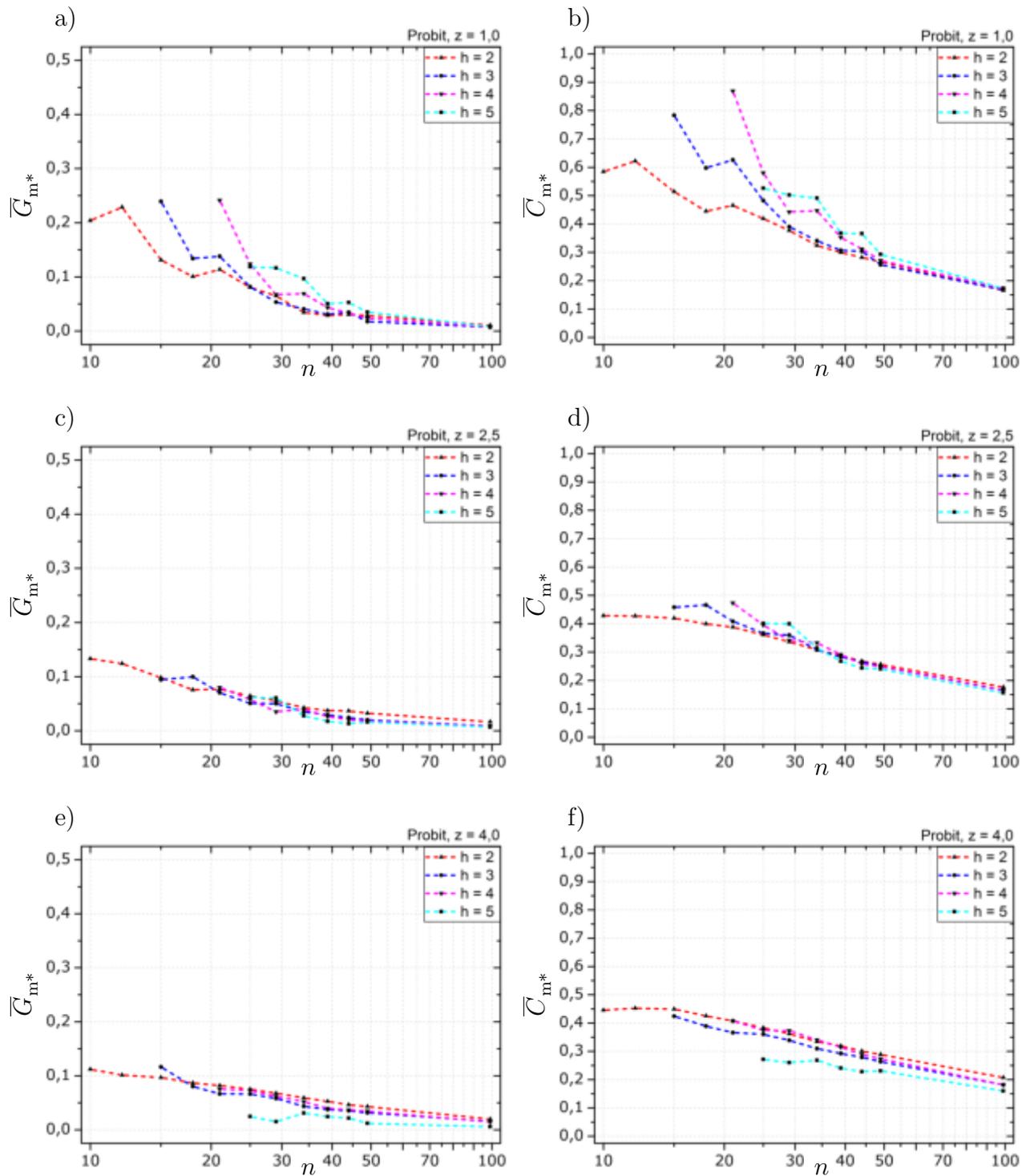


Abb. 4.38: Mittlere Verzerrungen \bar{G}_{m^*} und Standardfehler \bar{C}_{m^*} des Schätzers m^* für mittels PROBIT-METHODE ausgewertete Horizontversuche bei kleiner Horizontspreizung von $z' = 1,0$ in a) und b), mittlerer Horizontspreizung $z' = 2,5$ in c) und d) sowie großer Horizontspreizung $z' = 4,0$ in e) und f).

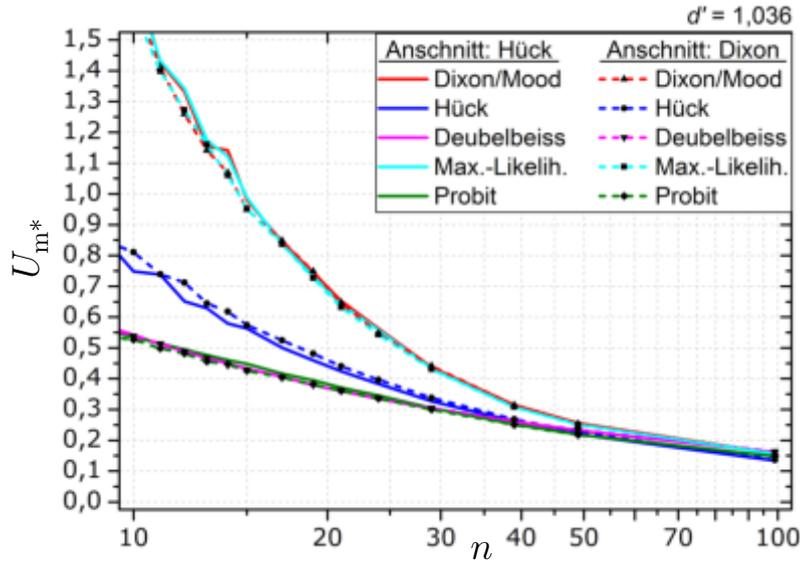


Abb. 4.39: Verallgemeinerter Standardfehler U_{m^*} für verschiedene Treppenstufenauswertungen mit einem mittleren Stufenabstand $d' \approx 1,0$.

4.5.2 Verallgemeinerter Standardfehler

Ist über den Startwert der Versuchsreihe nichts bekannt, d. h. wird die Versuchsreihe mit ihrem Startwert genauso oft zufällig über wie auch unter dem Erwartungswert der Festigkeit begonnen, kann zur Beschreibung der Unsicherheiten der verallgemeinerte Standardfehler verwendet werden, vgl. Abschnitt 3.4.4.2 ab S. 82. Das vollständige Messergebnis m_k , also der Schätzwert m für den Erwartungswert der Grundgesamtheit und dessen Messunsicherheit, kann dann durch

$$m_k = m \pm U_{m^*} \cdot s$$

angegeben werden.¹² Die verallgemeinerten Standardfehler für **Treppenstufenversuchen** sind in Abbildung 4.39 für „mittlere Stufenabstände“ und in Abbildung 4.40 für „kleine“ sowie „große“ Stufenabstände dargestellt. Für **Horizontversuche** werden die verallgemeinerten Standardfehler des Schätzers m^* in der Abbildung 4.41 und 4.42 festgehalten. Die berechneten Werte sind weitestgehend mit den oben berechneten mittleren Standardfehlern \bar{C}_{m^*} identisch. Der Grund hierfür liegt in der, im Vergleich zum Standardfehler, geringen Verzerrung des Schätzers.

¹²Auch hier lassen die Unterschiede in der Größe der ermittelten verallgemeinerten Standardfehler zwischen den einzelnen Auswertemethoden keine Rückschlüsse auf die tatsächliche Schätzgüte des Erwartungswertes der Grundgesamtheit durch die einzelnen Methoden zu. Die verallgemeinerten Standardfehler geben jeweils nur an, wie groß die Verschätzung anhand der im selben Versuch geschätzten Standardabweichung zu bewerten ist. So wird zum Beispiel für den im Treppenstufenverfahren mit 20 gültigen Versuchen bei einer Auswertung nach HÜCK ermittelten Schätzwert des Erwartungswertes der Grundgesamtheit eine Messunsicherheit von etwa $0,45 \cdot s$ abgelesen, vgl. Abbildung 4.39. Für die Schätzung des Erwartungswertes aus Horizontversuchen mit ebenso vielen Einzelversuchen auf 3 Horizonten bei einer mittleren Horizontspreizung und einer Auswertung mittels PROBIT-METHODE liegt die Messunsicherheit ebenfalls bei etwa $0,45 \cdot s$, vgl. Abbildung 4.41 e). Das ändert nichts an der Tatsache, dass die Schätzung des Erwartungswertes mit dem Treppenstufenverfahren nach HÜCK genauere Schätzwerte liefert, vgl. Abbildung 4.7 auf S. 107. Die Ergebnisse können nur aufgrund der geringeren Treffsicherheit der Schätzung der Standardabweichung in Treppenstufenversuchen nicht genauso gut wie bei Horizontverfahren beurteilt werden, siehe Abschnitt 4.3.

4.5 Schätzung des Erwartungswertes bei unbekannter Standardabweichung

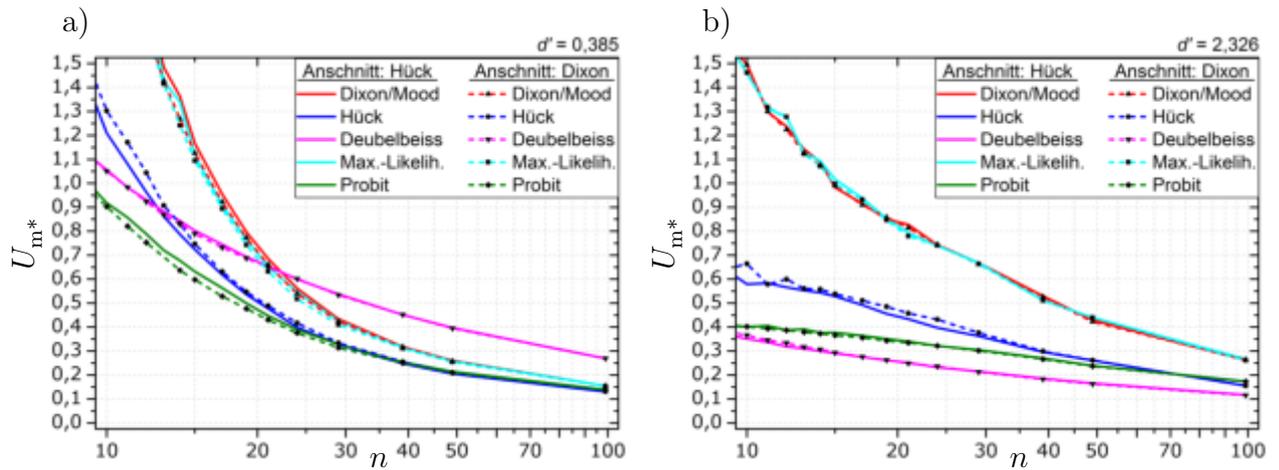


Abb. 4.40: Verallgemeinerter Standardfehler U_{m^*} für verschiedene Treppenstufenauswertungen bei einem kleinen Stufenabstand von $d' \approx 0,4$ in a) und einem großen Stufenabstand $d' \approx 2,3$ in b).

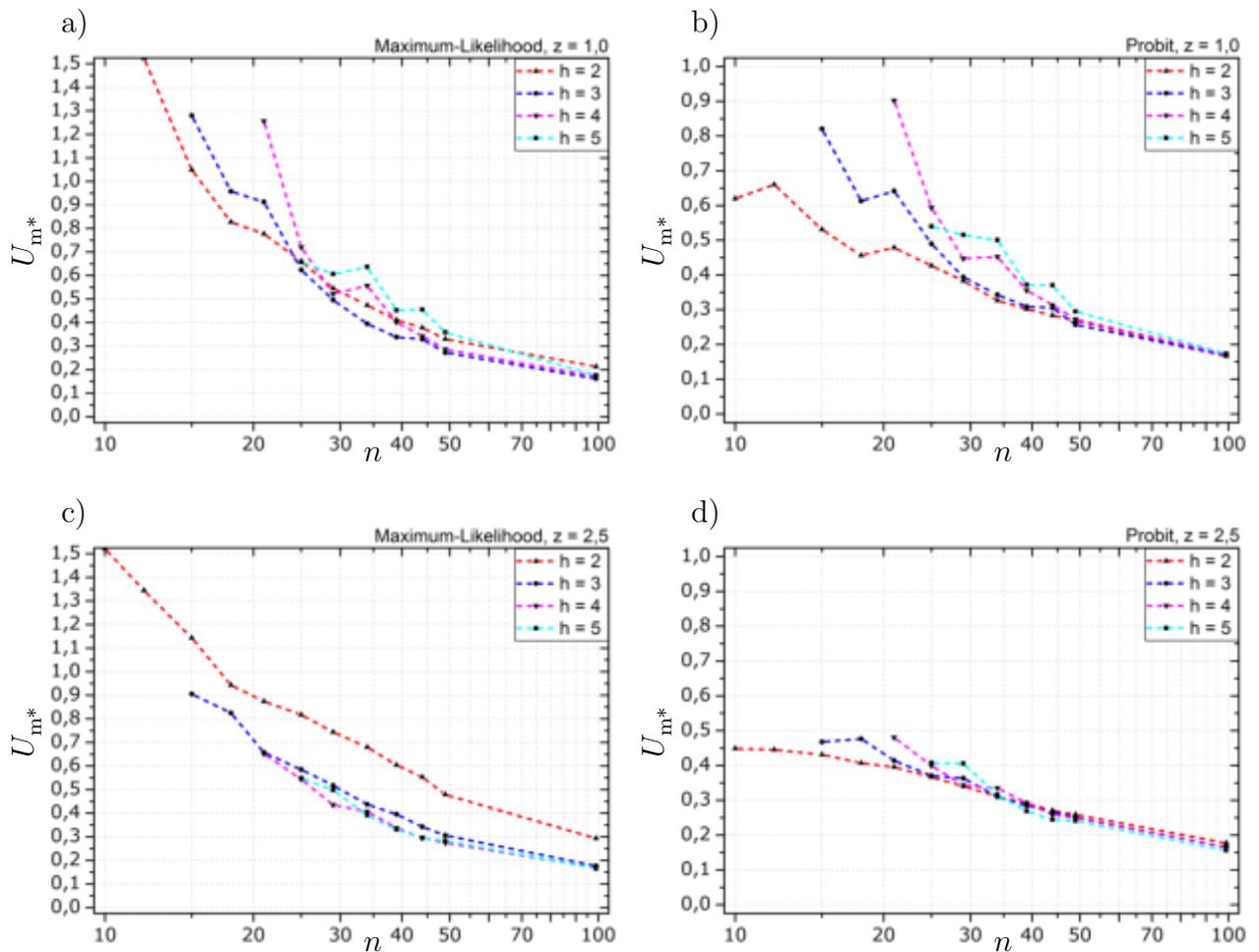


Abb. 4.41: Verallgemeinerter Standardfehler U_{m^*} des Schätzers m^* bei Auswertung von Horizontversuchen mittels PROBIT-METHODE und MAX.-LIKELIHOOD-METHODE. Für kleine Horizontspreizungen von $z' = 1,0$ in a) und b) sowie für mittlere Horizontspreizungen $z' = 2,5$ in c) und d).

4 Vergleich der Simulationsergebnisse

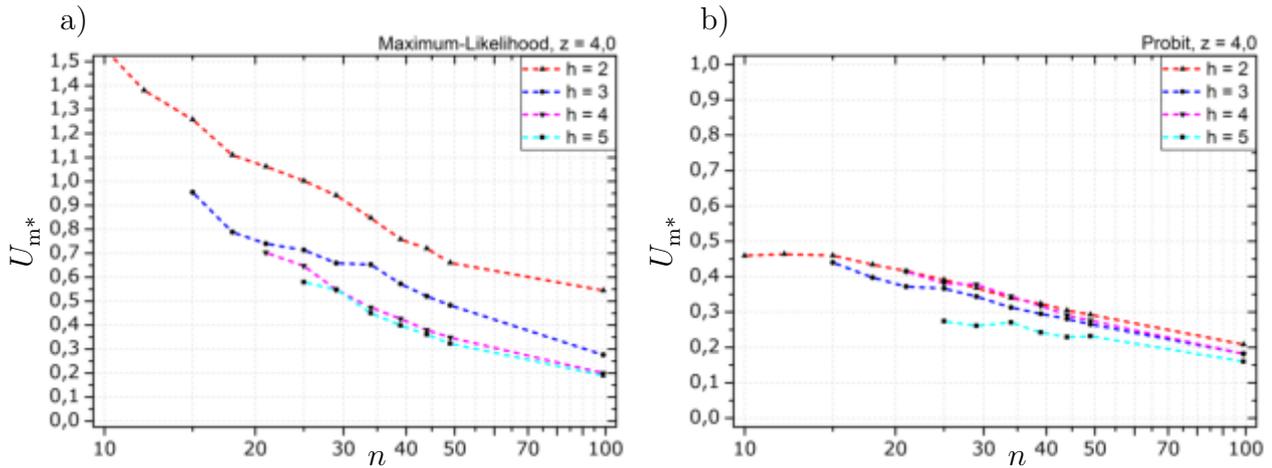


Abb. 4.42: Verallgemeinerter Standardfehler U_{m^*} für Horizontversuche. Auswertung mittels PROBIT-METHODE und MAX.-LIKELIHOOD-METHODE bei großen Horizontspreizungen von $z' = 4,0$ in a) und b).

4.5.3 Empirischer Vertrauensbereich des geschätzten Erwartungswertes

Aus den empirischen Quantilen der Simulationsergebnisse werden, wie in Abschnitt 3.4.4.4 beschrieben, die mittleren verallgemeinerten Quantile $\overline{m^*}_{|5\%|}$ ermittelt. Sie erlauben es, das Konfidenzintervall des Erwartungswertes zu den Irrtumswahrscheinlichkeiten $\alpha = 10\%$ mit

$$[m - \overline{m^*}_{|5\%|} \cdot s, m + \overline{m^*}_{|5\%|} \cdot s]$$

abzuschätzen. Die Ergebnisse der Treppenstufenversuche sind in Abbildung 4.43 und 4.45 dargestellt. Die Ergebnisse aus den Horizontversuchen sind der Abbildungen 4.44 zu entnehmen.

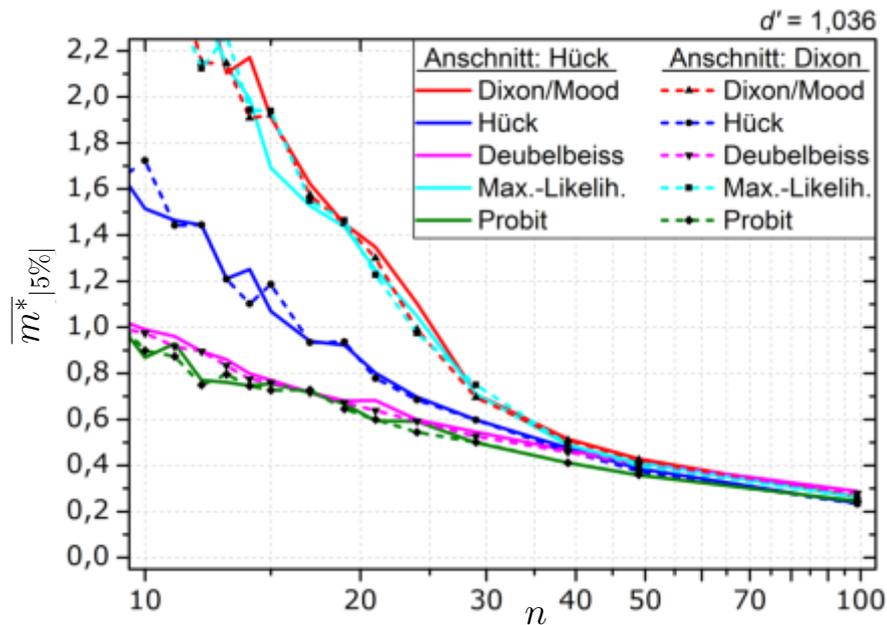


Abb. 4.43: Mittlere verallgemeinerte 5%-Quantile $\overline{m^*}_{|5\%|}$ des Schätzers m^* für Treppenstufenversuche bei einem mittleren Stufenabstand von $d' \approx 1,0$.

4.5 Schätzung des Erwartungswertes bei unbekannter Standardabweichung

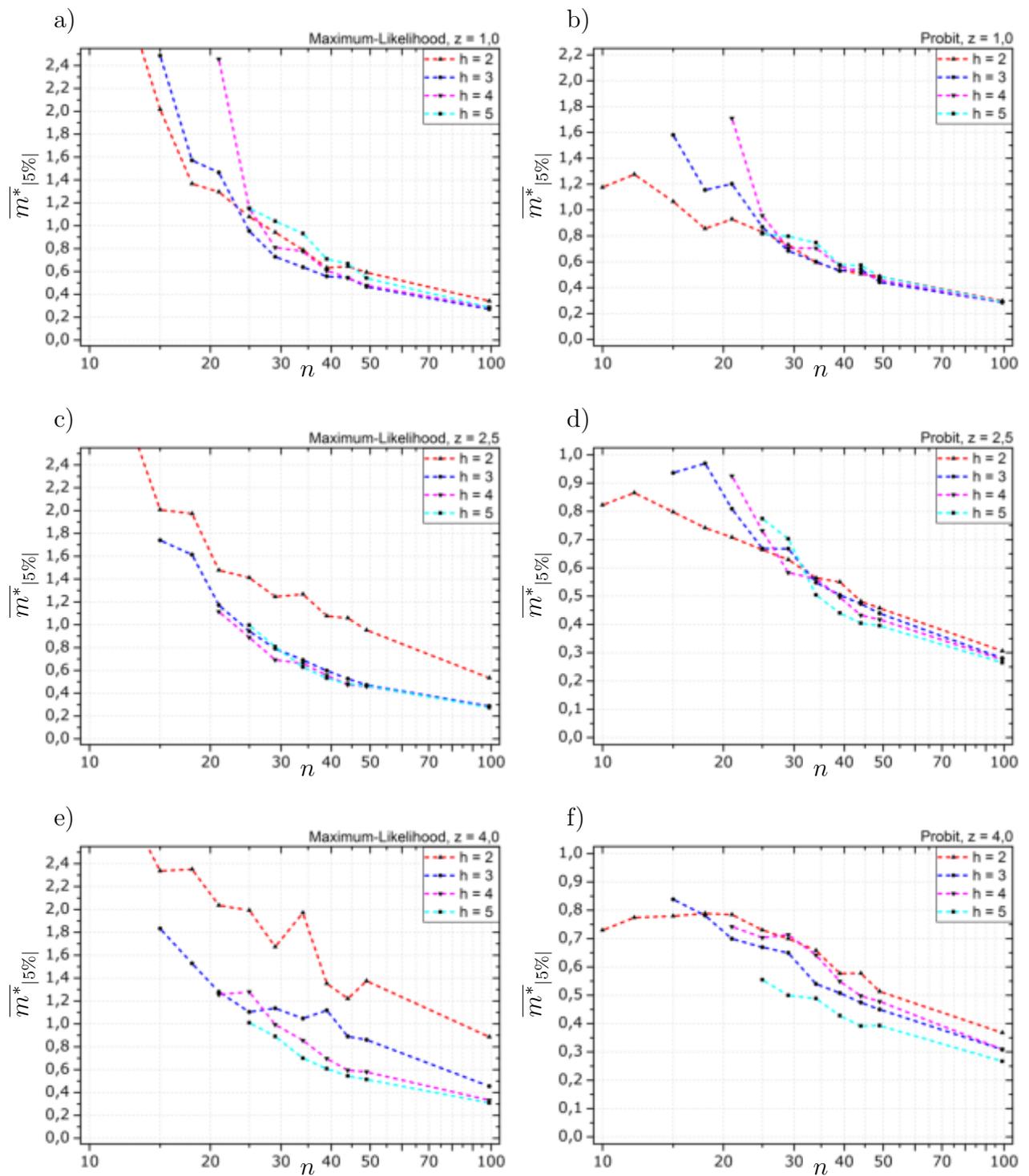


Abb. 4.44: Mittlere verallgemeinerte 95%-Quantile $\overline{m^*}_{|95\%|}$ des Schätzers m^* für Treppenstufenversuche mit Auswertung nach MAX.-LIKELIHOOD-METHODE und PROBIT-METHODE. Für kleine $z' = 1,0$ in a) und b), mittlere $z' = 2,5$ in c) und d) sowie große Horizontspreizungen $z' = 4,0$ in e) und f).

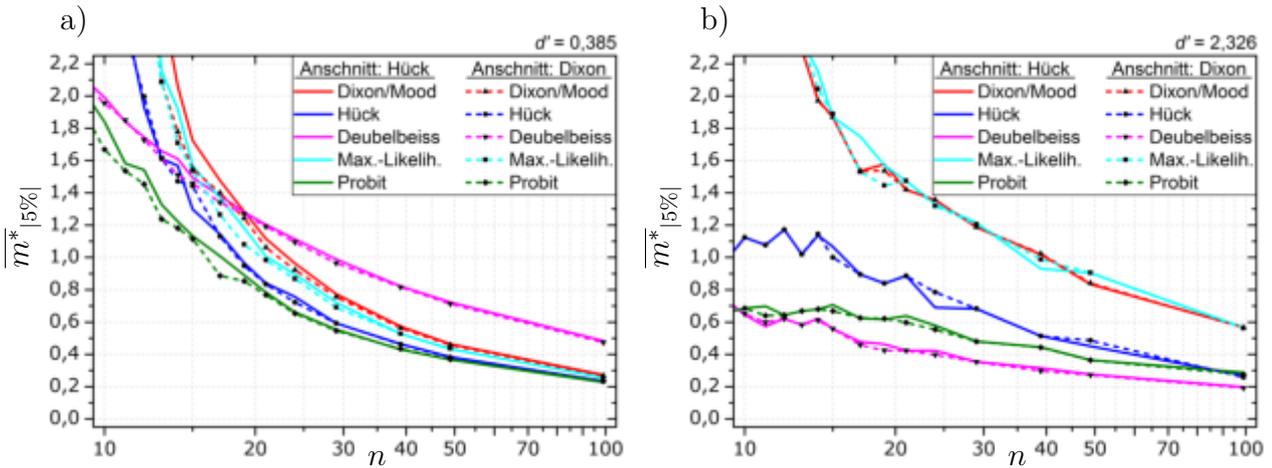


Abb. 4.45: Mittlere verallgemeinerte 5 %-Quantile $\overline{m^*}_{|5\%|}$ des Schätzers m^* für Treppenstufenversuche bei kleinem Stufenabstand $d' \approx 0,4$ in a) und großem Stufenabstand $d' \approx 2,3$ in b).

4.6 Schätzung des 2,5 %-Quantils bei unbekannter Standardabweichung

4.6.1 Mittlerer zufälliger und systematischer Schätzfehler

Mit dem Schätzer x^* wird der Abstand zwischen dem geschätzten und dem wahren 2,5 %-Quantil, normiert auf die geschätzte Standardabweichung, angegeben, vgl. Abschnitt 3.4.6 S. 90. Anhand der berechneten mittleren Verzerrungen \overline{G}_{x^*} und der mittleren Standardfehler \overline{C}_{x^*} ist es möglich, die Unsicherheit der Schätzungen einer jeden Versuchsreihe zu beurteilen, auch wenn die tatsächliche Standardabweichung nicht bekannt ist. Das vollständige Messergebnis $x_{2,5\%,k}$ für das experimentell ermittelte 2,5 %-Quantil lautet dann

$$x_{2,5\%,k} = x_{2,5\%} - \overline{G}_{x^*} \cdot s \pm \overline{C}_{x^*} \cdot s. \quad (4.2)$$

Für die Treppenstufenversuche sind die ermittelten Ergebnisse in Abbildung 4.46 dargestellt. Die Ergebnisse der Horizontversuche folgen in der Abbildung 4.47. Auch hier sind die Schätzfehler wieder vom tatsächlichen Stufenabstand bzw. von der tatsächlichen Horizontspreizung abhängig. Diese beiden Größen im Versuch exakt einzustellen ist nicht möglich. Es werden daher die Ergebnisse wieder für „kleine“, „mittlere“ sowie „große“ Stufenabstände bzw. Horizontspreizungen dargestellt. Es ist dann möglich, die experimentell ermittelten Werte um ihren systematischen Fehler zu korrigieren und eine einfache Messunsicherheit anzugeben. Die verschiedenen Auswertemethoden lassen sich anhand der Ergebnisse zwar nicht vergleichen, allerdings kann anhand der Größe der ermittelten Fehler abgewogen werden, welche Versuchs- und Auswertemethode geeignet ist, Vertrauen in die Ergebnisse zu erzeugen. So erscheint es wenig sinnvoll, anhand einer Versuchsreihe mit $n = 15$ Einzelversuchen und einer Auswertung nach DIXON/MOOD das 2,5 %-Quantil schätzen zu wollen. Hierfür würde zunächst mit $x_{2,5\%} = m - 1,96 \cdot s$ das 2,5 %-Quantil berechnet. Unter der Annahme eines mittleren Stufenabstandes wäre hierfür dann ein zufälliger Messfehler von $C_x > 3 \cdot s$ anzusetzen, vgl. Abbildung 4.46 d). Damit wäre das Ergebnis der Versuchsreihe durch $x_{2,5\%} = m - 1,96 \cdot s \pm 3 \cdot s$ anzugeben.

4.6 Schätzung des 2,5 %-Quantils bei unbekannter Standardabweichung

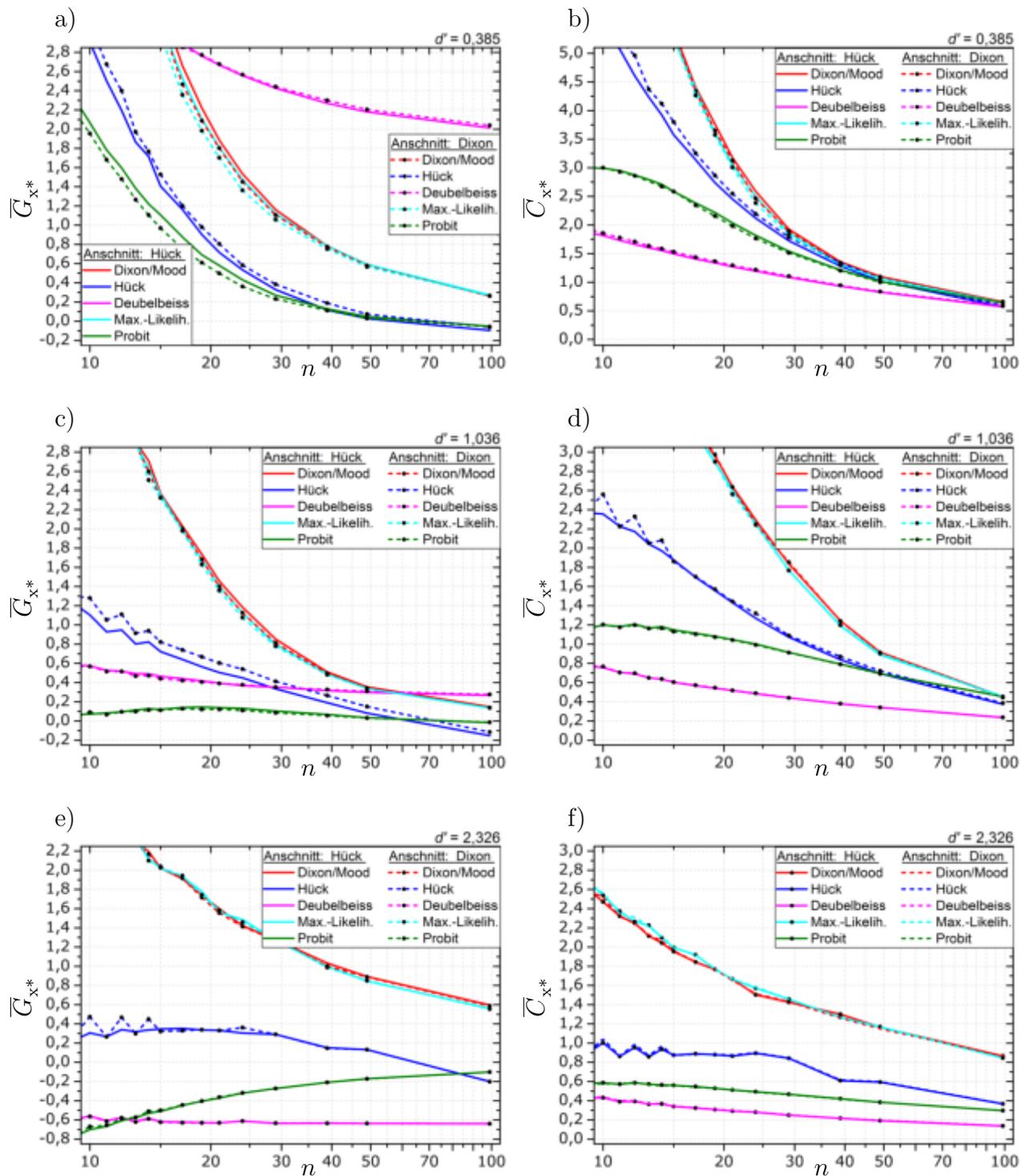


Abb. 4.46: Mittlere Verzerrungen \bar{G}_{x^*} und Standardfehler \bar{C}_{x^*} des Schätzers x^* für verschiedene bezogene Stufenabstände d' .

4 Vergleich der Simulationsergebnisse

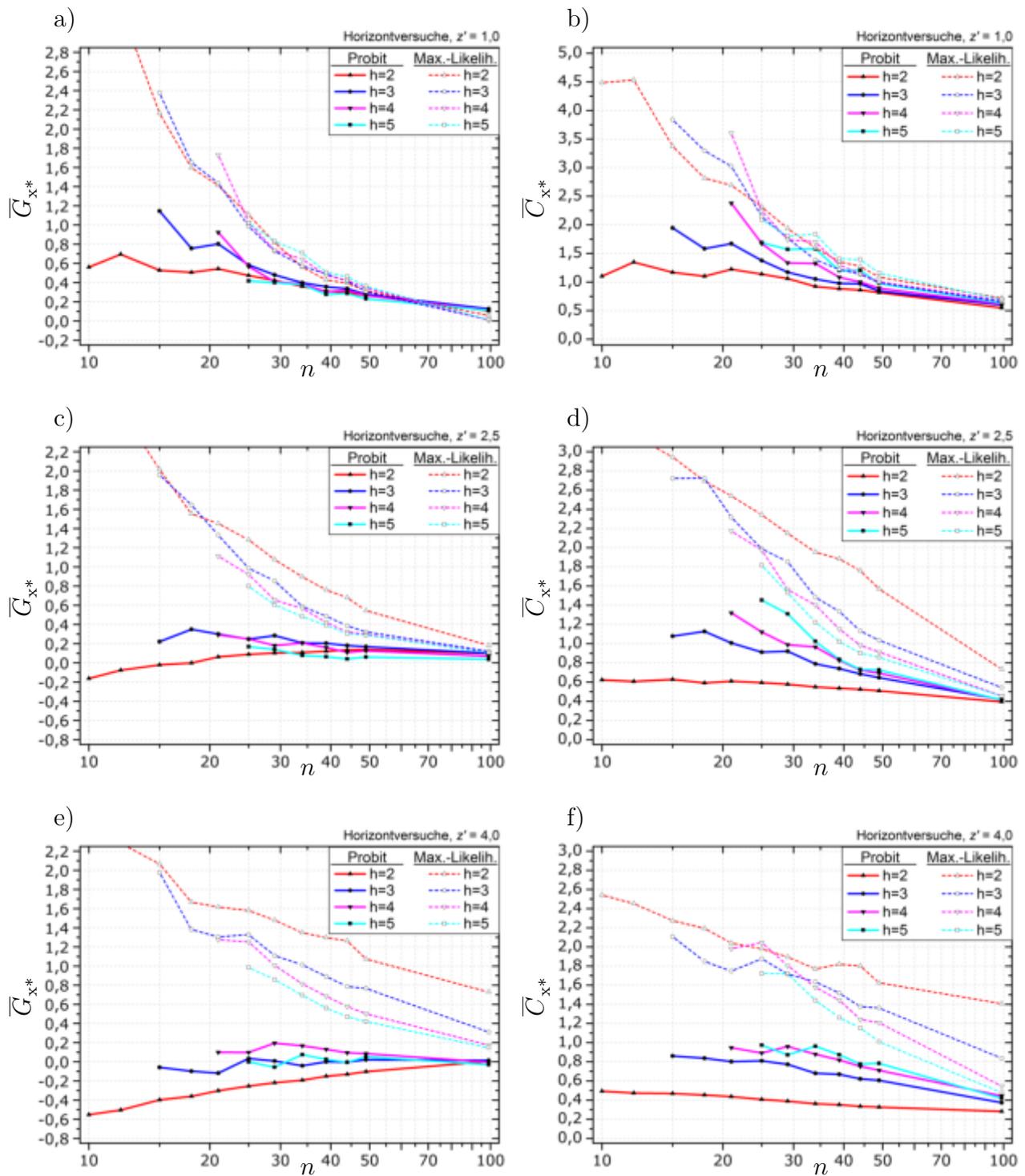


Abb. 4.47: Mittlere Verzerrungen \bar{G}_{x^*} des Schätzers x^* für verschiedene bezogene Horizontspreizungen z' .

4.6.2 RMS-Fehler

Der RMS-Fehler wird als die Quadratwurzel der mittleren quadratischen Abstände zwischen den geschätzten und den wahren Werten des 2,5 %-Quantils berechnet. Mit den nach Abschnitt 3.4.6.2 ab S. 93 berechneten mittleren RMS-Fehlern des Schätzers x^* kann als Messergebnis das 2,5 %-Quantil einschließlich des RMS-Fehler durch

$$x_{2,5\%_k} = x_{2,5\%} \pm \overline{RMS}_{x^*} \cdot s \quad (4.3)$$

angegeben werden. Die berechneten mittleren RMS-Fehler des Schätzers x^* sind in der Abbildung 4.48 festgehalten. Wie aus der grafischen Darstellung der Ergebnisse ersichtlich wird, ergeben sich teilweise sehr große RMS-Fehler. Sie zeigen deutlich, wie schwierig die Bewertung der Unsicherheit des ermittelten des 2,5 %-Quantils ist, wenn die Standardabweichung der Grundgesamtheit im selben Versuch geschätzt wird.

4 Vergleich der Simulationsergebnisse

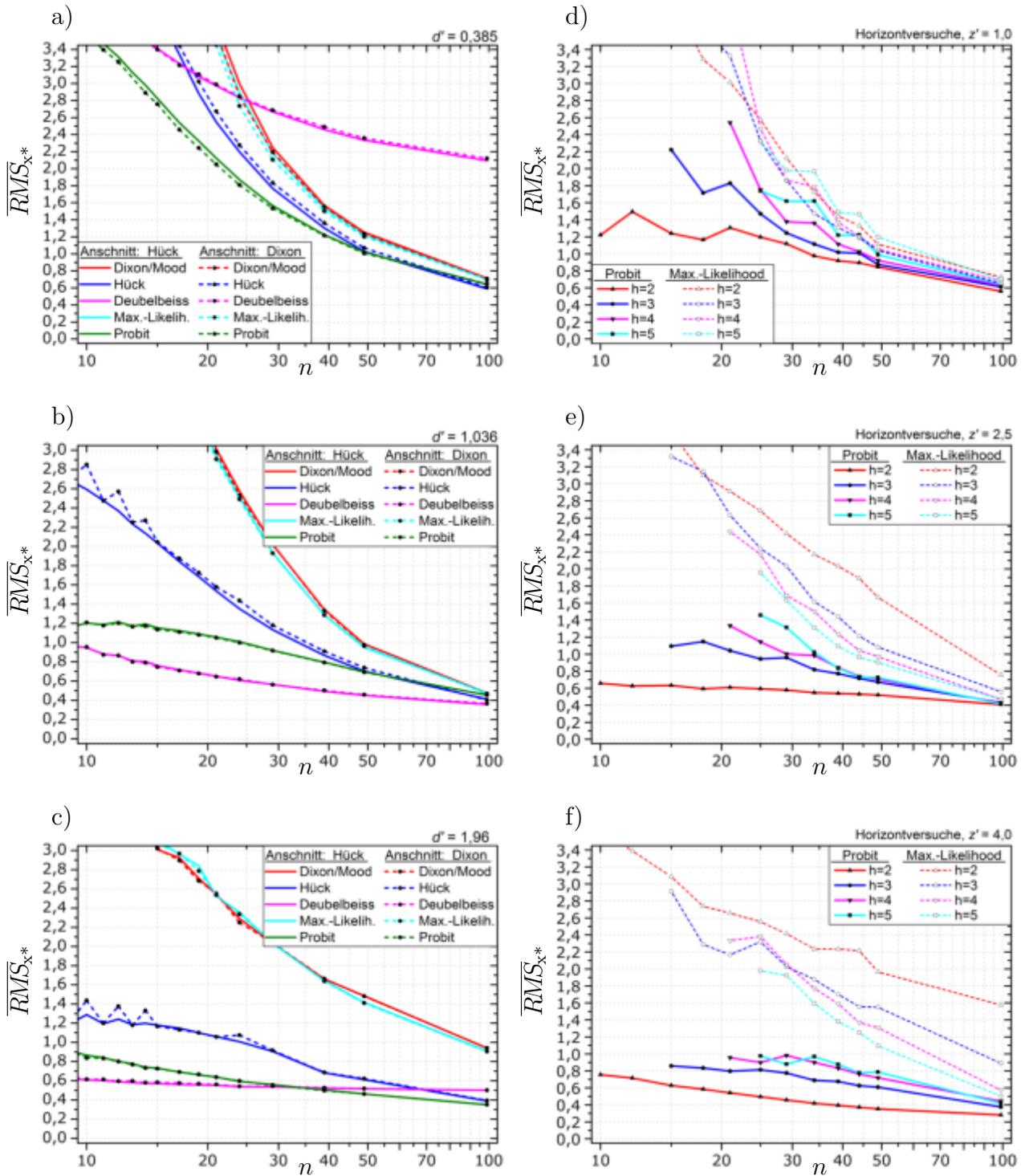


Abb. 4.48: Mittlere RMS-Fehler \overline{RMS}_{x^*} des Schätzers x^* für verschiedene bezogene Stufenabstände d' .

4.6.3 Empirischer Vertrauensbereich des geschätzten 2,5 %-Quantils

Die mittleren 95 %-Quantile des Schätzers x^* geben die zu erwartenden maximalen Überschätzungen des 2,5 %-Quantils zum Niveau $1 - \alpha = 0,95$ als Vielfaches der im selben Versuch geschätzten Standardabweichung an. Die untere Vertrauensgrenze für das tatsächliche 2,5 %-Quantil $X_{2,5\%,G}$ zum Niveau 0,95 wird durch

$$X_{2,5\%,G} \leq x_{2,5\%} - \bar{x}_{95\%}^* \cdot s$$

beschrieben. Die Ergebnisse für die mittleren 95 %-Quantile des Schätzers x^* aus Horizontversuchen und Treppenstufenversuchen sind in den Abbildungen 4.49 und 4.50 gegenübergestellt. Die Ergebnisse eignen sich nicht zum Vergleich der Verfahren, sondern nur zur Ermittlung der anzusetzenden mittleren maximalen Schätzfehler.

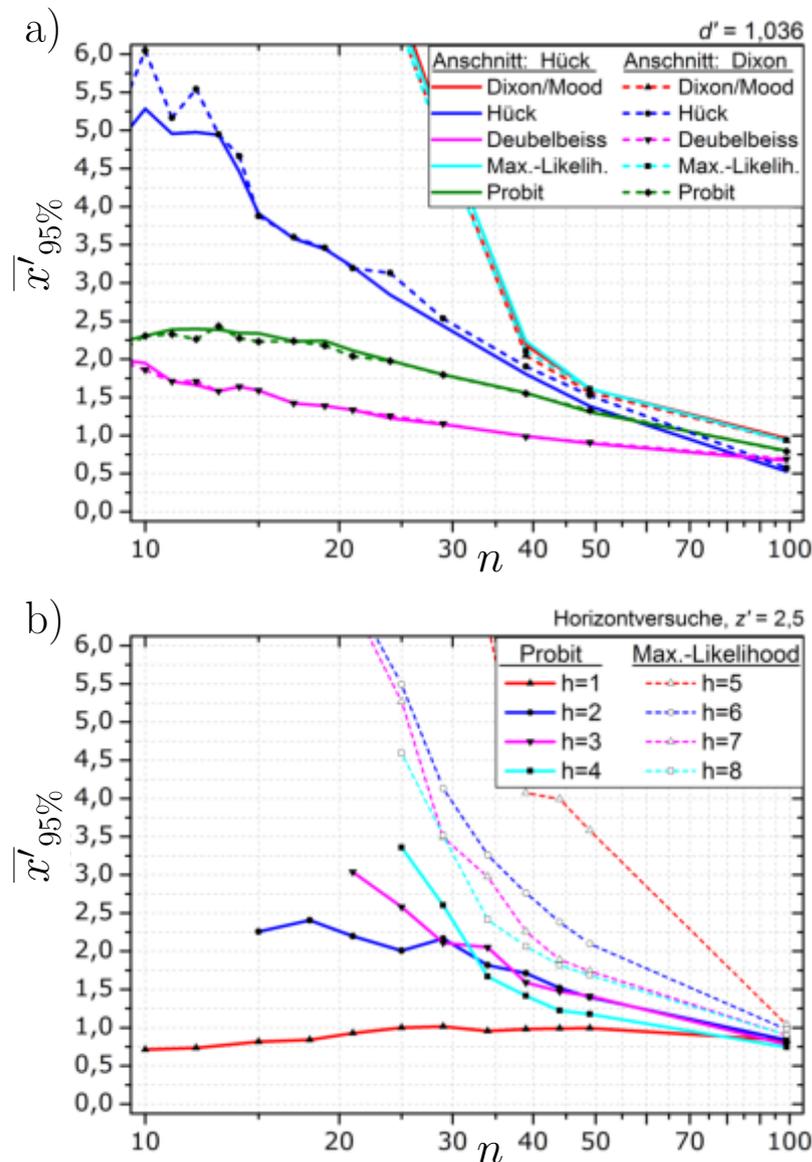


Abb. 4.49: Mittlere 95 %-Quantile $\bar{x}'_{95\%}$ des Schätzers x' für Treppenstufenversuche in a) und Horizontversuche in b) bei mittlerem Stufenabstand bzw. mittlerer Horizontspreizung.

4 Vergleich der Simulationsergebnisse

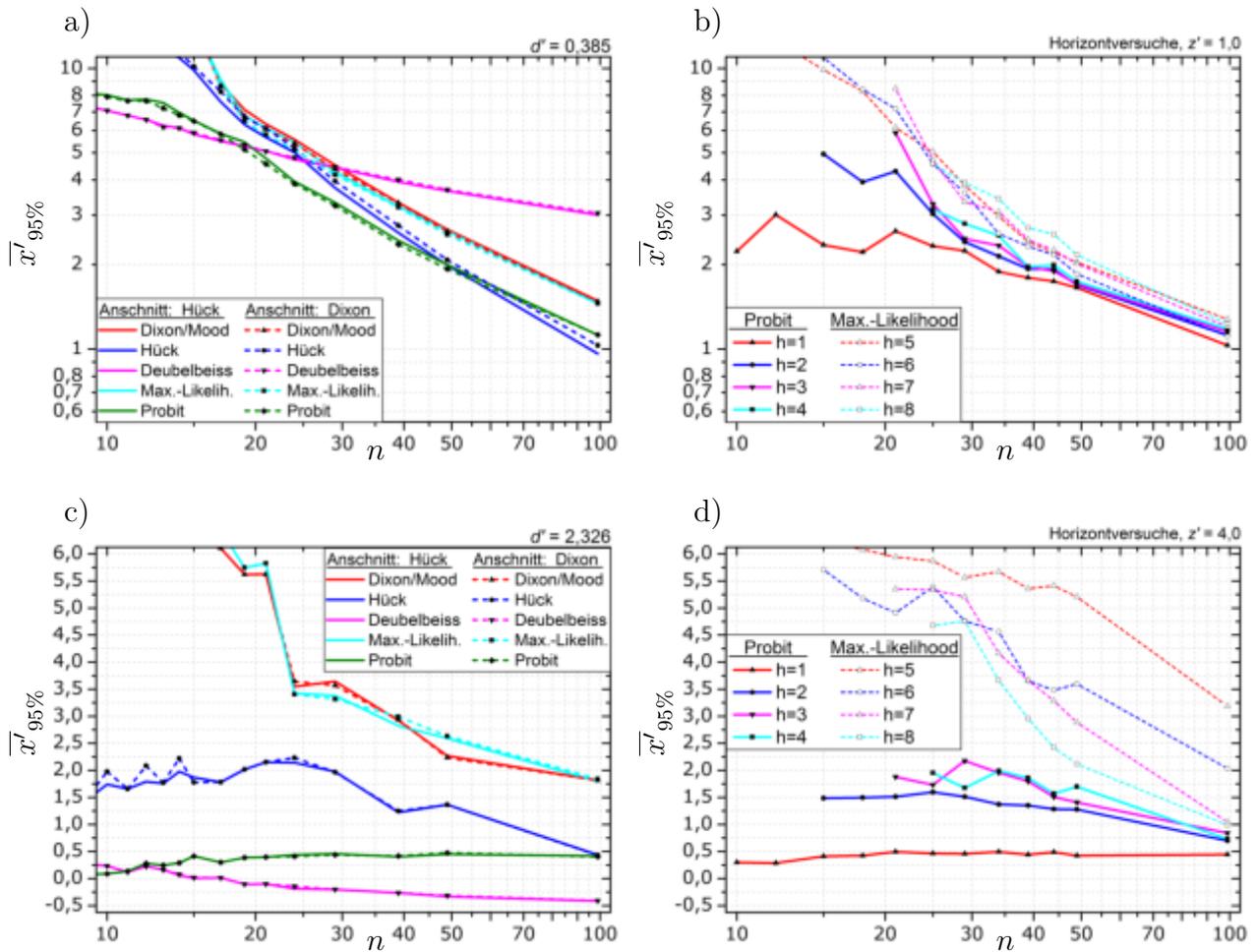


Abb. 4.50: Mittlere 95%-Quantile $\bar{x}'_{95\%}$ des Schätzers x' für Treppenstufenversuche und Horizontversuche. Für kleine Stufenabstände und Horizontspreizungen in a) und b) sowie für große Stufenabstände und Horizontspreizungen in c) und d).

5 Zusammenfassung der Simulationsergebnisse

5.1 Bewertung der Auswerteverfahren

Untersucht wurde die Treffsicherheit, mit der die Parameter einer normalverteilten Dauerfestigkeit oder Langzeitfestigkeit ermittelt werden können. Deren Bestimmung kann mit Horizontversuchen oder mit Treppenstufenversuchen erfolgen. Je nach Art der Versuchsführung stehen dann verschiedene Auswertemethoden zur Verfügung. **Horizontversuche** können mit der PROBIT-METHODE oder der MAXIMUM-LIKELIHOOD-METHODE ausgewertet werden. Für **Treppenstufenversuche** finden die Auswertemethoden nach DIXON/MOOD, HÜCK oder DEUBELBEISS Verwendung. Natürlich ist es auch möglich, Treppenstufenversuche mittels der PROBIT-METHODE oder der MAX.-LIKELIHOOD-METHODE auszuwerten, wenn die Lastniveaus der Treppenstufenversuche als Lasthorizonte interpretiert werden. Ergebnis der Auswertungen sind die Schätzer m und s für den Erwartungswert und die Standardabweichung der Grundgesamtheit. Darüber hinaus wird mit $x = m - 1,96 \cdot s$ noch das Quantil für eine Überlebenswahrscheinlichkeit von $P_{\bar{0}} = 97,5\%$ ermittelt. Die Treffsicherheit dieser drei Größen wird für die einzelnen Auswerteverfahren durch die Auswertung des Fehlers zwischen den aus Simulationsversuchen geschätzten und den vorgegebenen tatsächlichen Parametern der Grundgesamtheit abgeleitet. Die Größe des Fehlers ist dabei immer von der Größe der Streuung der Grundgesamtheit abhängig. Alle Ergebnisse werden daher auf die Streuung der Grundgesamtheit normiert. Es entstehen die bezogenen Schätzer m' , s' und x' , vgl. Tabelle 3.3 auf S. 61. Diese beschreiben alle den Schätzfehler in Vielfachen der tatsächlichen Standardabweichung σ_G der Grundgesamtheit. Sie besitzen also alle die selbe Bezugsgröße und können daher für einen Vergleich der verschiedenen Auswerteverfahren herangezogen werden. Für diesen Vergleich werden für die bezogenen Schätzer verschiedene Fehlergrößen verwendet. Der Erwartungswert der bezogenen Schätzer beschreibt die systematische Abweichung. Die Standardabweichung gibt den zufälligen Anteil des Schätzfehlers an. Mit dem RMS-Fehler kann die mittlere Abweichung einer einzelnen Schätzung beurteilt werden und mit den ermittelten Quantilen der bezogenen Schätzer lassen sich die auftretenden minimalen oder maximalen Schätzfehler beurteilen. Je nach betrachteter Fehlergröße ergeben sich bei einem Vergleich der Auswerteverfahren unterschiedliche Einschätzungen über deren Treffsicherheit. Eine qualitative Zusammenfassung dieser Betrachtung ist in Tabelle 5.1 finden.

Tabelle 5.1: Qualitativer Vergleich der Schätzmethoden je nach verwendeter Fehlergröße. Einteilung in geeignete Methoden mit dem geringsten Fehler (★★) und etwas größerem Fehler (★) sowie weniger geeignete Methoden (●) und keinesfalls geeignete Methoden (✗).

Versuchs- führung	Auswerte- methode	zu schätzende Größe und Schätzfehler											
		Erwartungswert				Standardabweichung				Quantil			
		\overline{G}_m	\overline{C}_m	\overline{RMS}_m	$\overline{m}'_{ 5\% }$	$\overline{G}_{s',lg}$	$\overline{C}_{s',lg}$	$\overline{RMS}_{s',lg}$	$\overline{s}'_{ lg,5\% }$	\overline{G}_x	\overline{C}_x	\overline{RMS}_x	$\overline{x}'_{95\%}$
Treppen- stufen	DIXON/MOOD	★	★	★	★	●	★ ²	●	●	★	★ ²	★★ ²	●
	HÜCK	★ ¹	★★	★★	★	★★	★ ³	★ ³	★ ³	★★ ³	★ ^{3,5}	★ ^{3,5}	★
	DEUBELBEISS	★ ¹	★★	★★	★	✗	★★	★★ ⁴	★ ⁴	✗	★★	★★ ⁴	★
	MAX.-LIKELI.H.	★★	★★	★★	★★	●	★ ²	●	●	★	★ ²	★★ ²	●
	PROBIT	★★	●	●	★	★★	★ ³	★	★★	★	★ ^{3,5}	★ ^{3,5}	★★
Hori- zonte	MAX.-LIKELI.H.	●	●	●	●	★	★ ^{7,8}	★	★	★★ ⁷	★ ^{7,8}	★ ^{7,8}	●
	PROBIT	●	●	●	●	★★	★★ ^{6,8}	★★ ⁶	★★	★★ ⁶	★★ ^{6,8}	★★ ^{6,8}	★★

¹ Anschnitt nach DIXON, ² kleiner Stufenabstand, ³ großer Stufenabstand, ⁴ $n \leq 20$ Einzelversuche

⁵ $n \geq 20$ Einzelversuche, ⁶ $h \leq 3$ Horizonte, ⁷ $h \geq 3$ Horizonten, ⁸ bei großer Horizontspreizung

5.1.1 Schätzung des Erwartungswertes

An die Schätzer der Parameter können, wie im Abschnitt 2.1.3 auf Seite 22 aufgeführt, verschiedene Anforderungen gestellt werden. Steht bei der Schätzung die Erwartungstreue im Vordergrund, dann ist der Erwartungswert der Grundgesamtheit aus Treppenstufenversuchen mittels der MAX.-LIKELIHOOD-METHODE zu schätzen. Mit dieser Methode fällt der aus einer etwaigen Startwertverschiebung resultierende mittlere systematische Schätzfehler am geringsten aus. Zwar lassen sich auch mit Horizontverfahren nahezu erwartungstreue Schätzer finden, allerdings weisen diese eine deutlich größere Varianz bei kleineren Stichprobenumfängen auf als alle Treppenstufenauswertemethoden und sind daher als weniger effizient zu bezeichnen. Im Rahmen der hier untersuchten Stichprobenumfänge sind alle betrachteten Schätzer des Erwartungswertes konsistent. Wird für die Bewertung der Schätzgüte der RMS-Fehler verwendet, zeigen sich alle Treppenstufenauswertemethoden außer der PROBIT-METHODE etwa gleich treffsicher. Die Auswertung der Horizontversuche führt bei der Schätzung des Erwartungswertes immer zu größeren RMS-Fehlern, als das bei der Auswertung der Treppenstufenversuche der Fall ist. Anhand des ermittelten Vertrauensbereichs zum Niveau $1 - \alpha = 0,9$ können die größten und kleinsten Verschätzungen des Erwartungswertes beurteilt werden. Auch hier sind mit den Treppenstufenversuchen die größten Treffsicherheiten zu erzielen. Vor allem bei sehr kleinen Stufenabständen treten hier bei einer Auswertung mittels der MAX.-LIKELIHOOD-METHODE die kleinsten Schätzfehler auf.

Die Wahl der Anschnittsdefinition hat nur bei den Auswertungen nach HÜCK und DEUBELBEISS einen beachtenswerten Einfluss. Bei beiden Verfahren kann die Erwartungstreue der Schätzung verbessert werden, wenn die Anschnittsdefinition nach DIXON Verwendung findet. Der Grund für die bessere Schätzung ist, dass nach dieser Definition mehr Versuche dem Anschnitt der Versuchsreihe zugeordnet werden und somit insgesamt mehr Proben für das Erreichen einer vorgegebenen Anzahl an gültigen Einzelversuchen nötig sind.

Der Einfluss einer Startwertverschiebung, also der Abweichung des Versuchsreihenstartwertes

vom tatsächlichen Erwartungswert der Grundgesamtheit, zeigt sich vor allem bei den Auswertungen der Treppenstufenversuche nach HÜCK und DEUBELBEISS. Die Güte der Schätzung ist davon abhängig ob das Niveau des tatsächlichen Erwartungswertes der Festigkeit zufällig genau mit einem Lastniveau der Treppenstufen zusammen fällt oder genau zwischen zwei Treppenstufenniveaus liegt. Die Simulationen zeigen aber, dass die sich ergebenden Schätzfehler bei festgehaltenem Stufenabstand und variierender Startwertverschiebung nicht unbegrenzt anwachsen und daher die oben dargestellten mittleren Schätzfehler gebildet werden können. Eine systematische Verschätzung des Erwartungswertes ist vor allem dann zu berücksichtigen, wenn in praktischen Versuchen aus Gründen der Zeitersparnis die Versuchsreihen systematisch auf einem hohen Lastniveau, auf dem nur Brüche erwartet werden, begonnen werden.

5.1.2 Schätzung der Standardabweichung

Bei der Schätzung der Standardabweichung der Grundgesamtheit zeigen alle Auswertemethoden der Horizontversuche und der Treppenstufenversuche bis auf die Auswertung nach DEUBELBEISS mindesten asymptotisch mediantreues Verhalten. Die Auswertungen der Treppenstufenversuche nach DIXON/MOOD und mittels MAX.-LIKELIHOOD-METHODE liefern stets die gleichen mittleren Schätzfehler. Beide Methoden unterschätzen bei kleinen Stichprobenumfängen die Standardabweichung systematisch. Die Auswertungen nach HÜCK und mittels PROBIT-METHODE überschätzen die Standardabweichung im Mittel leicht. Die Auswertung nach HÜCK liefert hier meist nahezu mediantreue Schätzungen. Die Auswertungen der Horizontversuche zeigen ähnliche systematische Schätzfehler, wie sie auch mit den gleichen Auswerteverfahren bei Treppenstufenversuche beobachtet werden. Die Schätzungen mittels PROBIT-METHODE überschätzen die Standardabweichungen in geringerem Maße als die Auswertung mittels MAX.-LIKELIHOOD-METHODE die Standardabweichungen unterschätzen. Auch die Standardfehler der Schätzer fallen bei einer Auswertung der Horizontversuche mittels PROBIT-METHODE stets kleiner aus als bei einer Auswertung mittels MAX.-LIKELIHOOD-METHODE. Bei der Auswertung der Treppenstufenversuche sind nur bei sehr kleinen Stufenabständen mit den Auswertungen nach DIXON/MOOD oder mittels MAX.-LIKELIHOOD-METHODE kleine Standardfehler zu beobachten. Bei Stufenabständen $d' > 1$ sind mit den anderen Auswertemethoden geringere Standardfehler zu erreichen. Das spiegelt sich auch in den berechneten mittleren RMS-Fehlern oder maximalen Unterschätzungen zum Niveau $1 - \alpha = 0,95$ der Standardabweichung wieder. Mit den Auswertungen nach HÜCK oder der PROBIT-METHODE treten hier deutlich geringere Schätzfehler auf. Schätzungen aus Horizontversuchen sind nur bei einer Auswertung mittels PROBIT-METHODE und nur für Versuchsreihen mit $h \leq 3$ Horizonten treffsicherer als Auswertungen von Treppenstufenversuchen. Schätzungen der Standardabweichung sollten daher entweder bei Treppenstufenversuchen mit der Auswertemethode nach HÜCK oder der PROBIT-METHODE erfolgen oder bei Horizontversuchen mit zwei oder drei Horizonten und der PROBIT-METHODE durchgeführt werden.

Eine Sonderrolle bei der Schätzung der Standardabweichung nimmt die Auswertemethode nach DEUBELBEISS ein. Mit ihr ist es nur für einen sehr begrenzten Bereich vorgegebener Stufenabstände von $1 \leq d' \leq 1,4$ möglich, erwartungstreue Schätzungen der Standardabweichung zu erreichen. Die Schätzungen sind also im Allgemeinen nicht mediantreu oder konsistent. Allerdings sind die Standardfehler der Schätzungen sehr gering. Dies führt bei

geringen Versuchszahlen zu durchweg geringeren RMS-Fehlern. Auch die mittleren maximalen Unterschätzungen zum Niveau $1 - \alpha = 0,95$ fallen hier geringer als mit allen anderen Auswerteverfahren aus. Daher können die Auswertung nach DEUBELBEISS bei Stichprobengrößen $n \leq 20$ durchaus zur Schätzung der Standardabweichung geeignet sein.

5.1.3 Schätzung des 2,5 %-Quantils

Die Güte der Standardabweichungsschätzung dominiert die Güte der Schätzung des 2,5 %-Quantils der Grundgesamtheit. Das Quantil wird bei der Auswertung von Treppenstufenversuchen nach DIXON/MOOD sowie mittels der MAX.-LIKELIHOOD-METHODE im Mittel überschätzt. Beide Verfahren schätzen das 2,5 %-Quantil asymptotisch erwartungstreu. Mit der PROBIT-METHODE wird das 2,5 %-Quantil ebenso asymptotisch erwartungstreu aber im Mittel zu gering, d. h. also konservativ, geschätzt. Auch die Auswertung nach HÜCK liefert im Mittel konservative Schätzungen des Quantils. Bei Stufenabständen von $d' \geq 1$ und Stichprobenumfängen von $15 \leq n \leq 50$ sind diese nahezu erwartungstreu. Damit auch die Auswertungen der Horizontversuche erwartungstreue Schätzer liefern, dürfen die Horizontspreizungen nicht zu groß oder zu klein gewählt werden. Mit der MAX.-LIKELIHOOD-METHODE wird das Quantil systematisch über-, mit der PROBIT-METHODE systematisch unterschätzt. Beide Verfahren liefern asymptotisch erwartungstreue Schätzungen des 2,5 %-Quantils. Die kleinsten Standardfehler zeigen sich bei den Auswertungen der Treppenstufenversuche nach DIXON/MOOD oder mittels der MAX.-LIKELIHOOD-METHODE. Auch mit Horizontversuchen werden keine Schätzer mit kleineren Standardfehlern erreicht.¹ Das führt dazu, dass für die Auswertungen der Treppenstufenversuche nach DIXON/MOOD oder mit der MAX.-LIKELIHOOD-METHODE auch die kleinsten RMS-Fehler bestimmt werden. Bei Horizontversuchen sind solche kleinen RMS-Fehler nur bei einer Auswertung von Versuchsreihen mit zwei oder drei Horizonten und der PROBIT-METHODE oder mit Versuchsreihen mit $n \geq 100$ Einzelversuchen möglich. Allerdings liefert das Abgrenzungsverfahren hier die geringsten RMS-Fehler.

Obwohl die Standardfehler und RMS-Fehler bei einer Auswertung der Treppenstufenversuche mittels MAX.-LIKELIHOOD-METHODE bzw. nach DIXON/MOOD die geringsten Werte annehmen, sind die mittleren maximalen Schätzfehler zum Niveau $1 - \alpha = 0,95$ dieser beiden Verfahren deutlich größer als mit den anderen Treppenstufenauswertungen.² Treppenstufenversuche sollten daher mit der Methode nach HÜCK oder der PROBIT-METHODE ausgewertet werden. Bei Horizontversuchen sollten möglichst viele Versuche auf wenige Horizonte verteilt und nach der PROBIT-METHODE ausgewertet werden. Das Abgrenzungsverfahren ist im Rahmen der hier betrachteten Simulationsergebnisse das genaueste Verfahren, um das 2,5 %-Quantil zu schätzen. Alle anderen Horizontauswertungen bringen zumindest bei den betrachteten Stichprobenumfängen keinen hervorzuhebenden Vorteil.

¹Ausgenommen davon ist das Abgrenzungsverfahren bei sehr großen Horizontspreizungen, mit dem aber wieder größere systematische Schätzfehler einhergehen.

²Der Grund hierfür ist zum einen die ohnehin vorhandene systematische Überschätzung, zum anderen neigen die PROBIT-METHODE wie auch die Auswertung nach HÜCK dazu, die Quantile stärker zu unterschätzen als zu überschätzen. Die statistischen Verteilungen der Schätzer sind nicht symmetrisch und durch die Angabe eines Standardfehlers nur bedingt zu beschreiben.

5.2 Vorgehen zur experimentellen Bestimmung der Verteilungsparameter

5.2.1 Bei bekannter Standardabweichung

Ist die Streuung σ_G der Festigkeit bereits bekannt, oder wird sie als bekannt vorausgesetzt, ist nur der Erwartungswert der Festigkeit zu ermitteln. Mittels Treppenstufenversuchen kann dieser mit dem kleinsten Fehler ermittelt werden, wenn der Stufenabstand der Treppenstufenversuche mit $0,5 \leq d' \leq 1,0$ möglichst klein gewählt wird. Bei einer Auswertung der Versuche nach der MAX.-LIKELIHOOD-METHODE ergeben sich außerdem über alle Fehlerarten hinweg die treffsichersten Schätzungen. Diese Auswertemethode sollte daher bevorzugt Anwendung finden.

Die Unsicherheit der Erwartungswertschätzung kann mithilfe des RMS-Fehlers nach Abbildung 5.1 und der tatsächlichen Standardabweichung σ_G ermittelt werden.³ Das vollständige Ergebnis der Erwartungswertschätzung setzt sich aus der Angabe des geschätzten Erwartungswertes sowie der Unsicherheit der Schätzung zusammen und lautet:

$$m_k = m \pm \overline{RMS}_{m'} \cdot \sigma_G .$$

Für gekerbte Stahlbauteile liegt die Größe der Festigkeitsstreuung im Bereich von etwa 5 % bis 10 % der mittleren Festigkeit, vgl. [1] und [19].⁴ Bei Versuchsreihen mit $n = 20$ gültigen Einzelversuchen sind also Unsicherheiten bei der Bestimmung der mittleren Festigkeit von 1,5 % bis 3 % der mittleren Festigkeiten zu erreichen. Auch bei Versuchsreihen mit nur $n = 10$ gültigen Einzelversuchen ergeben sich Schätzungen mit ausreichend geringen Unsicherheiten von 2 % bis 4 % der mittleren Festigkeit. Zu bedenken bleibt aber, dass die dargestellten Fehler für die Anzahl der gültigen Einzelversuche dargestellt sind. Wird der Versuchsreihenstartwert weit entfernt von der tatsächlichen mittleren Festigkeit platziert, so ist vor allem bei kleinen Stufenabständen eine deutlich größere Probenzahl für die Erstellung einer gültigen Treppenstufenversuchsreihe nötig. Die Unsicherheiten bei der Schätzung von Quantilen der Festigkeitsverteilung fallen bei bekannter Standardabweichung genauso groß wie die Unsicherheiten der mittleren Festigkeit aus.

³Die MAX.-LIKELIHOOD-METHODE schätzt den Erwartungswert der Festigkeit recht erwartungstreu. Der RMS-Fehler $\overline{RMS}_{m'}$ unterscheidet sich kaum vom Standardfehler \overline{C}_m bzw. vom verallgemeinerten Standardfehler $U_{m'}$ des Schätzers m' .

⁴Die Anzahl an Literaturangaben zu mittleren Streuungen der Festigkeit von Bauteilen bei hohen Schwingungszahlen ist sehr gering. In [1] sind hierzu Angaben für stark gekerbte Bauteile (Zahnräder Schrauben und Schweißverbindungen) und in [19] für Probestäbe zu finden. Die Streuungen von Bauteilen sind dabei größer als die von Probestäben oder Werkstoffproben. In [19] wird gezeigt, dass die Unterschiede der ermittelten Streuung zwischen Bauteilen und daraus entnommenen Werkstoffproben den Faktor 4 erreichen können

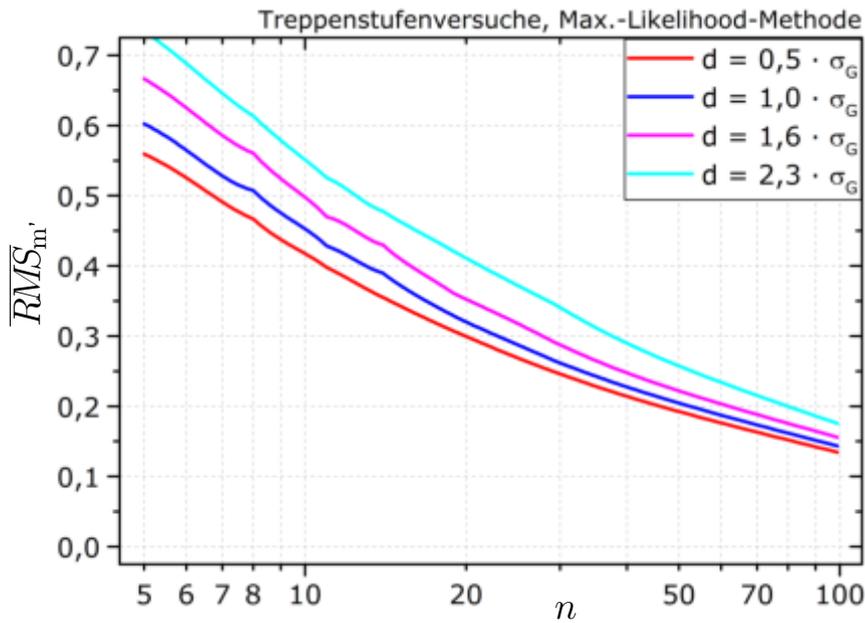


Abb. 5.1: RMS-Fehler der Erwartungswertschätzung für die Auswertung von Treppenstufenversuchen mittels der MAX.-LIKELIHOOD-METHODE bei bekannter Standardabweichung σ_G .

5.2.2 Bei unbekannter Standardabweichung

Ist die Standardabweichung nicht bekannt, muss für die Versuchsplanung zunächst eine Standardabweichung der Festigkeit geschätzt oder geraten werden. Literaturangaben zur Streuung der Festigkeit im Bereich großer Schwingspielzahlen sind rar, Hinweise für Stahlwerkstoffe sind in [1] oder [19] zu finden. Auch Angaben zur Streuung der Festigkeit im Zeitfestigkeitsbereich [14] können für eine erste Schätzung herangezogen werden. Andernfalls ist eine Streuung zwischen 5% und 10% des Versuchsreihenstartwertes als Festigkeitsstreuung anzunehmen.

Im Falle einer unbekannt Standardabweichung, verändert sich nicht die Genauigkeit der Erwartungswertschätzung, sondern nur die subjektive Einschätzung der Unsicherheit fällt je nach verwendetem Verfahren anders aus. Die Auswertung der Treppenstufenversuche mittels der MAX.-LIKELIHOOD-METHODE liefert aber weiterhin objektiv die treffsichersten Ergebnisse und sollte daher verwendet werden, wenn ausschließlich die mittlere Festigkeit zu ermitteln ist. Die Auswertungen nach HÜCK sowie DEUBELBEISS liefern nur unwesentlich größere Schätzfehler. Mit dem RMS-Fehler \overline{RMS}_{m^*} aus Abbildung 5.2 kann die Unsicherheit der Schätzung näherungsweise bestimmt werden. Das vollständige Ergebnis der Erwartungswertschätzung setzt sich aus der Angabe des geschätzten Erwartungswertes sowie der Unsicherheit der Schätzung zusammen und lautet:

$$m_k = m \pm \overline{RMS}_{m^*} \cdot s.$$

Da bei einer unbekannt Standardabweichung auch der tatsächliche bezogene Stufenabstand d' unbekannt ist, können die Unsicherheiten nicht exakt angegeben werden. Meist ist nur eine Unterscheidung in Versuchsreihen mit einem kleinen Stufenabstand ($d' < 1$), einem mittleren ($d' \approx 1$) oder einem großen Stufenabstand ($d' > 1$) möglich. Daher sind in Abbildung 5.2 die Ergebnisse für drei vorgegeben Stufenabstände $d \approx 0,5$, $d \approx 1,0$, $d \approx 2,0$ dargestellt. In der praktischen Versuchsführung ist ein bezogener Stufenabstand von $d' = 1$ anzustreben. Weicht

5.2 Vorgehen zur experimentellen Bestimmung der Verteilungsparameter

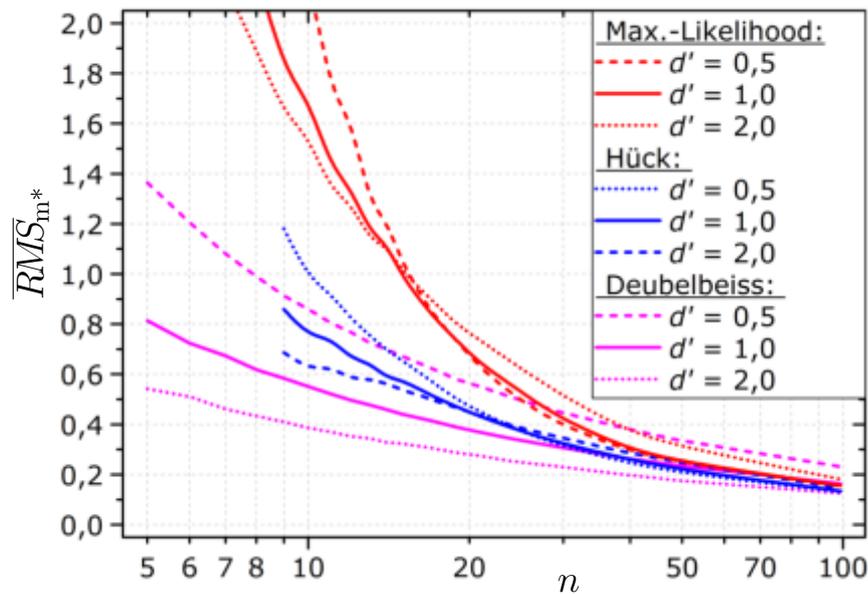


Abb. 5.2: RMS-Fehler der Erwartungswertschätzung für die Auswertung von Treppenstufenversuchen bei unbekannter Standardabweichung.

der geschätzte Stufenabstand der experimentell ermittelten Versuchsreihe stark von $d' = 1$ ab, so sind die entsprechenden Kurven für den großen oder kleinen bezogene Stufenabstand zu wählen.

Im Gegensatz zur Schätzung der mittleren Festigkeit ist die Schätzung des 2,5 %-Quantils der Festigkeitsverteilung direkt von dem Schätzwert der Standardabweichung abhängig. Hier verändern sich sowohl die Genauigkeit der Schätzung als auch die Einschätzung der Unsicherheit.

Die Ermittlung der Streuung und des 2,5 %-Quantils der Festigkeit sollte mittels Treppenstufenversuchen und der Auswertung nach HÜCK oder mit Horizontversuchen auf zwei oder drei Horizonten und der Auswertung mittels der PROBIT-METHODE erfolgen. In beiden Fällen ist ein möglichst großer Bereich der Festigkeitsverteilung mit Versuchen zu belegen. Dafür ist aus der anfänglichen grob geschätzten Standardabweichung ein Stufenabstand von $d' \approx 1,5$ bzw. eine Horizontspreizung von $z' \approx 2,5$ festzulegen. Mit dem RMS-Fehler \overline{RMS}_{x^*} aus den Abbildung 5.4 und 5.3 kann die Unsicherheit der Schätzung näherungsweise bestimmt werden. Auch hier sind zunächst mit der aus dem Versuch abgeschätzten Standardabweichung der tatsächliche Stufenabstand bzw. die Horizontspreizung zu ermitteln und je nach Ergebnis die entsprechenden Kurven aus den Abbildungen zu wählen. Das vollständige Ergebnis der Schätzung des 2,5 %-Quantils x_k setzt sich aus der Angabe des geschätzten 2,5 %-Quantils x sowie Unsicherheit der Schätzung zusammen und lautet:

$$x_k = x \pm \overline{RMS}_{x^*} \cdot s.$$

Im Falle einer begrenzten Probenanzahl ($n < 20$) kann auch eine Schätzung des 2,5 %-Quantils aus Treppenstufenversuchen und der Auswertung nach DEUBELBEISS erfolgen. Dazu ist ein ausreichend großer Stufenabstand von $d' \approx 1,5$ anzustreben. Die Schätzungen sind im mittel

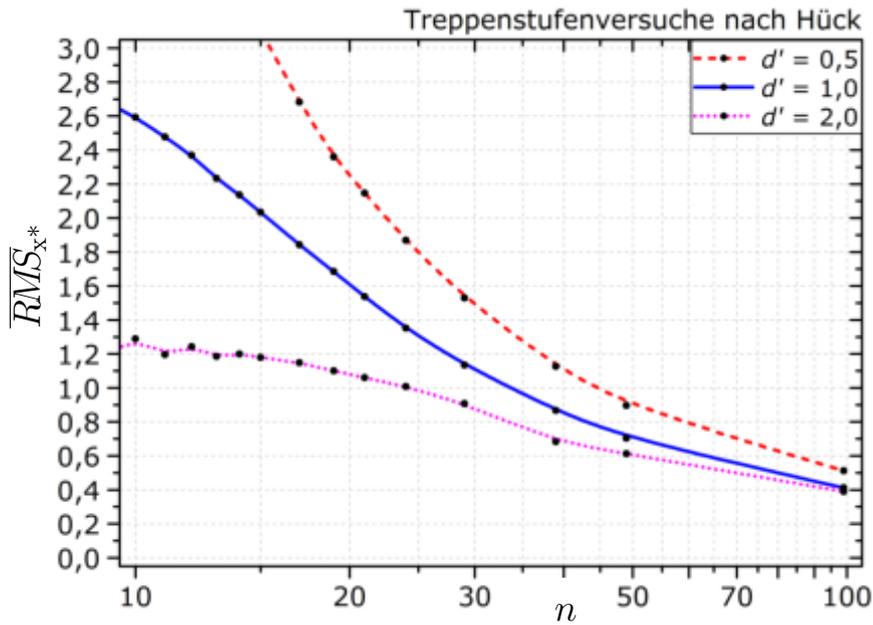


Abb. 5.3: RMS-Fehler des 2,5%-Quantils für die Auswertung von Treppenstufenversuchen bei unbekannter Standardabweichung.

stark konservativ, besitzen aber geringere RMS-Fehler als alle anderen Schätzungen. Eine Mittelung so geschätzter Standardabweichungen oder Quantile führt nicht zu treffsichereren Ergebnissen.

Stehen für die Ermittlung der Quantile hingegen sehr viele Proben ($n > 50$) zur Verfügung, sind Treppenstufenversuche bei einem Stufenabstand von $d' \approx 1$ nach der MAX.-LIKELIHOOD-METHODE auszuwerten. Die systematische Verschätzung ist dann vernachlässigbar und die Größe der zufälligen Schätzfehler verändert sich kaum, wenn der tatsächliche Stufenabstand von $d' = 1$ verschieden ist, vgl. Abschnitt 4.4.

Die Vergrößerung der Probenanzahl ist bei allen Auswerteverfahren (außer der Auswertung nach DEUBELBEISS) und für alle zu schätzenden Größen die einzige Möglichkeit, die Schätzfehler unabhängig vom Stufenabstand oder der Horizontspreizung zu verringern.

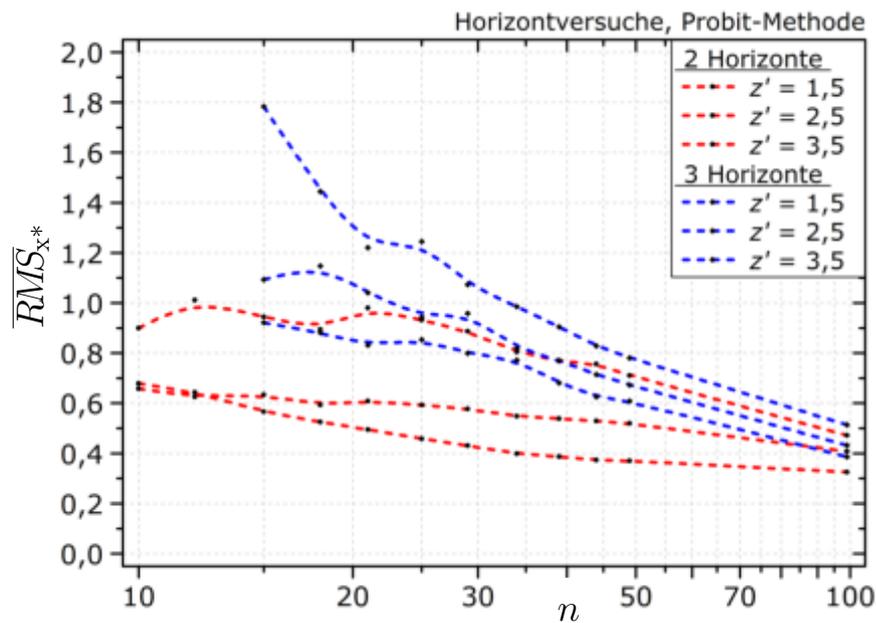


Abb. 5.4: RMS-Fehler des 2,5%-Quantils für die Auswertung von Horizontversuchen bei unbekannter Standardabweichung.

5.2.3 Anwendungsbeispiel

Mit den ermittelten Schätzfehlern kann die Genauigkeit der Ergebnisse der beispielhaften Auswertung von Horizontversuchen aus Abschnitt 2.2.1.4 ab S. 44 sowie der Treppenstufenversuche aus Abschnitt 2.2.2.5 ab S. 51 beurteilt werden.⁵ Unter der Annahme einer unbekannt Standardabweichung ist es unmöglich, den exakten Stufenabstand der Treppenstufenversuche oder die exakte Horizontspreizung der Horizontversuche anzugeben⁶. Dadurch ist es ebenfalls unmöglich, die Fehler der Standardabweichungsschätzung exakt zu bestimmen, da diese Fehler selbst jeweils vom Stufenabstand bzw. der Horizontspreizung abhängig sind. Es ist also nur möglich, den Fehler der Standardabweichungsschätzung für die geschätzten Stufenabstände oder Horizontspreizungen anzugeben. Zunächst werden die bei der Schätzung der Standardabweichung auftretenden Verzerrungen des Schätzers ermittelt, Tabelle 5.2. Mit diesen kann eine Neubewertung der Stufenabstände bzw. Horizontspreizungen erfolgen. Anschließend wird der Fehler für den bezogenen Schätzer des Erwartungswertes (Tabelle 5.3) und den bezogenen Schätzer des 2,5%-Quantils (Tabelle 5.4) bestimmt. Mit den bezogenen Schätzfehlern und der im selben Versuch ermittelten Standardabweichung können dann die in den Tabellen 5.5 und 5.6 angegebenen Schätzfehler für den Erwartungswert und das 2,5%-Quantil der Festigkeit ermittelt werden.

⁵Die den Auswertungen zugrunde liegende zufällig gezogene Stichprobe entstammt einer Grundgesamtheit mit dem Erwartungswert $\mu_G = 150$ MPa und einer Standardabweichung $\sigma_G = 10$ MPa. Der Mittelwert der Stichprobe ist $\bar{x} = 149,3$ MPa und die korrigierte Stichprobenstandardabweichung beträgt $\sqrt{s^2} = 9,54$ MPa.

⁶Für das Auswertungsbeispiel mit $\sigma_G = 10$ MPa beträgt der tatsächliche Stufenabstand $d' = 0,6$ und die tatsächliche Horizontspreizung $z' = 2$.

Tabelle 5.2: Fehler des Schätzers s'_{lg} für Auswertungsbeispiele.

Versuchs- führung	Auswerte- Methode	d/s bzw. z/s	$\bar{G}_{s',lg}$	$\bar{C}_{s',lg}$	$d/s \cdot 10^{\bar{G}_{s',lg}}$ bzw. $z/s \cdot 10^{\bar{G}_{s',lg}}$
Treppen- stufen	DIXON/MOOD	1,027	-0,1791	0,2875	0,680
	HÜCK	0,579	0,0350	0,3485	0,628
	DEUBELBEISS	1,020	-0,0639	0,0783	0,881
	MAX.-LIKELI.H.	0,995	-0,1797	0,2836	0,658
	PROBIT	0,484	0,0025	0,3120	0,487
Hori- zonte	MAX.-LIKELI.H.	3,140	-0,1485	0,2422	2,230
	PROBIT	2,378	0,0473	0,2058	2,650

Tabelle 5.3: Fehler des Schätzers m^* für Auswertungsbeispiele

Versuchs- führung	Auswerte- Methode	$d/s \cdot 10^{\bar{G}_{s',lg}}$ bzw. $z/s \cdot 10^{\bar{G}_{s',lg}}$	\bar{G}_{m^*}	\bar{C}_{m^*}	\overline{RMS}_{m^*}	$\overline{m^*}_{5\%}$	$\overline{m^*}_{95\%}$
Treppen- stufen	DIXON/MOOD	0,680	0,082	0,778	0,781	-0,990	1,389
	HÜCK	0,628	0,082	0,499	0,518	-0,586	0,937
	DEUBELBEISS	0,881	0,094	0,420	0,439	-0,589	0,792
	MAX.-LIKELIHOOD	0,658	0,040	0,764	0,764	-1,015	1,275
	PROBIT	0,487	0,032	0,488	0,490	-0,683	0,855
Hori- zonte	MAX.-LIKELIHOOD	2,230	0,158	0,811	0,385	-1,078	1,550
	PROBIT	2,650	0,091	0,444	0,457	-0,599	0,923

Tabelle 5.4: Fehler des Schätzers x^* für Auswertungsbeispiele

Versuchs- führung	Auswerte- Methode	$d/s \cdot 10^{\bar{G}_{s',lg}}$ bzw. $z/s \cdot 10^{\bar{G}_{s',lg}}$	\bar{G}_{x^*}	\bar{C}_{x^*}	\overline{RMS}_{x^*}	$\overline{x^*}_{95\%}$
Treppen- stufen	DIXON/MOOD	0,680	1,898	3,399	3,894	10,23
	HÜCK	0,628	0,665	2,197	2,295	5,53
	DEUBELBEISS	0,881	0,722	0,641	0,973	1,76
	MAX.-LIKELIHOOD	0,658	1,831	3,396	3,927	12,24
	PROBIT	0,487	0,584	2,035	2,053	4,93
Hori- zonte	MAX.-LIKELIHOOD	2,230	1,597	2,778	3,163	7,84
	PROBIT	2,650	0,207	1,076	1,085	1,96

Tabelle 5.5: Fehler der Erwartungswertschätzung bei unbekannter Standardabweichung für Auswertungsbeispiele.

Versuchsführung	Auswertemethode	Schätzung m in MPa	vollst. Messergebnis $m - s \cdot \overline{G}_{m^*} \pm s \cdot \overline{C}_{m^*}$ in MPa	Erg. mit RMS-Fehler $m \pm s \cdot \overline{RMS}_{m^*}$ in MPa	Konfidenzintervall zu $\alpha = 0,1$ $[m - s \cdot \overline{m}_{95\%}^* ; m - s \cdot \overline{m}_{5\%}^*]$ in MPa
Treppenstufen	DIXON/MOOD	147,0	146,3 ± 6,9	147,0 ± 4,6	[138,9 ; 152,8]
	HÜCK	147,0	146,2 ± 4,8	147,0 ± 5,4	[137,3 ; 153,1]
	DEUBELBEISS	147,0	146,4 ± 2,9	147,0 ± 2,6	[142,3 ; 150,5]
	MAX.-LIKELIHOOD	145,7	145,3 ± 7,0	145,7 ± 4,6	[138,0 ; 151,8]
	PROBIT	144,4	144,0 ± 6,0	144,4 ± 6,1	[133,8 ; 152,9]
Horizonte	MAX.-LIKELIHOOD	152,6	150,2 ± 7,3	151,6 ± 2,5	[141,7 ; 158,5]
	PROBIT	152,7	152,0 ± 3,4	152,7 ± 3,8	[144,9 ; 157,7]

Tabelle 5.6: Fehler der Schätzung des 2,5%-Quantils bei unbekannter Standardabweichung für Auswertungsbeispiele.

Versuchsführung	Auswertemethode	Schätzung $x_{2,5\%}$ in MPa	vollst. Messergebnis $x_{2,5\%} - s \cdot \overline{G}_{x^*} \pm s \cdot \overline{C}_{x^*}$ in MPa	Erg. mit RMS-Fehler $m \pm s \cdot \overline{RMS}_{x^*}$ in MPa	Konfidenzgrenze zu $\alpha = 0,05$ $X_{2,5\%,G} \geq x_{2,5\%} - s \cdot \overline{x}_{95\%}^*$ in MPa
Treppenstufen	DIXON/MOOD	135,6	118,9 ± 30,0	135,6 ± 22,7	75,9
	HÜCK	126,7	120,3 ± 21,0	126,7 ± 23,8	69,4
	DEUBELBEISS	135,5	130,6 ± 4,4	135,5 ± 5,7	125,2
	MAX.-LIKELIHOOD	133,9	117,2 ± 31,0	133,9 ± 23,7	60,1
	PROBIT	120,1	112,9 ± 25,1	120,1 ± 25,4	59,0
Horizonte	MAX.-LIKELIHOOD	139,1	124,8 ± 24,9	139,1 ± 20,1	89,2
	PROBIT	136,2	134,6 ± 8,1	136,2 ± 9,1	119,7

5.2.4 Wahl der Fehlergröße

In den vorherigen Abschnitten werden verschiedene Arten von Fehlern zur Bewertung der Ergebnisse von Treppenstufen- und Horizontversuchen vorgestellt. Es ist natürlich nicht ratsam, bei der Veröffentlichung eines Messergebnisses verschiedene Fehler anzugeben, da dieses Vorgehen geeignet ist, unnötige Verunsicherung hervorzurufen. Vielmehr sollten entweder der systematische und der zufällige Anteil des Schätzfehlers, der RMS-Fehler oder das Konfidenzintervall angegeben werden. Mit den ermittelten systematischen und zufälligen Fehleranteilen ist es möglich, das Messergebnis zu korrigieren und eine Unsicherheit anzugeben. Die so korrigierten Ergebnisse sind dann erwartungstreue Schätzungen. Der Mittelwert der Ergebnisse aus mehreren auf gleiche Weise wiederholten Experimenten liegt dann näher am wahren Wert als die Einzelergebnisse und ist mit einer deutlich geringeren Unsicherheit behaftet. Das ist vor allem bei der Durchführung von Ring-Versuchen oder bei der Erstellung von Versuchs-Datenbanken wichtig. Wird der RMS-Fehler als Maß der Treffsicherheit verwendet, ist es nicht in jedem Fall möglich, den Fehler durch Mittelung mehrerer experimenteller Ergebnisse zu verringern. Ein großer systematischer Fehler und ein kleiner zufälliger Fehler, wie es zum Beispiel bei der Standardabweichungsschätzung nach DEUBELBEISS zu beobachten ist, können zu einem kleinen RMS-Fehlern führen. Der Mittelwert der Schätzungen liegt aber nicht näher am wahren Wert als die Ergebnisse der einzelnen Experimente. Liegt nur eine einzelne Schätzung vor, ist dieser Umstand allerdings unbedeutend. Wird das Konfidenzintervall der Schätzung als Maß der Schätzgüte angegeben, können damit besser die größten oder kleinsten auftretenden Abweichungen eingeschätzt werden. Nicht erfasst werden dabei systematische Abweichungen. Werden die Ergebnisse mit den angegebenen systematischen Schätzfehlern korrigiert, verschieben sich natürlich auch die Grenzen des Konfidenzintervalls.

5.2.5 Anwendung bei nicht normalverteilten Grundgesamtheiten

Alle aus den Simulationen abgeschätzten Fehlergrößen gelten unter der Annahme einer normalverteilten Grundgesamtheit. Ist dies in der praktischen Anwendung nicht der Fall, sind andere Fehler zu erwarten. Für Verteilungen, die sich näherungsweise durch die Normalverteilung wiedergeben lassen, sollten diese Fehler vernachlässigbar sein. Die Unterschiede in den Verteilungsmodellen zeigen sich hier hauptsächlich in den Verteilungsenden, also bei sehr kleinen oder großen Merkmalswerten. Diese treten zum einen bei einem begrenzten Stichprobenumfang nur selten auf, zum anderen werden sie ohnehin nicht direkt beobachtet. Es wird lediglich festgestellt, ob ein Merkmalswert über oder unter einem festgelegten Wert liegt. Lässt sich die angenommene Verteilung durch Transformation auf eine Normalverteilung zurückführen, so können die vorgestellten Ergebnisse direkt verwendet werden, wenn die gesamte Auswertung mit den transformierten Werten erfolgt. Für Treppenstufenversuche werden die Beanspruchungshöhen dann so platziert, dass sich für die transformierten Beanspruchungshöhen äquidistante Stufenabstände ergeben. Ebenso wird bei Horizontversuchen die Lage der Lastniveaus für die transformierten Größen berechnet. Werden zum Beispiel Treppenstufenversuche unter der Annahme einer logarithmischen Normalverteilung durchgeführt, ist direkt mit den Logarithmen der Laststufen zu rechnen. Bei der Durchführung der Versuche sollte dann die Differenz zwischen den Logarithmen zweier benachbarter Laststufen

konstant sein.⁷ Die geschätzten Verteilungsparameter können direkt mit den vorgestellten Schätzfehlern angegeben werden, bevor das gesamte Messergebnis durch Potenzieren wieder zurück transformiert wird.

5.3 Kritische Betrachtung zu Simulationsversuchen

Mit Monte-Carlo-Simulationen ist es möglich, komplizierte Sachverhalten durch „einfaches Ausprobieren“ nachzuvollziehen. Es ist möglich, die Auswirkungen von Unsicherheiten der Eingangsgrößen auf die Unsicherheit von Ausgangsgrößen eines Modells oder Prozesses abzubilden, ohne die genauen Wechselwirkungen zwischen den Eingangsgrößen und den Ausgangsgrößen zu kennen. Problematisch ist dabei die genaue Beschreibung der Eingangsgrößen, da hier oft einige nicht zu unterschätzende Annahmen getroffen werden. Im Rahmen der hier vorgestellten Simulationsversuche wird für die Verteilung der Bauteilfestigkeit eine Normalverteilung gewählt. Diese Verteilung ist nach beiden Enden offen und lässt damit auch beliebig kleine Festigkeiten zu. Bei einer vorhandenen Qualitätskontrolle werden solche Bauteile oder Proben aber wahrscheinlich nicht bei einer Prüfung zum Einsatz kommen. Weiterhin wird das Simulationsgebiet durch die frei vorgegebenen Simulationsparameter festgelegt. Die Eingangsparameter der Simulation sind dabei so gewählt, dass vermutlich das gesamte Gebiet der in praktischen Versuchen auftretenden Kombinationen abgebildet wird. Ob dies tatsächlich so ist, kann nicht gezeigt werden. Auch die Frage, ob die „Ecken“ des Simulationsgebietes, also gleichzeitig auftretende maximale oder minimale Werte der Eingangsparameter, in praktischen Versuchen erreicht werden, kann nicht beantwortet werden. Alle erstellten und als gültig befundenen virtuellen Versuchsreihen werden ausgewertet und haben damit Einfluss auf das Ergebnis der Simulation. Es ist denkbar, dass in praktischen Versuchen ein Teil der Versuchsreihen durch den Versuchsführenden nicht ausgewertet werden. Das ist mit einer Zensur der Stichproben der Schätzer gleichzusetzen. Solch eine Zensur bzw. subjektive Bewertung guter und schlechter Versuchsreihen in der Simulation zu berücksichtigen, ist mit vertretbarem Aufwand kaum umzusetzen. Genauso führt die beim Versuchsführenden vorhandene Erfahrung dazu, dass in praktischen Versuchen ein Teil der ermittelten Parameterschätzungen als unglaubwürdig eingestuft würden. Im Sinne der Simulation würden dabei gerade die besonders kleinen oder großen Verschätzungen eliminiert und damit die Streuung der Schätzer herabgesetzt.

Zu hinterfragen ist auch das Vorgehen, bei der Simulation der Versuchsreihen den Versuchsreihenstartwert und den Stufenabstand als feste Parameter vorzugeben. Das Verhältnis beider Größen beeinflusst die Treffsicherheit der Schätzungen. In praktischen Versuchen können diese Größen nicht exakt eingestellt werden. Es wäre daher denkbar, beide Größen als Zufallsgröße bei der Simulation zu berücksichtigen. Das würde aber zum einen die Anzahl der Eingangsparameter der Simulation erhöhen, zum anderen müssten für diese Größen zunächst sinnvolle Verteilungsparameter identifiziert werden.

Für weitere Arbeiten zur Ermittlung der Schätzfehler bei der Bestimmung der Festigkeiten durch Treppenstufenversuche ist das bisherige Vorgehen zu überdenken. Für die Auswertung-

⁷Das Lastverhältnis zwischen zwei benachbarten Laststufen ist also konstant, der Stufenabstand wird zu einem Stufensprung.

en der Treppenstufenversuche nach HÜCK und DIXON/MOOD werden die Schätzungen der Standardabweichung und des Erwartungswertes anhand der Mittelwerte und der Varianzen der Einzelversuche ermittelt. Die Schätzung der Standardabweichung wird dabei noch durch eine Korrekturfunktion modifiziert. Es wäre zum einen denkbar, auch für die Schätzung des Erwartungswertes eine solche Korrekturfunktion einzuführen. Zum anderen könnten auch die höheren Stichprobenmomente, also die Schiefe und die Wölbung der Einzelversuche explizit in die Korrekturfunktion einfließen. Es ist anzunehmen, dass sich durch diese genauere Beschreibung der Stichprobe auch der Einfluss der Startwertverschiebung besser erfassen lässt. Folgt man dieser Überlegung, dann bleibt aber auch festzuhalten, dass die tatsächliche Abfolge der Einzelversuche also die dokumentierte Folge von Brüchen und Durchläufern die genaueste Beschreibung des Treppenstufenversuches liefert. Es sollte daher versucht werden, durch Monte-Carlo-Simulationen zu ermitteln, unter welcher Kombination der Parameter der Grundgesamtheit das Eintreten einer spezifischen Abfolge von Durchläufern und Brüchen, also z. B. die Folge „xxxoxoxoooxo“ am wahrscheinlichsten ist. Ob sich diese wahrscheinlichsten Parameterkombination dann als Schätzer eignen und wie groß die dabei entstehenden Schätzfehler sind, könnte in einer zweiten Simulation überprüft werden. Hier wäre es für weitere Untersuchungen denkbar, die simulierten Treppenstufenverläufe als einen Referenzdatensatz abzulegen. Dann könnten alle weiteren Untersuchungen anhand ein und des selben Datenmaterials durchgeführt und somit die Vergleichbarkeit der Ergebnisse gewährleistet werden.

Literatur

- [1] R. Adenstedt: „Streuung der Schwingfestigkeit“, Diss., Fakultät für Mathematik/Informatik und Maschinenbau, Technische Universität Clausthal, 2001.
- [2] J. Böhm: „Zur Vorhersage von Dauerschwingfestigkeiten ungekerbter und gekerbter Bauteile unter Berücksichtigung des statischen Größeneinflusses“, Diss., TU München, Lehrstuhl A für Mechanik, 1980.
- [3] H. Bomas, T. Linkewitz und P. Mayr: „Anwendung des Fehlstellenmodells auf die Dauerfestigkeit des Stahls 100Cr6 im bainitischen Zustand“, in: *Materialwissenschaft und Werkstofftechnik* Band 29, Nr. 4 (1998), S. 185–192.
- [4] K. A. Brownlee, J. L. Hodges Jr. und M. Rosenblatt: „The Up-and-Down Method with Small Samples“, in: *Journal of the American Statistical Association* Band 48, Nr. 262 (1953), S. 262–277.
- [5] E. Deubelbeiss: „Dauerfestigkeitsversuche mit einem modifizierten Treppenstufenversuch“, in: *Materialprüfung* Band 16, Nr. 8 (1974), S. 240–244.
- [6] DIN 1319: *Grundlagen der Meßtechnik*, Deutsches Institut für Normung e.V., Beuth Verlag, 1996.
- [7] DIN 53804: *Statistische Auswertungen*, Deutsches Institut für Normung e.V., Beuth Verlag, 2002.
- [8] W. J. Dixon: „The Up-and-Down Method for Small Samples“, in: *Journal of the American Statistical Association* Band 60, Nr. 312 (1965), S. 967–978.
- [9] W. Dixon und A. Mood: „A Method for Obtaining and Analyzing Sensitivity Data“, in: *Journal of the American Statistical Association* Band 43, Nr. 241 (1948), S. 109–126.
- [10] D. Dorff: „Vergleich verschiedener statistischer Transformationsverfahren auf ihre Anwendbarkeit zur Ermittlung der Dauerschwingfestigkeit“, Diss., Technische Universität Berlin, Fakultät für Maschinenwesen, 1961.
- [11] K.-G. Eulitz: *Beurteilung der Zuverlässigkeit von Lebensdauervorhersagen nach dem Nennspannungskonzept und dem Örtlichen Konzept anhand einer Sammlung von Betriebsfestigkeitsversuchen*, Habilitation, Technische Universität Dresden, 1999.
- [12] R. A. Fisher und F. Yates: *Statistical Tables for Biological, Agricultural and Medical Research*, hrsg. von O. Boyd, 1963.
- [13] S. Götz: „Weiterentwicklung eines bruchmechanischen Konzeptes zur formzahlfreien Abschätzung der Dauerfestigkeit gekerbter Strukturen am Beispiel verschiedener Sinterstähle“, Diss., Technische Universität Dresden, 2012.
- [14] E. Haibach: *Betriebsfestigkeit: Verfahren und Daten zur Bauteilberechnung*, 3. (korrigierte und ergänzte) Auflage, Springer, Berlin, 2006.

- [15] P. Hantschke: „Routinen zur Auswertung von Wöhlerlinien im Zeit und Dauerfestigkeitsbereich.“, in: *Institutinterne Routinenbibliothek 'Zenfab'*, persönliche Mitteilungen (2012).
- [16] J. Hartung, B. Elpelt und K.-H. Klösener: *Statistik, Lehr- und Handbuch der angewandten Statistik*, 8., durchgesehene Auflage, R. Oldenbourg Verlag, 1991.
- [17] J. Hartung, B. Elpelt und K.-H. Klösener: *Statistik: Lehr- und Handbuch der angewandten Statistik*, 14., unwesentlich veränd. Aufl., R. Oldenbourg Verlag, 2005.
- [18] S. Hauschild: *Reibkorrosion. Gestaltung und Festigkeitsberechnung reibkorrosionsgefährdeter Fügeverbindungen*, Abschlussbericht Vorhaben Nr. 1142, Rechnerische Festigkeitsbewertung zur Auslegung von reibdauerbeanspruchten Maschinenbauteilen, FVV Abschlussbericht 1097, 2016.
- [19] M. Hück, J. Bergmann und W. Schütz: *Bewertung der Schwingfestigkeit der mikrolegierten Stähle 27 MnVS 6 und 38 MnVS 5*, IABG-Bericht TF-2804/1, Bd. 1, 1990.
- [20] M. Hück: „Auswertung von Stichproben normalverteilter, quantitativer Merkmalsgrößen“, in: *Materialwissenschaft und Werkstofftechnik* Band 25, Nr. 1 (1994), S. 20–29.
- [21] M. Hück: „Ein verbessertes Verfahren für die Auswertung von Treppenstufenversuchen“, in: *Materialwissenschaft und Werkstofftechnik* Band 14, Nr. 12 (1983), S. 406–417.
- [22] M. Hück: *Statistik im Ingenieurwesen*, Manuskript zur Vorlesungsreihe, Institut für Hüttenmaschinen und Maschinelle Anlagentechnik, TU Clausthal, 1990.
- [23] M. Hück: *Unveröffentlichte Mitteilung Herrn Manfred Hücks an Prof. Dr.-Ing. Erwin Haibach*, Aug. 2000.
- [24] T. Jung: „Schwingfestigkeit unter Berücksichtigung des Spannungsbedingten und des Technologiebedingten Größeneinflusses“, Diss., Technische Universität Clausthal, Fakultät für Bergbau, Hüttenwesen und Maschinenwesen, 1997.
- [25] F. Klubberg und P. Beiss: „Modifizierte Prüf- und Auswertemethodik im Übergangsbereich zur Dauerschwingfestigkeit“, in: *VDI Berichte*, Nr. 1151 (1995), S. 777–780.
- [26] J. Liu: *Dauerfestigkeitsberechnung metallischer Bauteile*, Habilitation, Fakultät für Bergbau, Hüttenwesen und Maschinenwesen der Technischen Universität Clausthal, 2001.
- [27] W.-W. Maennig: „Bemerkungen zur Beurteilung des Dauerschwingverhaltens von Stahl und einige Untersuchungen zur Bestimmung des Dauerfestigkeitsbereichs“, in: *Materialprüfung* Band 12, Nr. 4 (Apr. 1970), S. 124–131.
- [28] W.-W. Maennig: „Das Abgrenzungsverfahren, eine kostensparende Methode zur Ermittlung von Schwingfestigkeitswerten - Theorie, Praxis und Erfahrungen“, in: *Materialprüfung* Band 19, Nr. 8 (Aug. 1977).
- [29] H. Mauch: „Lebensdauerstatistik“, in: *Forschungsheft Forschungsvereinigung Antriebstechnik e. V.* Band 591 (1999).
- [30] H. Mogwitz: „Anwendung statistischer Methoden bei der Berechnung und experimentellen Bestimmung der Ermüdungsfestigkeit“, Diss., Technische Universität Dresden. Fakultät Maschinenwesen, 1988.

- [31] C. Müller: „Zur statistischen Auswertung von Wählerlinien“, Diss., Fakultät für Mathematik/Informatik und Maschinenbau, Technische Universität Clausthal, 2015.
- [32] C. Müller u. a.: *Zur Treffsicherheit der experimentellen Dauerfestigkeitsschätzung*, hrsg. von A. Esderts, Technical Report Series, Fakultät für Mathematik/Informatik und Maschinenbau, Technische Universität Clausthal, 2014.
- [33] L. Papula: *Mathematik für Ingenieure und Naturwissenschaftler - 3 : Vektoranalysis, Wahrscheinlichkeitsrechnung, mathematische Statistik, Fehler- und Ausgleichsrechnung*, 3., verb. Aufl., Vieweg, 1999.
- [34] D. Radaj und M. Vormwald: *Ermüdungsfestigkeit*, Springer-Verlag Berlin Heidelberg New York, 2007.
- [35] R. Rennert u. a.: *Rechnerischer Festigkeitsnachweis für Maschinenbauteile - FKM-Richtlinie*, (6. Auflage), VDMA-Verlag GmbH, 2012.
- [36] E. Rossow: „Eine einfache rechenschiebernäherung an die normal scores entsprechenden Prozentpunkte“, in: *Zeitschrift für wirtschaftliche Fertigung* Band 59, Nr. 12 (1964), S. 146–147.
- [37] A. Rotem: „Accelerated fatigue testing method“, in: *International Journal of Fatigue* Band 3, Nr. 4 (1981), S. 211–215.
- [38] P. Royston: „Approximating the Shapiro-Wilk W-test for non-normality“, in: *Statistics and Computing* Band 2 (1992), S. 117–119.
- [39] L. Sachs: *Statistische Methoden 2. Planung und Auswertung*, Springer, 1990.
- [40] L. Sachs: *Statistische Methoden. Planung und Auswertung*, 6., neubearb. u. erw. Aufl., Springer-Verlag, 1988.
- [41] L. Sachs und J. Hedderich: *Angewandte Statistik. Methodensammlung mit R*, 13., aktualisierte und erw. Aufl., Springer, 2009.
- [42] W. Uhlmann: *Statistische Qualitätskontrolle*, Teubner Verlag, 1988.
- [43] VDI 4009: *VDI-Richtlinie*, Verein Deutscher Ingenieure e. V., 1984.

A Grafische Darstellung der Simulationsergebnisse

A.1 Vom Auswerteverfahren unabhängige Ergebnisse

A.1.1 Anzahl der mit gültigen Versuchen belegten Horizonte bei Treppenstufenversuchen

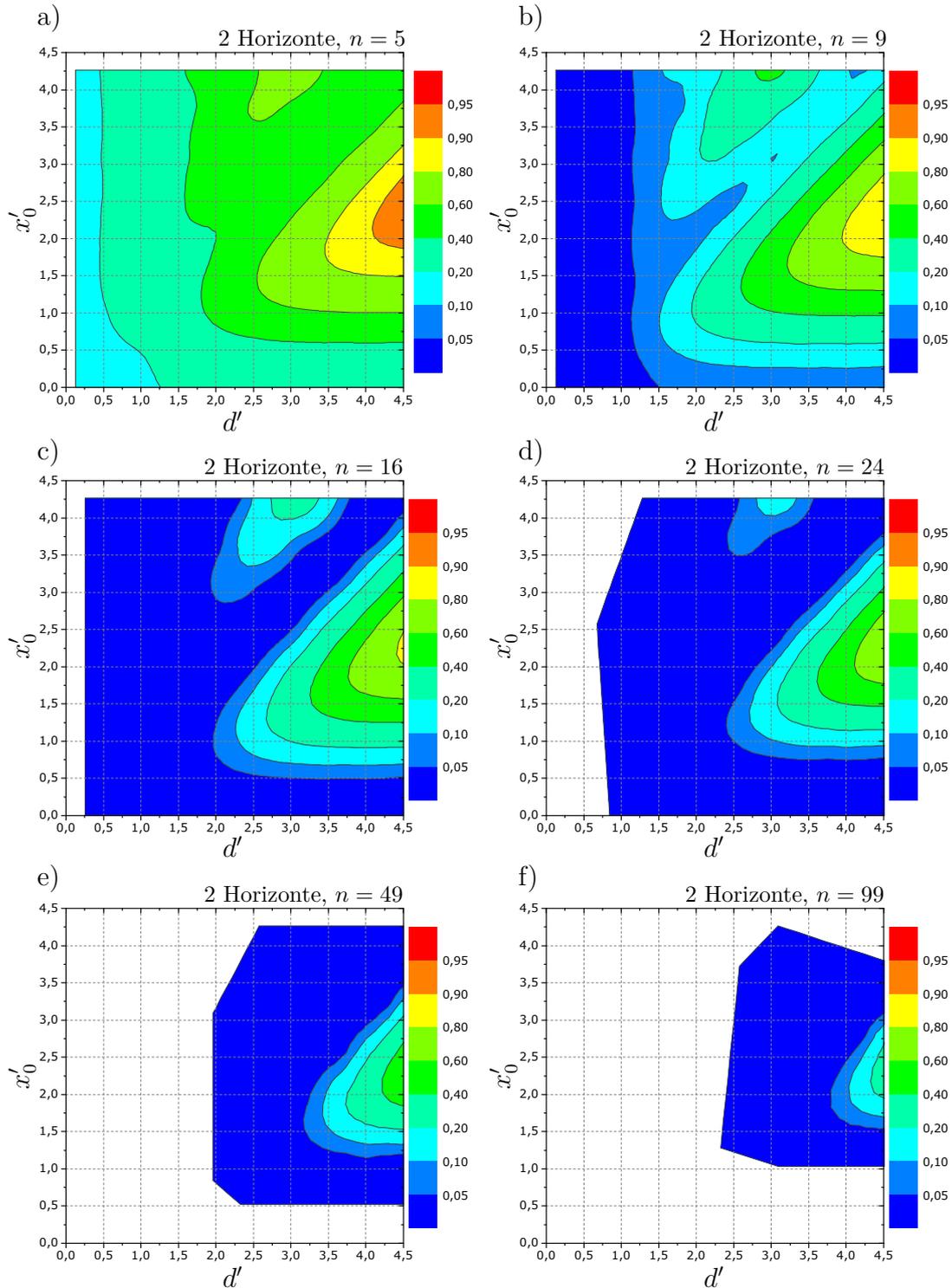


Abb. A.1: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit im Ergebnis genau 2 belegten Horizonten, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

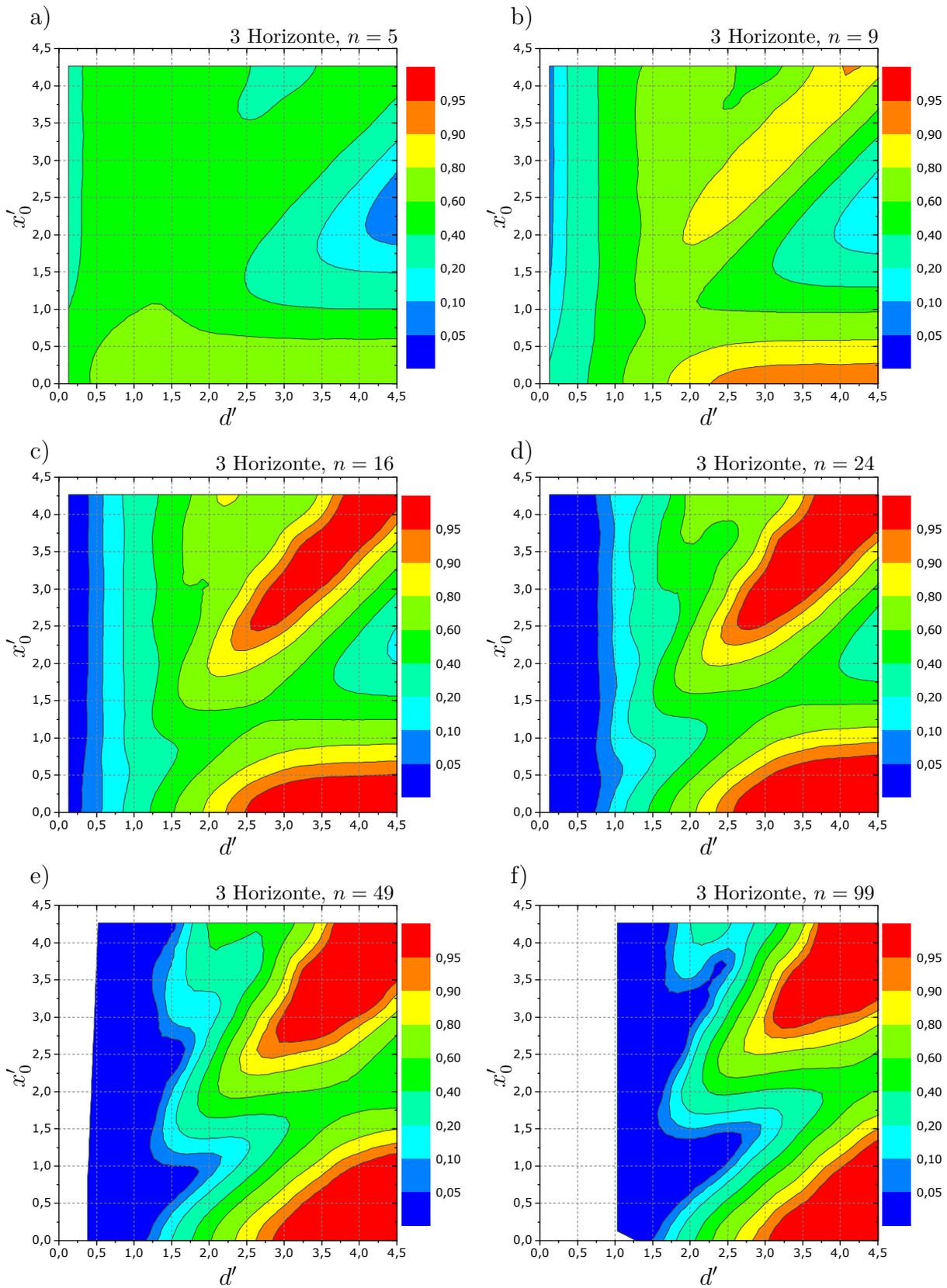


Abb. A.2: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit im Ergebnis genau 3 belegten Horizonten, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

A Grafische Darstellung der Simulationsergebnisse

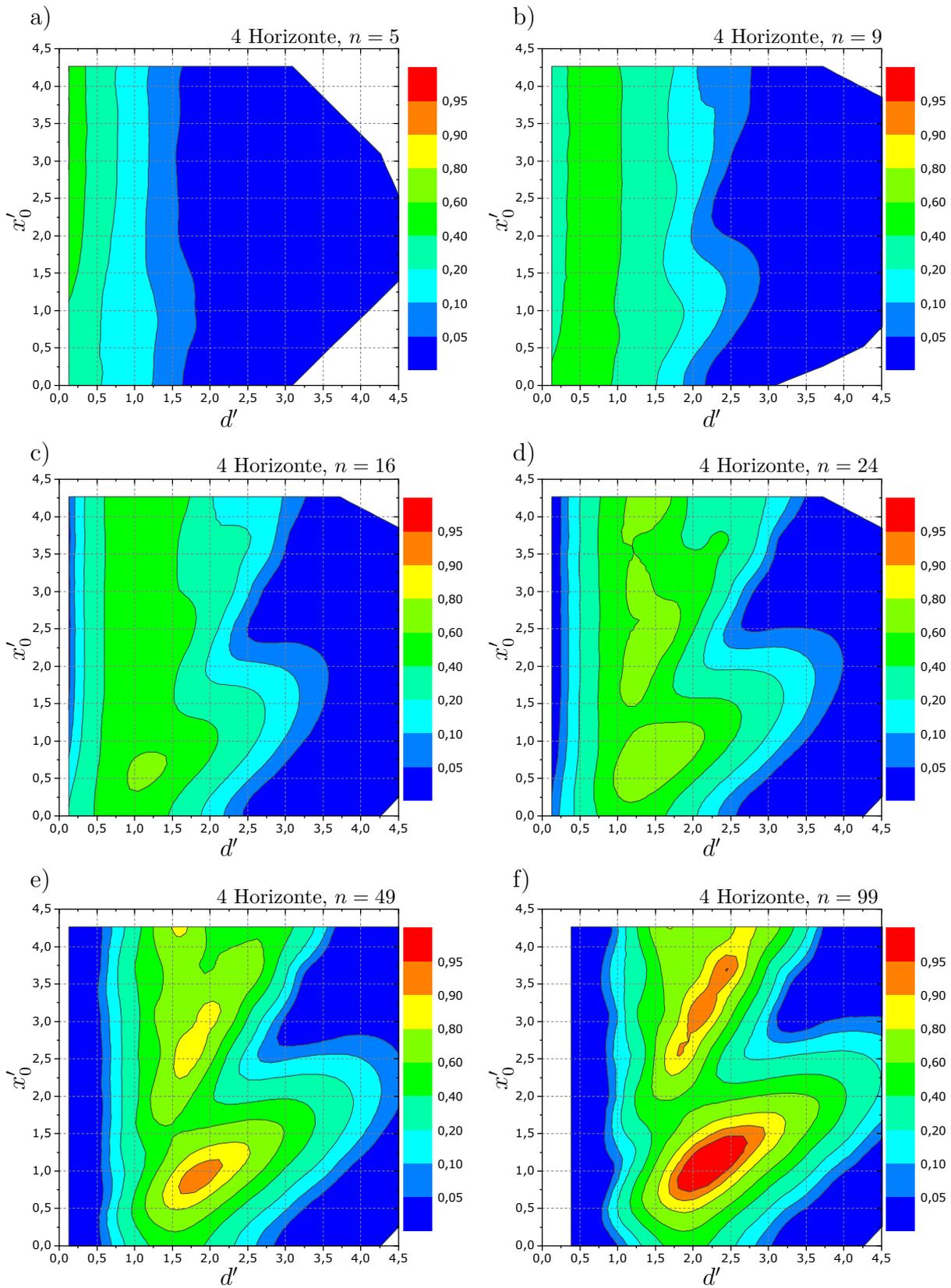


Abb. A.3: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit im Ergebnis genau 4 belegten Horizonten, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

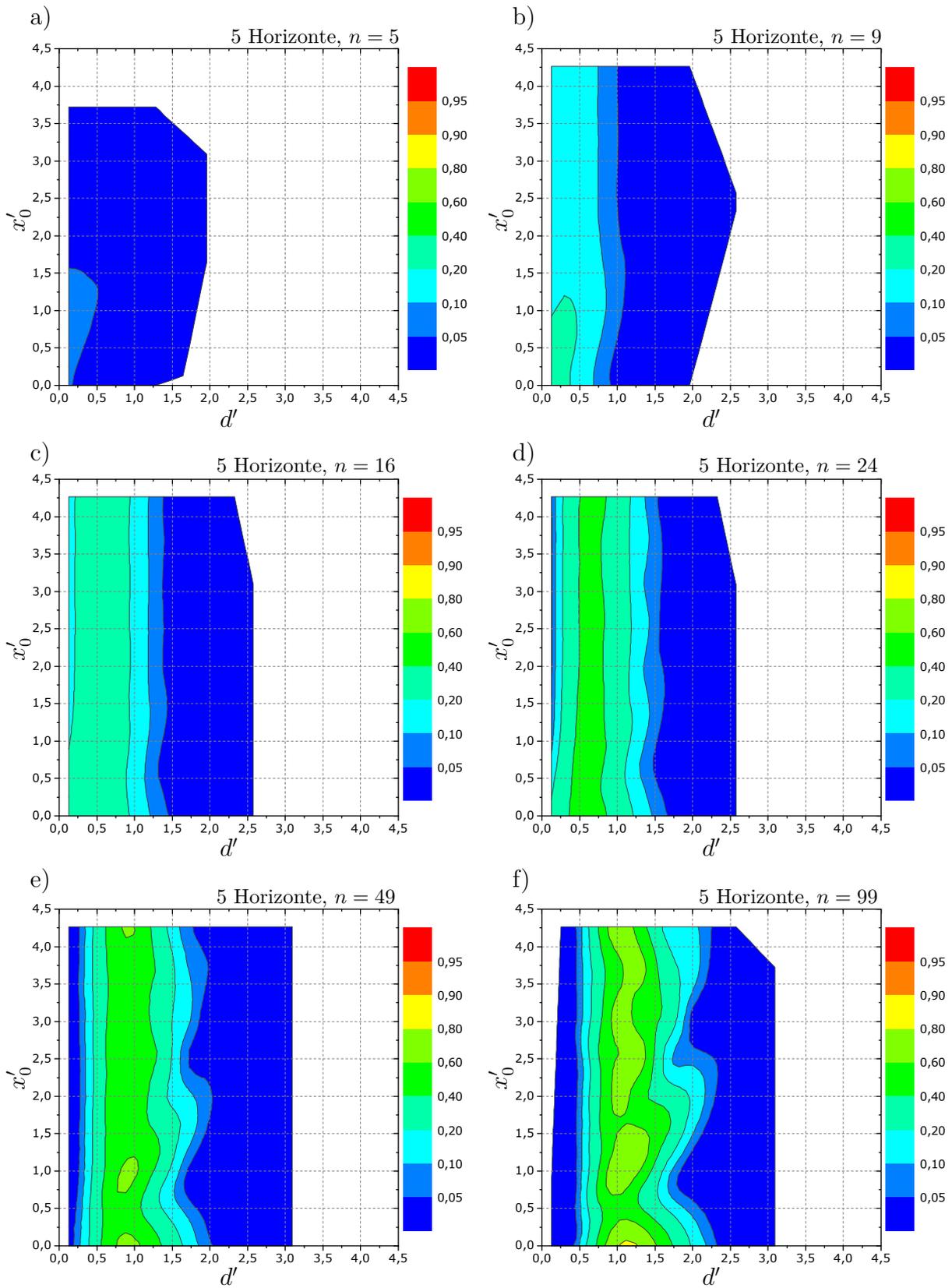


Abb. A.4: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit im Ergebnis genau 5 belegten Horizonten, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

A Grafische Darstellung der Simulationsergebnisse

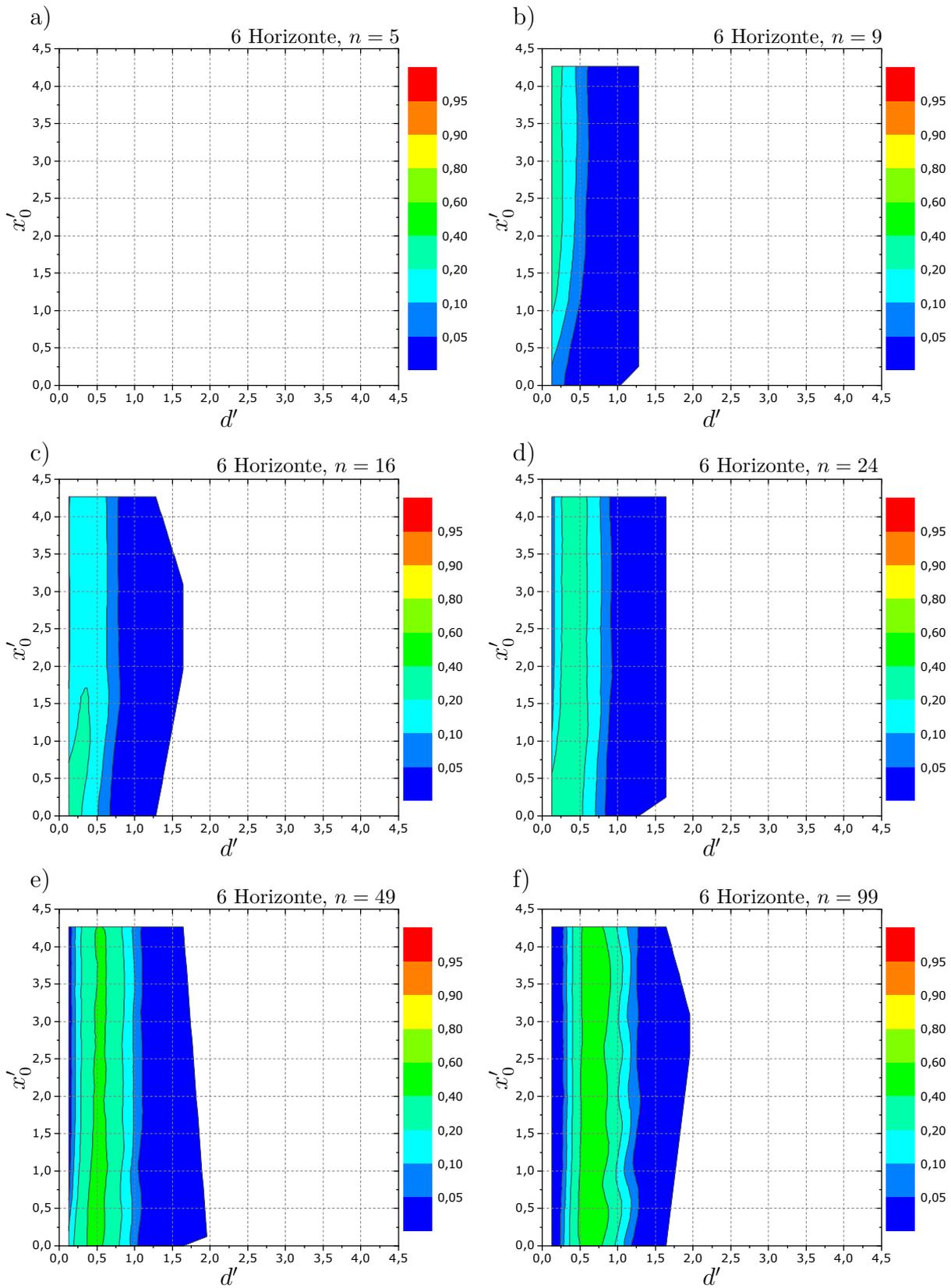


Abb. A.5: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit im Ergebnis genau 6 belegten Horizonten, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

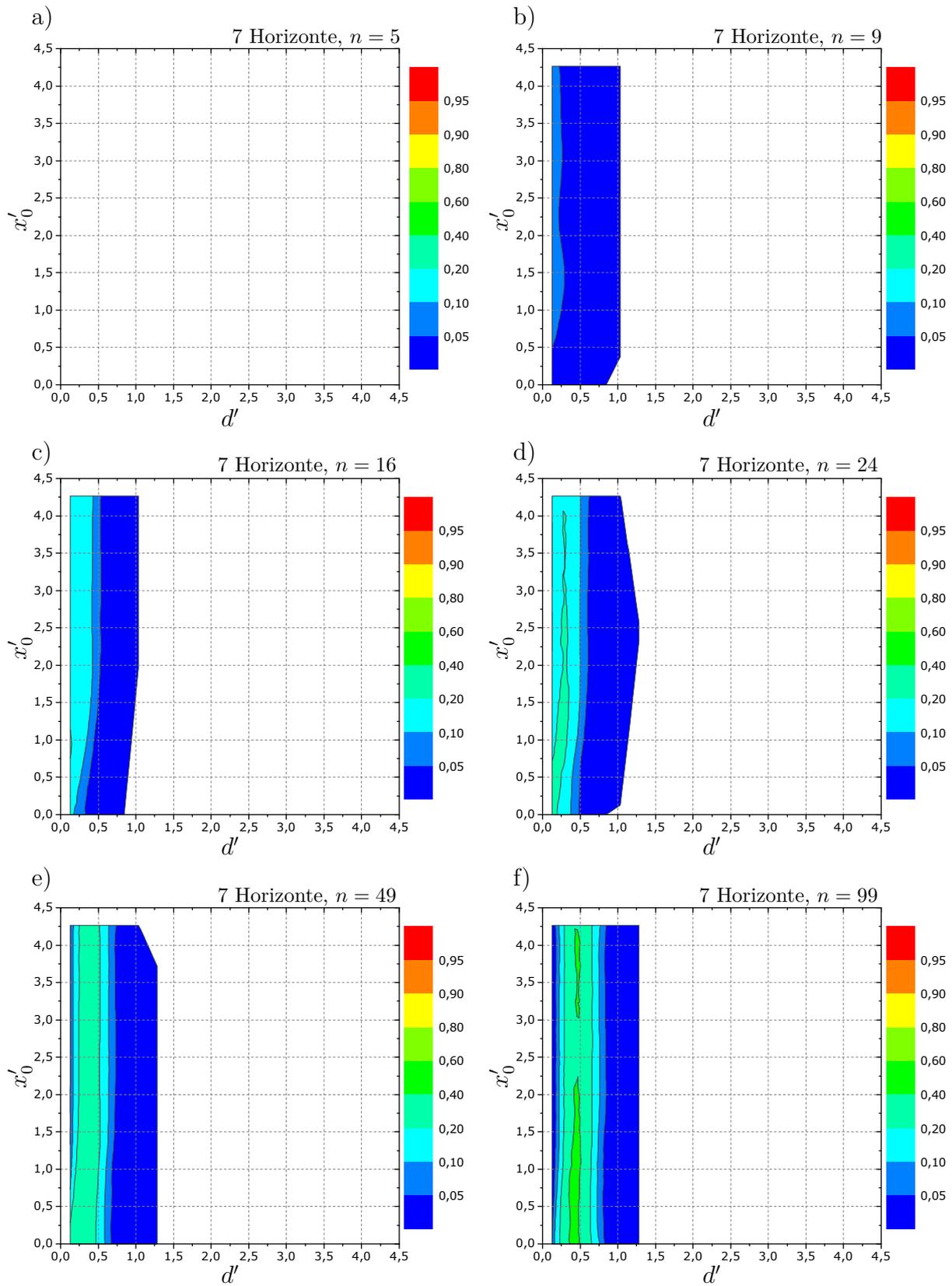


Abb. A.6: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit im Ergebnis genau 7 belegten Horizonten, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

A.1.2 Anzahl ungültiger Versuche in einer Versuchsreihe

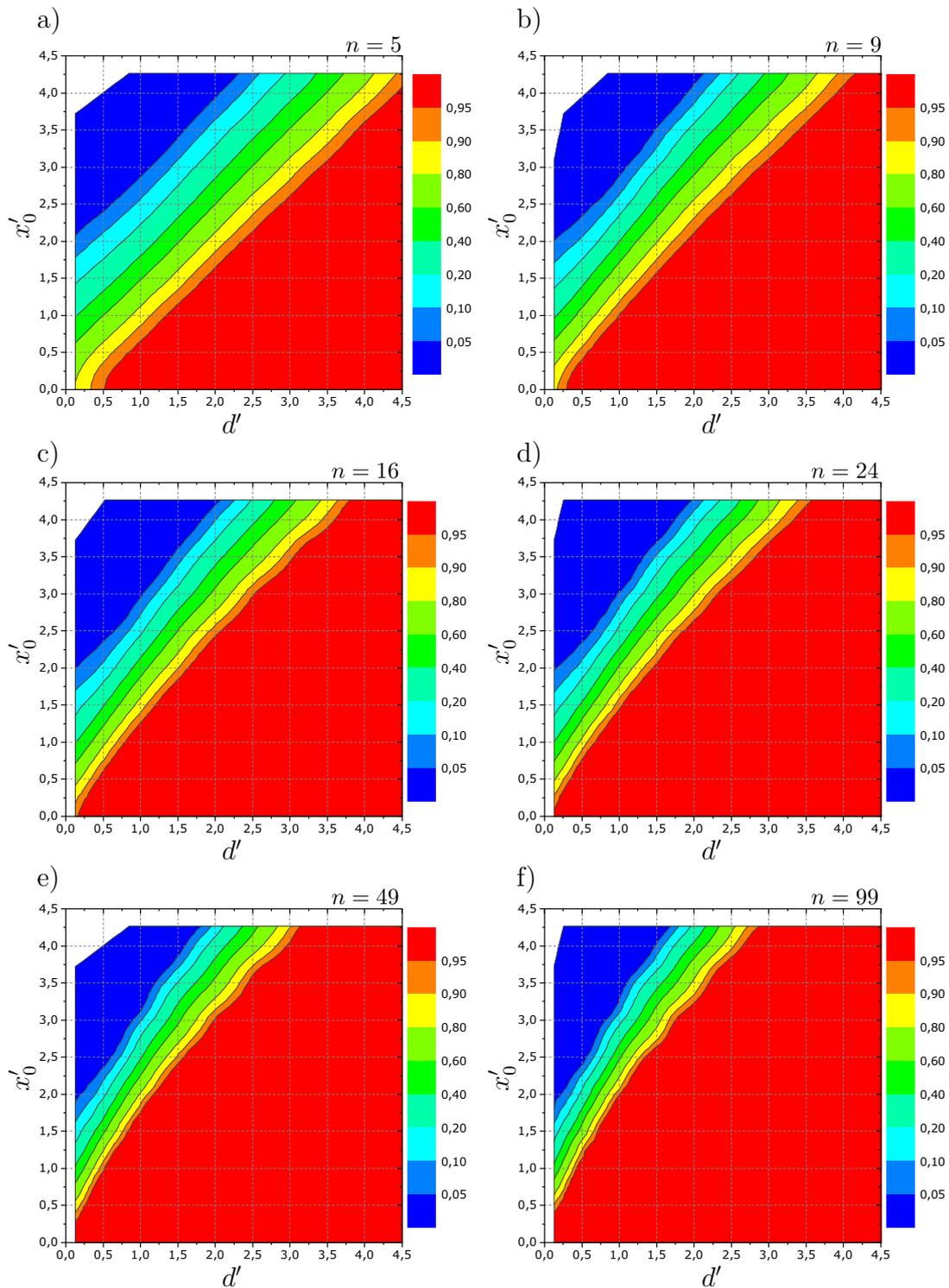


Abb. A.7: Relative Häufigkeit simulierter Treppenstufenversuchsreihen ohne ungültige Versuche im Abschnitt für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

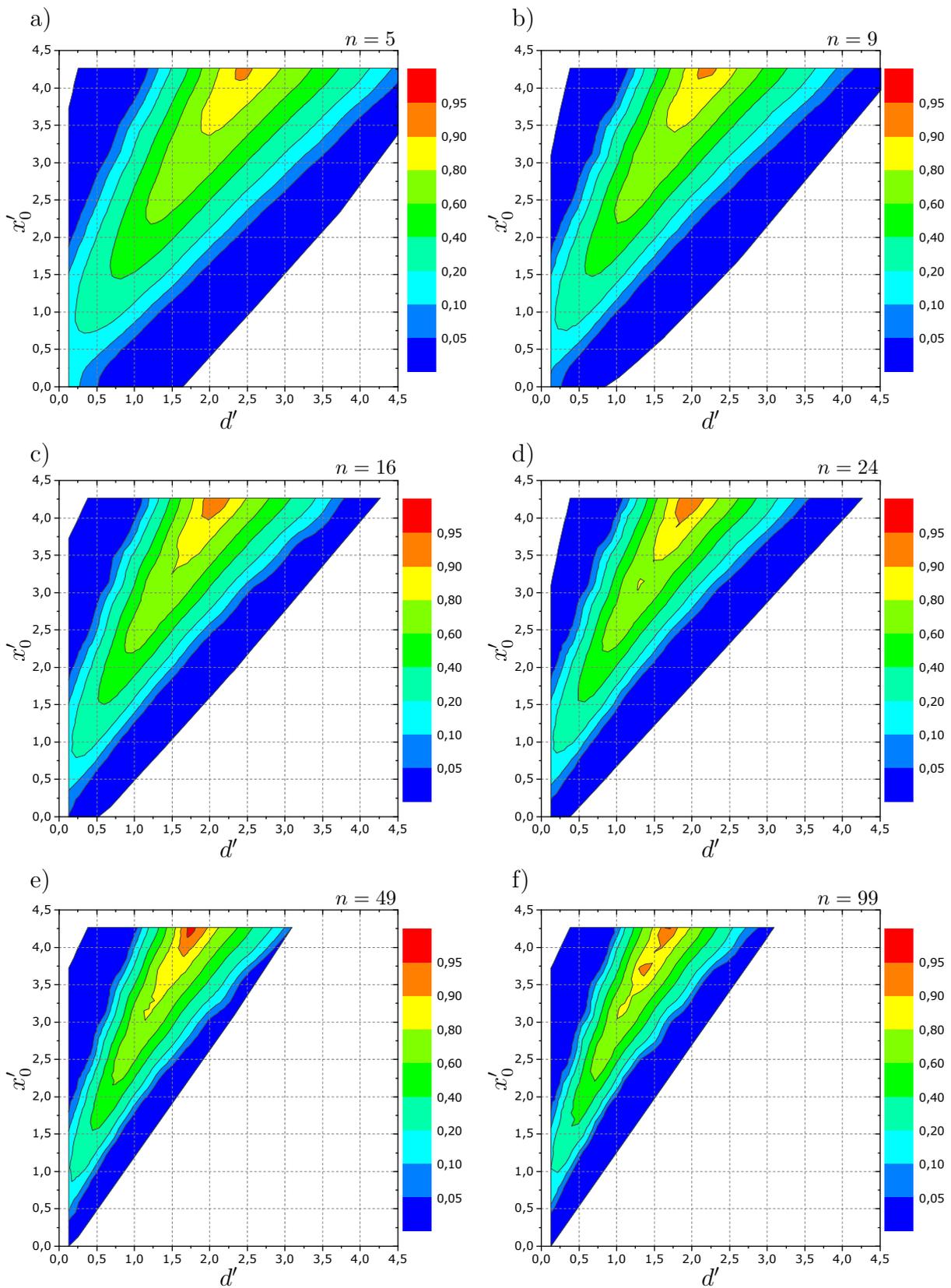


Abb. A.8: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit einem Versuch im Anschnitt, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

A Grafische Darstellung der Simulationsergebnisse

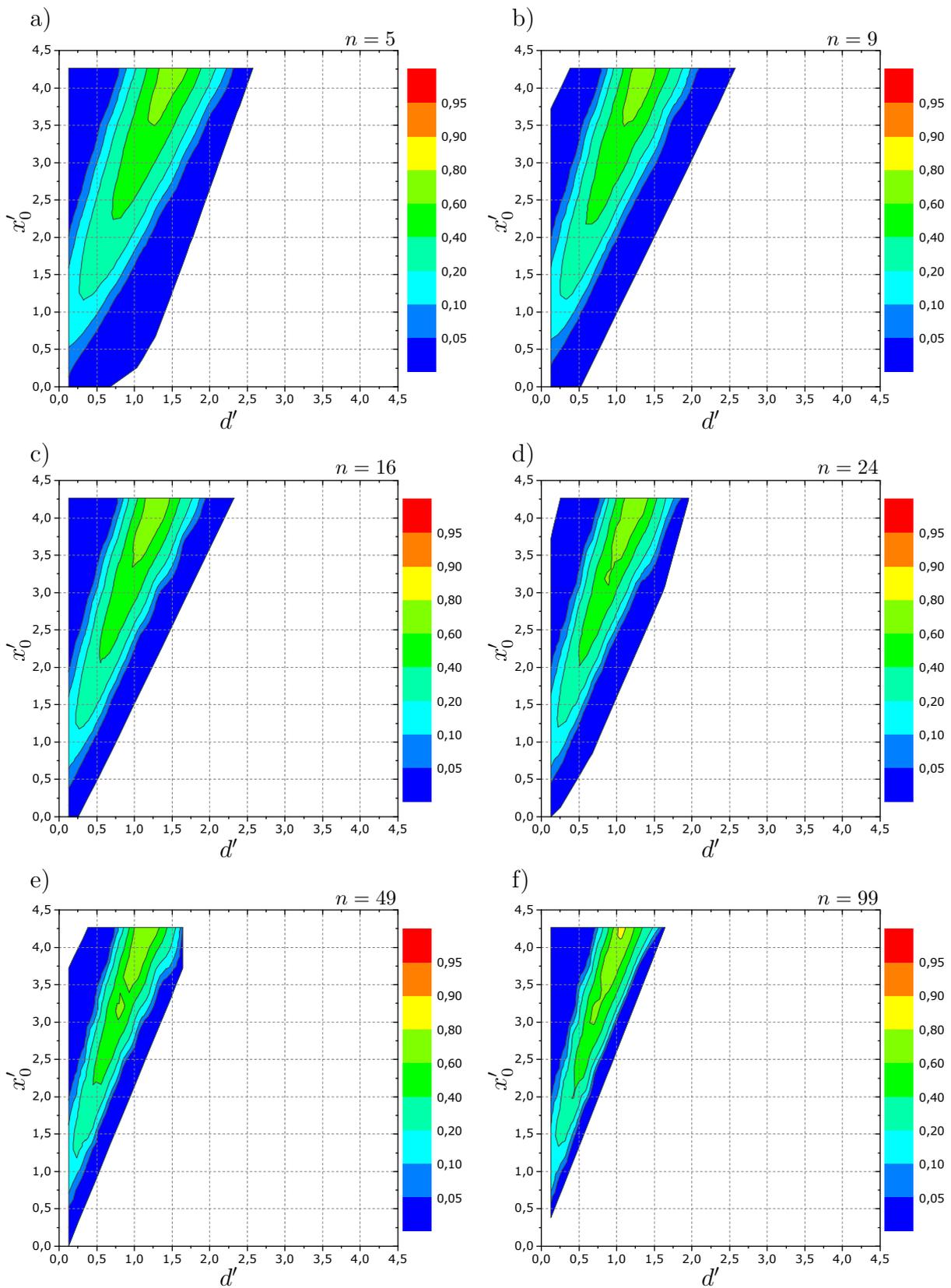


Abb. A.9: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit 2 Versuchen im Anschnitt, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

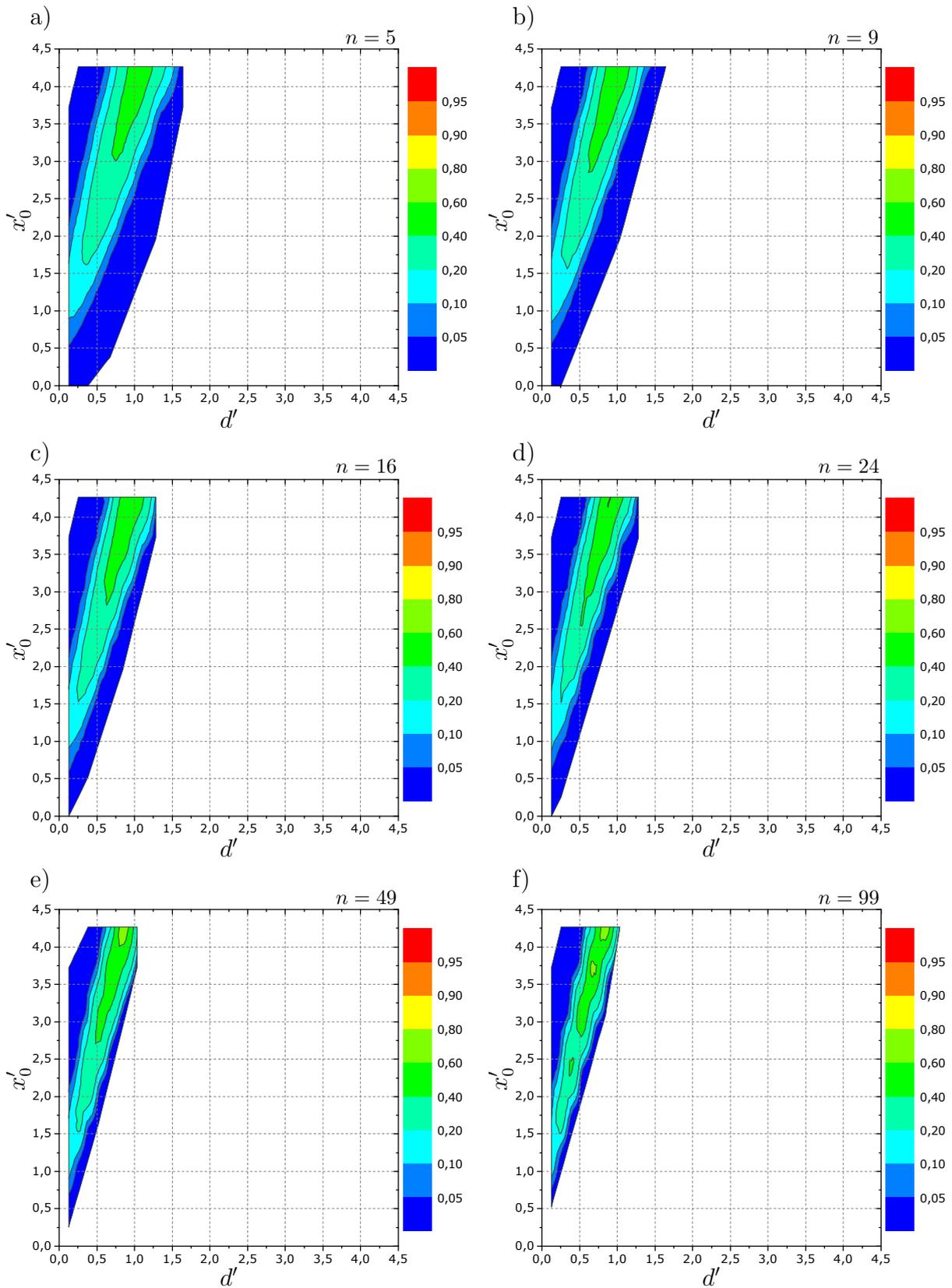


Abb. A.10: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit 3 Versuchen im Anschnitt, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

A Grafische Darstellung der Simulationsergebnisse

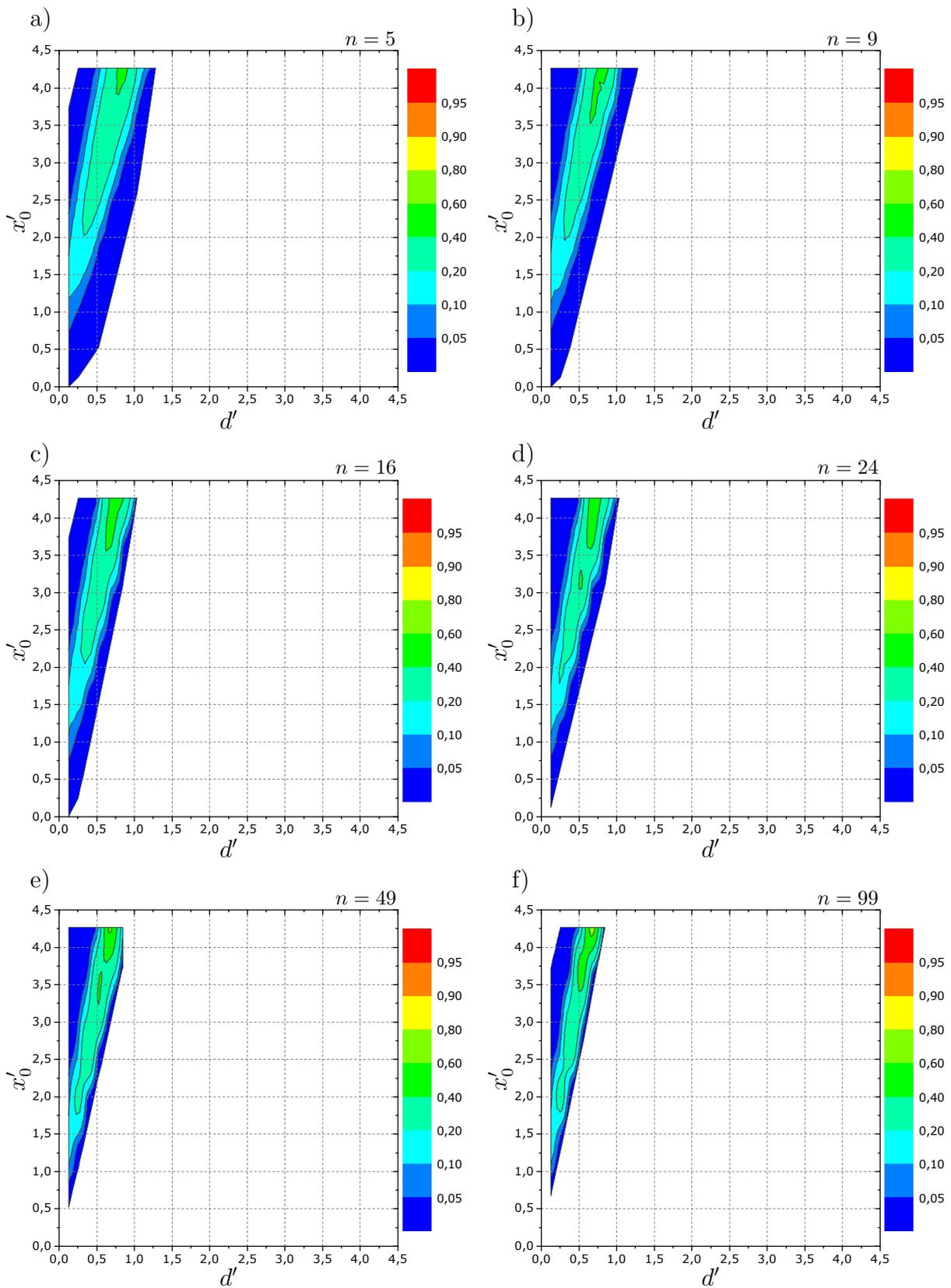


Abb. A.11: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit 4 Versuchen im Abschnitt, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

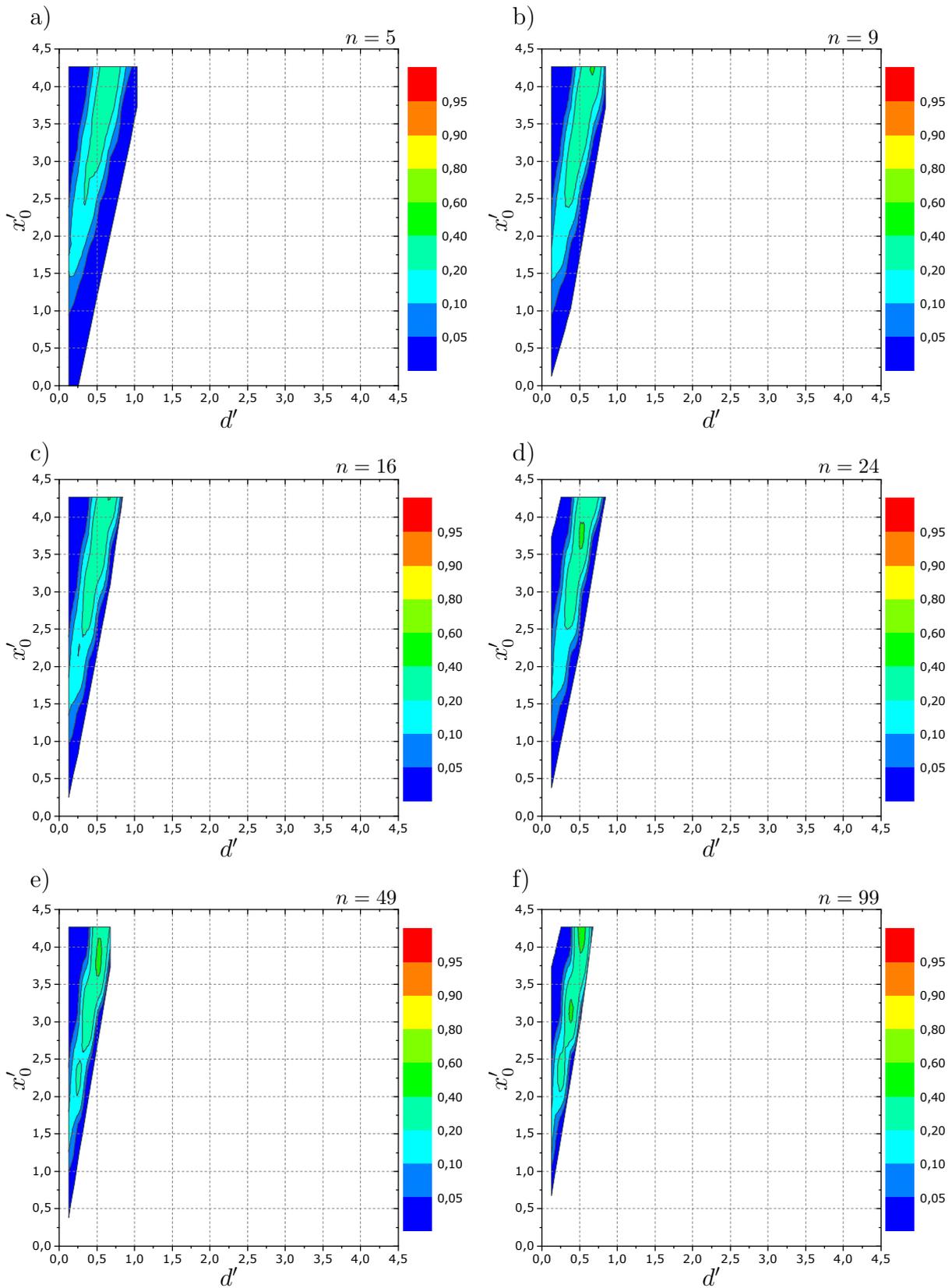


Abb. A.12: Relative Häufigkeit simulierter Treppenstufenversuchsreihen mit 5 Versuchen im Abschnitt, für $n = 5$ in a), $n = 9$ in b), $n = 16$ in c), $n = 24$ in d), 49 in e) oder $n = 99$ gültige Einzelversuche in f).

A.2 Erwartungswertschätzung bei bekannter Standardabweichung

A.2.1 Mittlere Verzerrungen und Standardfehler des Schätzers m'

Simulationsergebnisse der Treppenstufenversuche

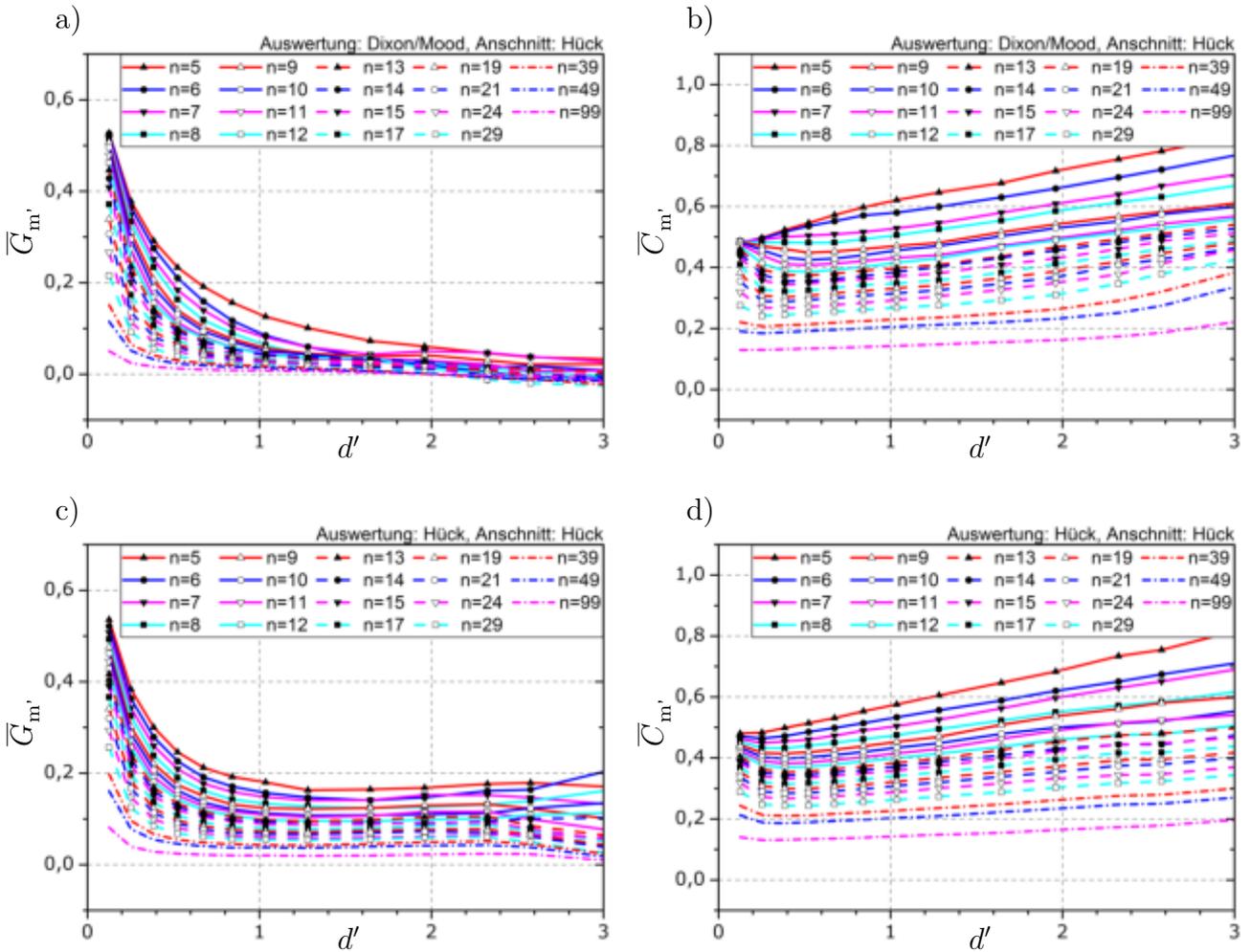


Abb. A.13: Mittlere Verzerrungen und Standardfehler des Schätzers m' für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethode nach DIXON/MOOD in a) und b) sowie HÜCK in c) und d).

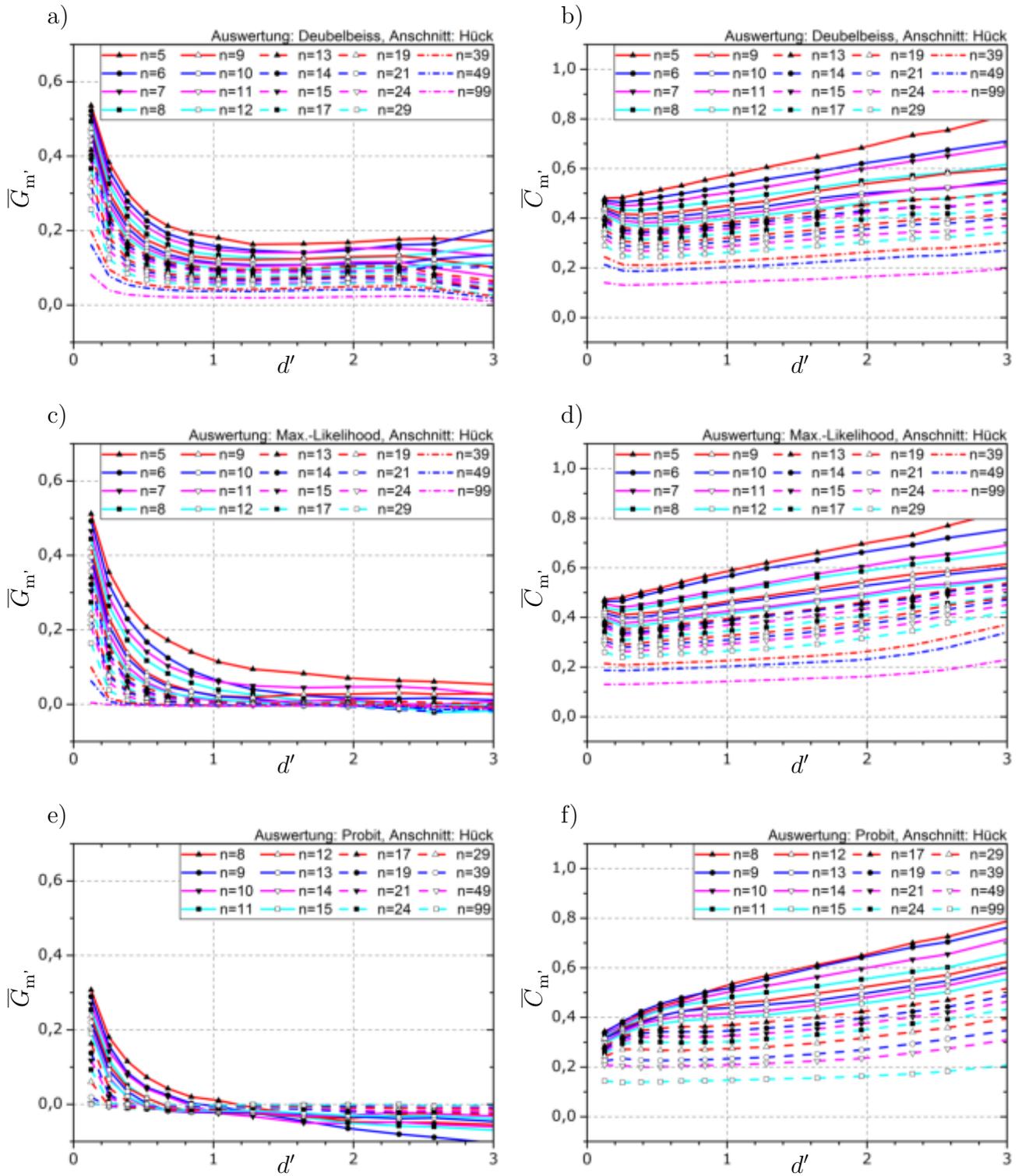


Abb. A.14: Mittlere Verzerrungen und Standardfehler des Schätzers m' für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

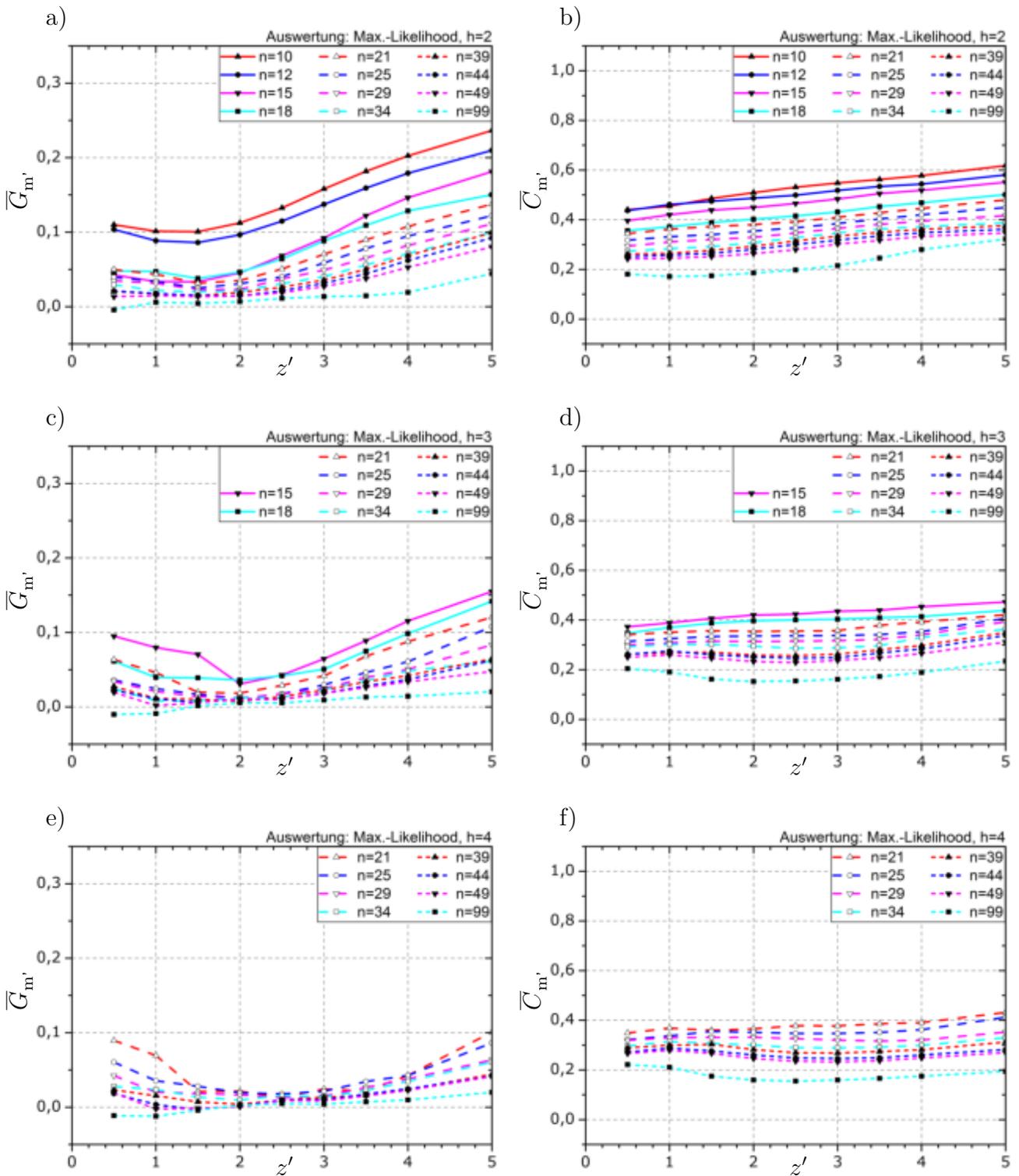


Abb. A.15: Mittlere Verzerrungen und Standardfehler des Schätzers m' für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

A.2 Erwartungswertschätzung bei bekannter Standardabweichung

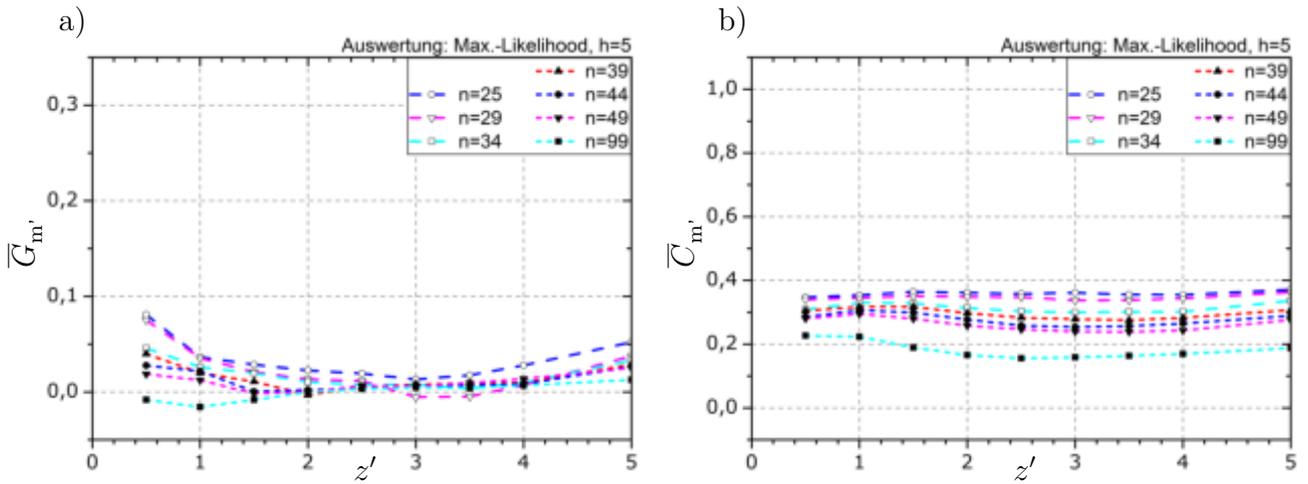


Abb. A.16: Mittlere Verzerrungen und Standardfehler des Schätzers m' für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

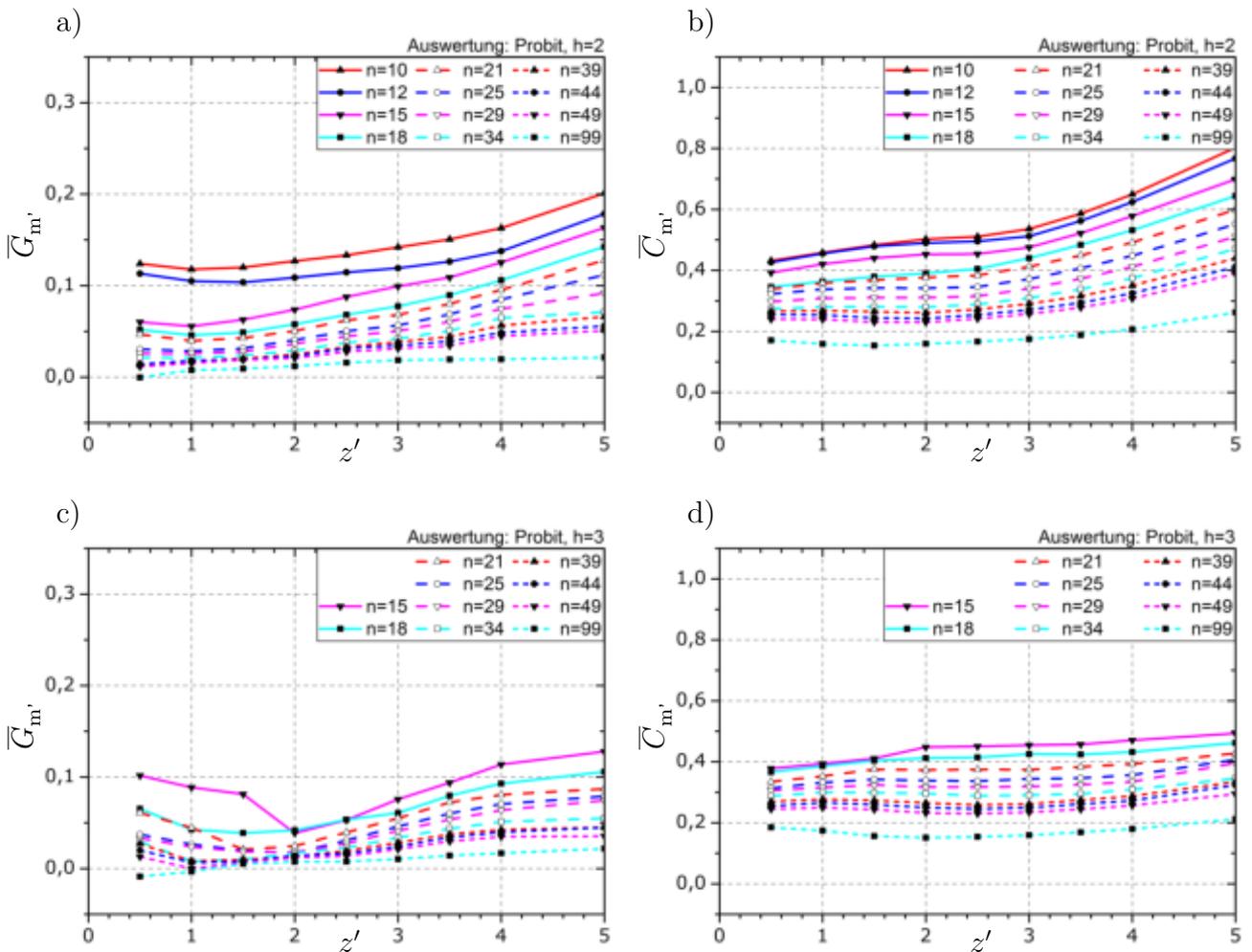


Abb. A.17: Mittlere Verzerrungen und Standardfehler des Schätzers m' für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

A Grafische Darstellung der Simulationsergebnisse

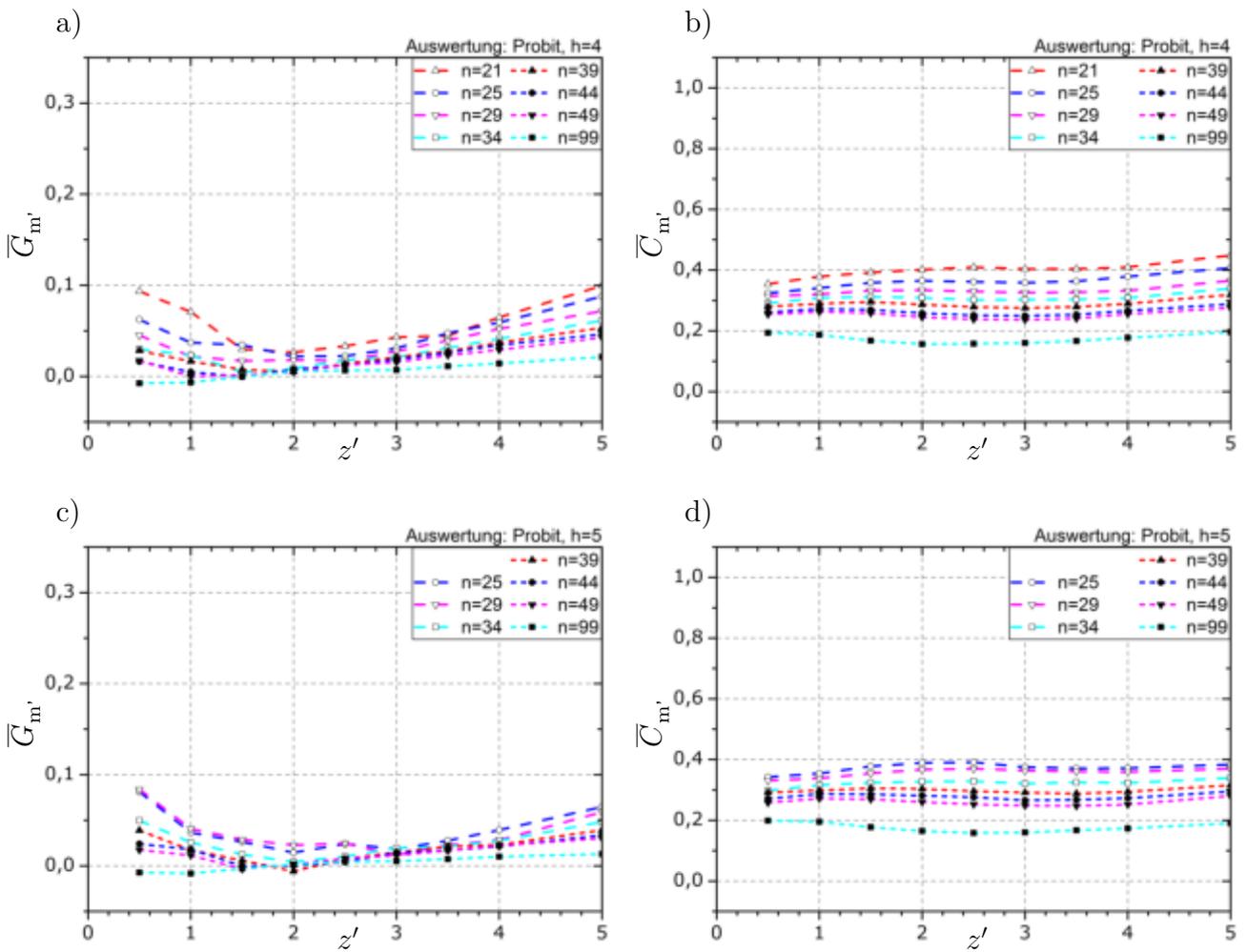


Abb. A.18: Mittlere Verzerrungen und Standardfehler des Schätzers m' für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.2.2 Verallgemeinerte Standardfehler des Schätzers m'

Simulationsergebnisse der Treppenstufenversuche

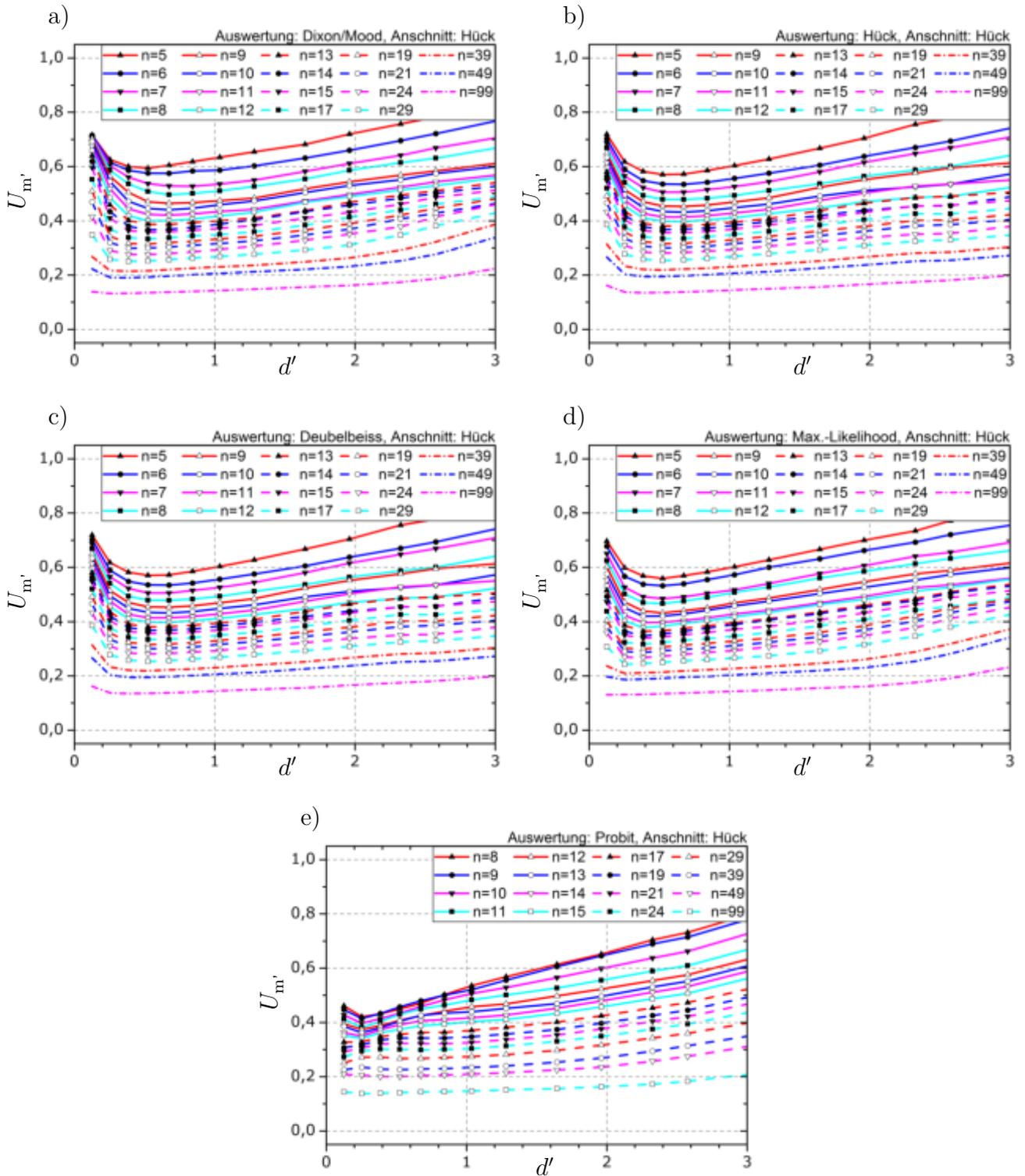


Abb. A.19: Verallgemeinerte Standardfehler $U_{m'}$ des Schätzers m' für Treppenstufenversuche mit der Anschnittsdefinition nach HÜCK, für die Auswertemethode nach DIXON/MOOD in a), nach HÜCK in b), nach DEUBELBEISS in c), MAX.-LIKELIHOOD-METHODE in d) sowie PROBIT-METHODE in e).

Simulationsergebnisse der Horizontversuche

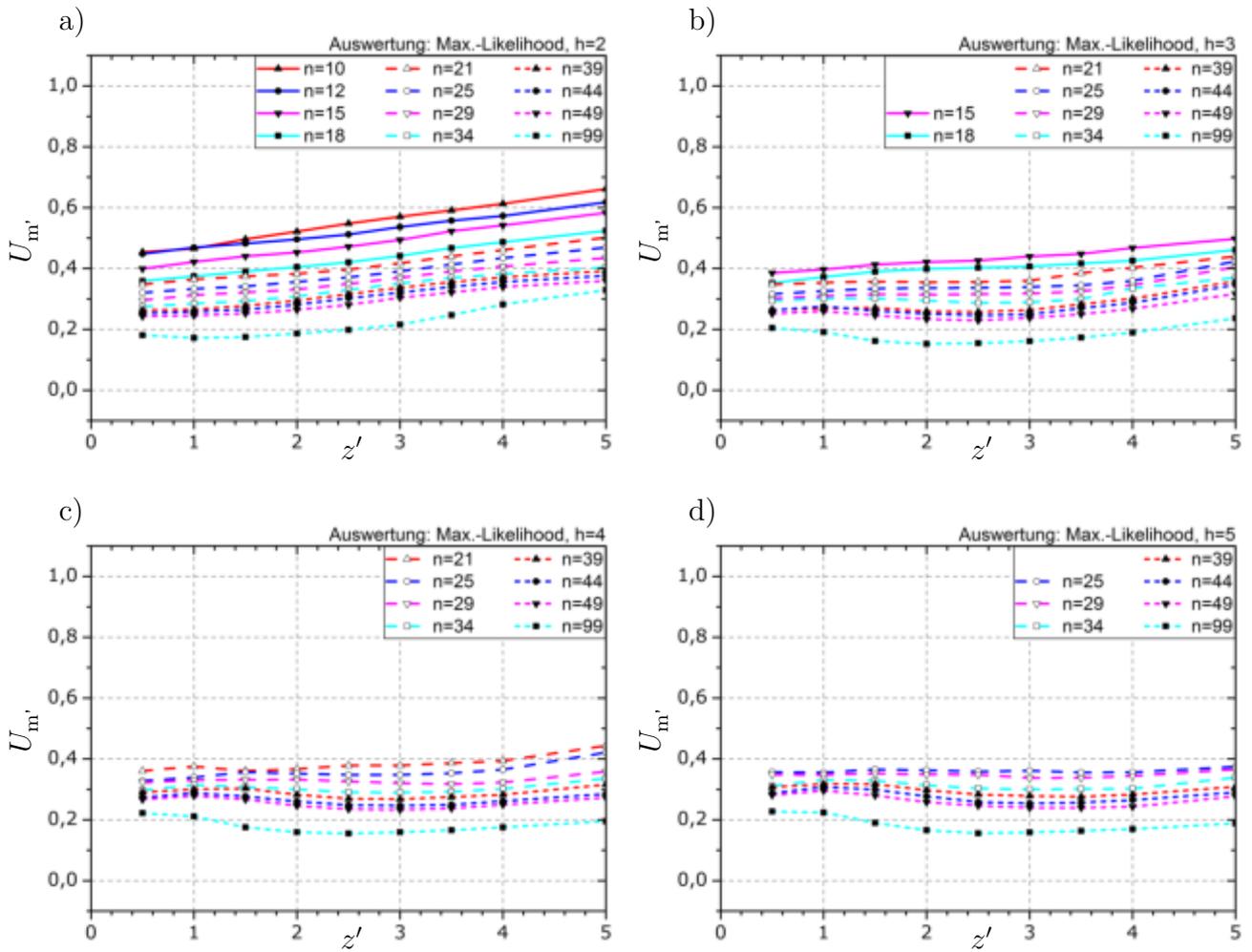


Abb. A.20: Verallgemeinerte Standardfehler $U_{m'}$ des Schätzers m' für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

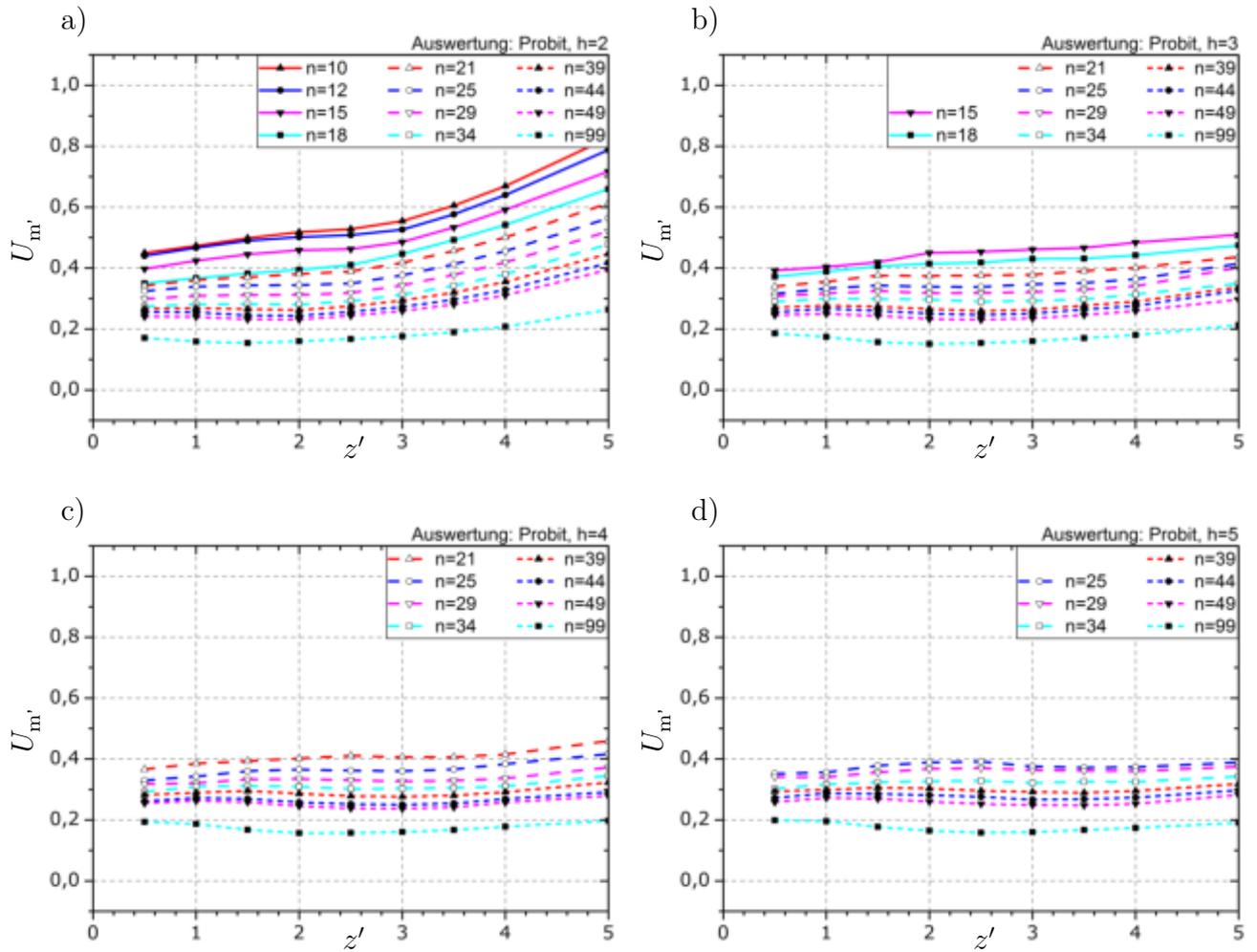


Abb. A.21: Verallgemeinerte Standardfehler $U_{m'}$ des Schätzers m' für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

A.2.3 RMS-Schätzfehler des Schätzers m'

Simulationsergebnisse der Treppenstufenversuche

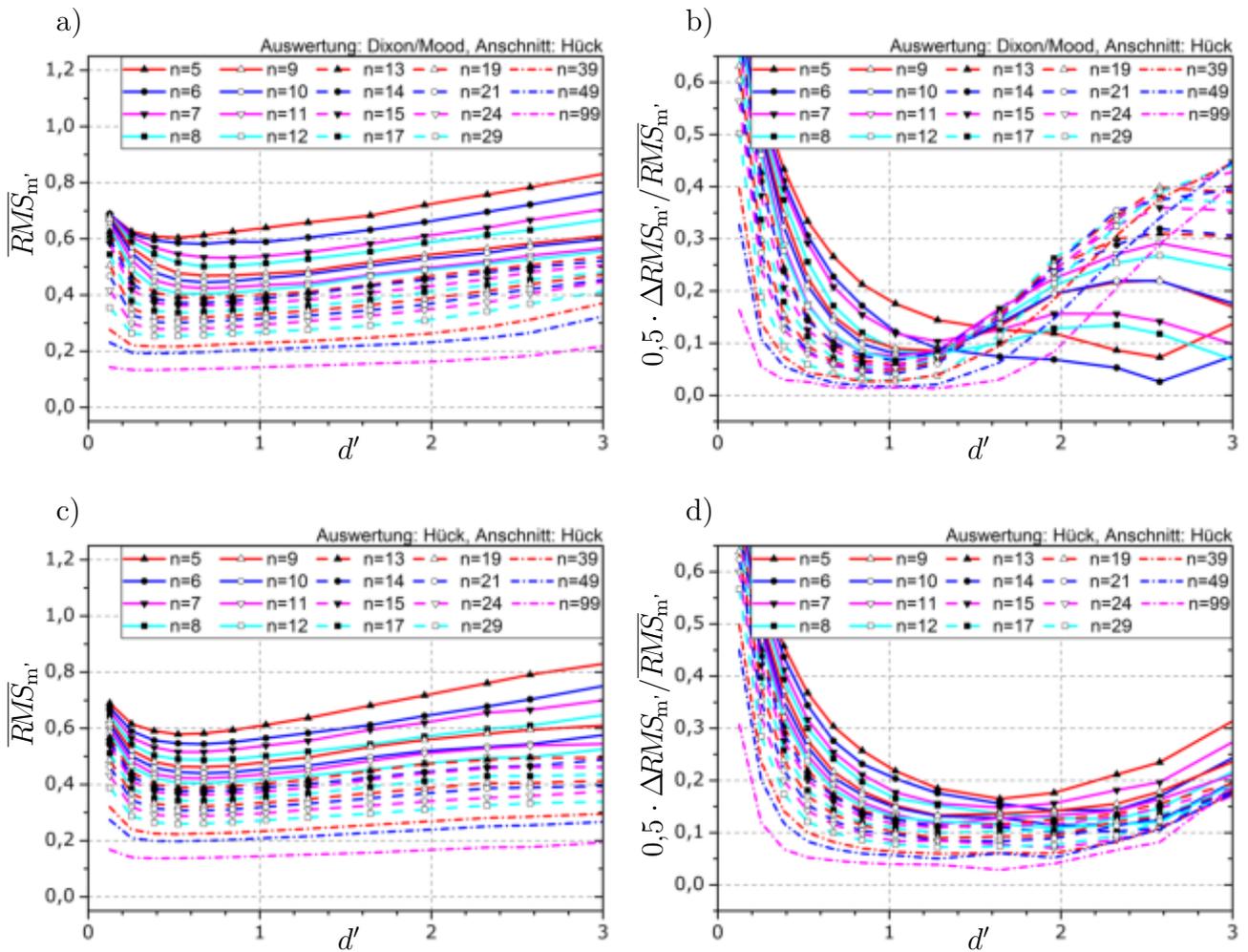


Abb. A.22: Mittlerer RMS-Fehler $\overline{RMS}_{m'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{m'} / \overline{RMS}_{m'}$ des Schätzers m' für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für die Auswertemethode nach DIXON/MOOD in a) und b) sowie HÜCK in c) und d).

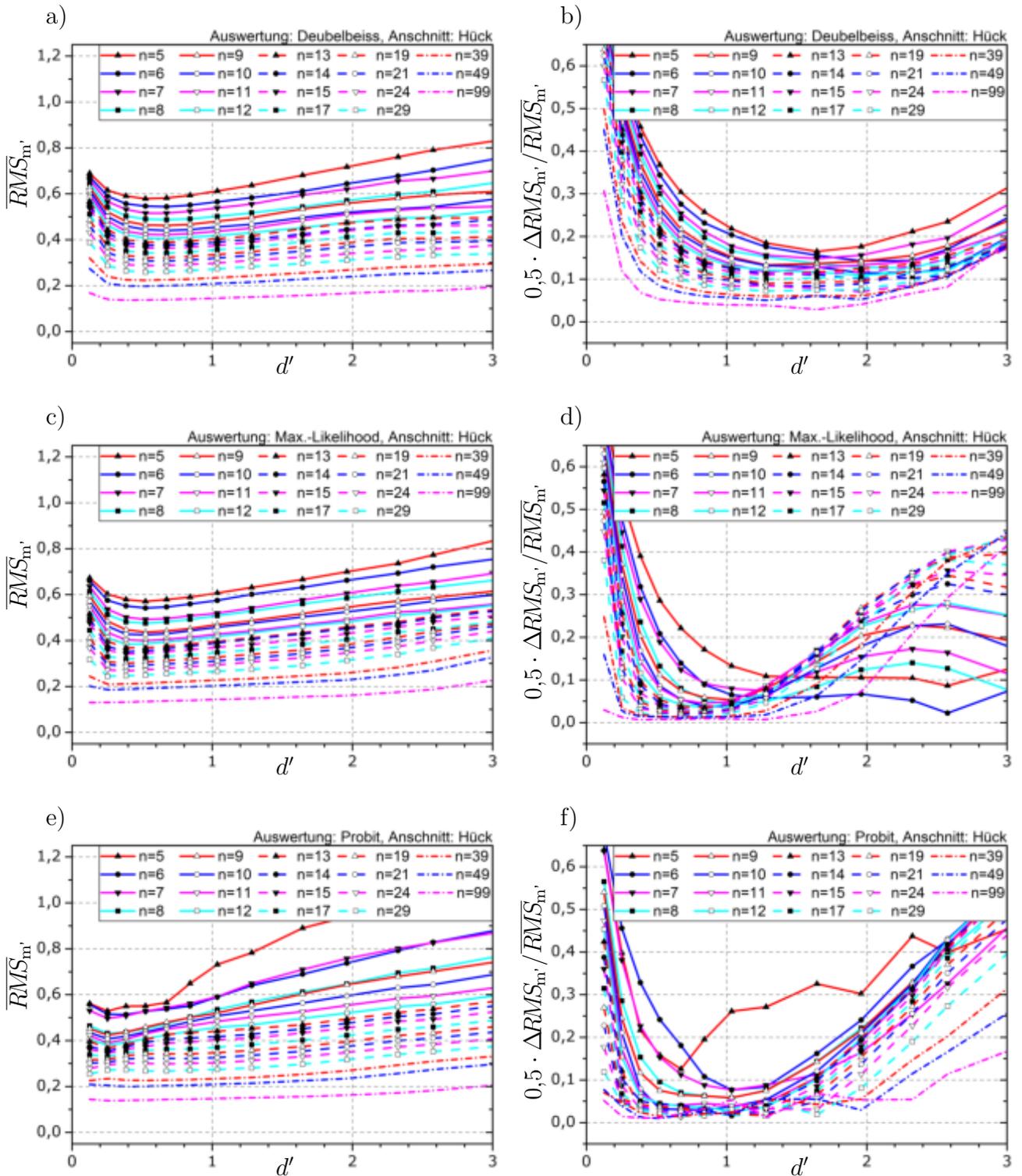


Abb. A.23: Mittlerer RMS-Fehler $\overline{RMS}_{m'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{m'} / \overline{RMS}_{m'}$ des Schätzers m' für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

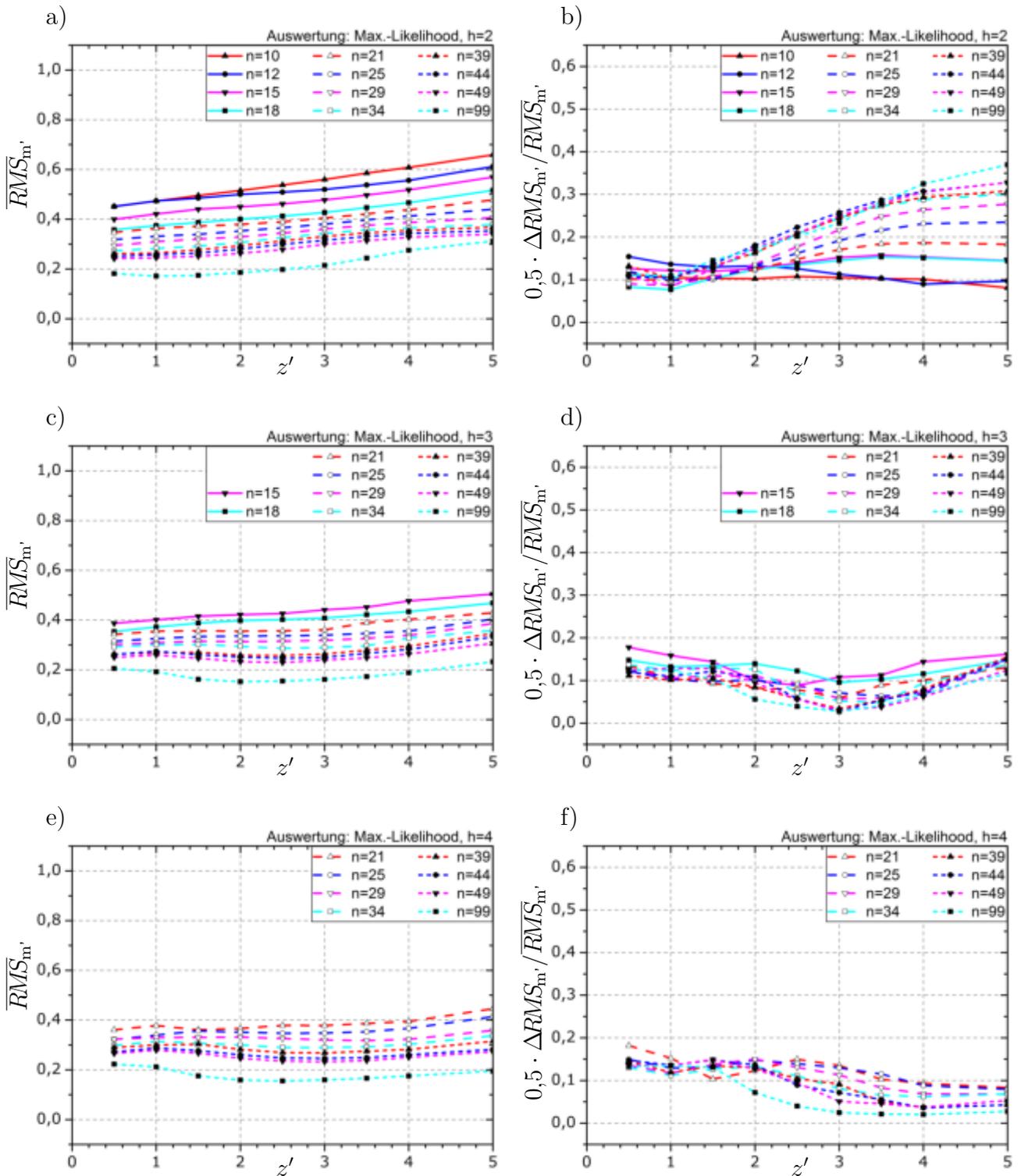


Abb. A.24: Mittlerer RMS-Fehler $\overline{RMS}_{m'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{m'} / \overline{RMS}_{m'}$ des Schätzers m' für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

A.2 Erwartungswertschätzung bei bekannter Standardabweichung

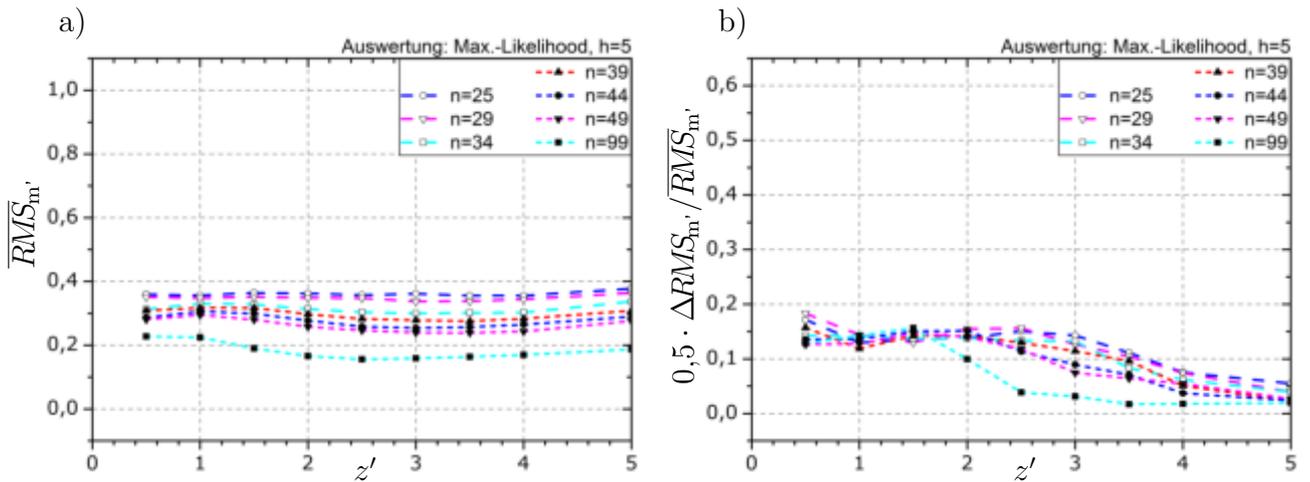


Abb. A.25: Mittlerer RMS-Fehler $\overline{RMS}_{m'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{m'} / \overline{RMS}_{m'}$ des Schätzers m' für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

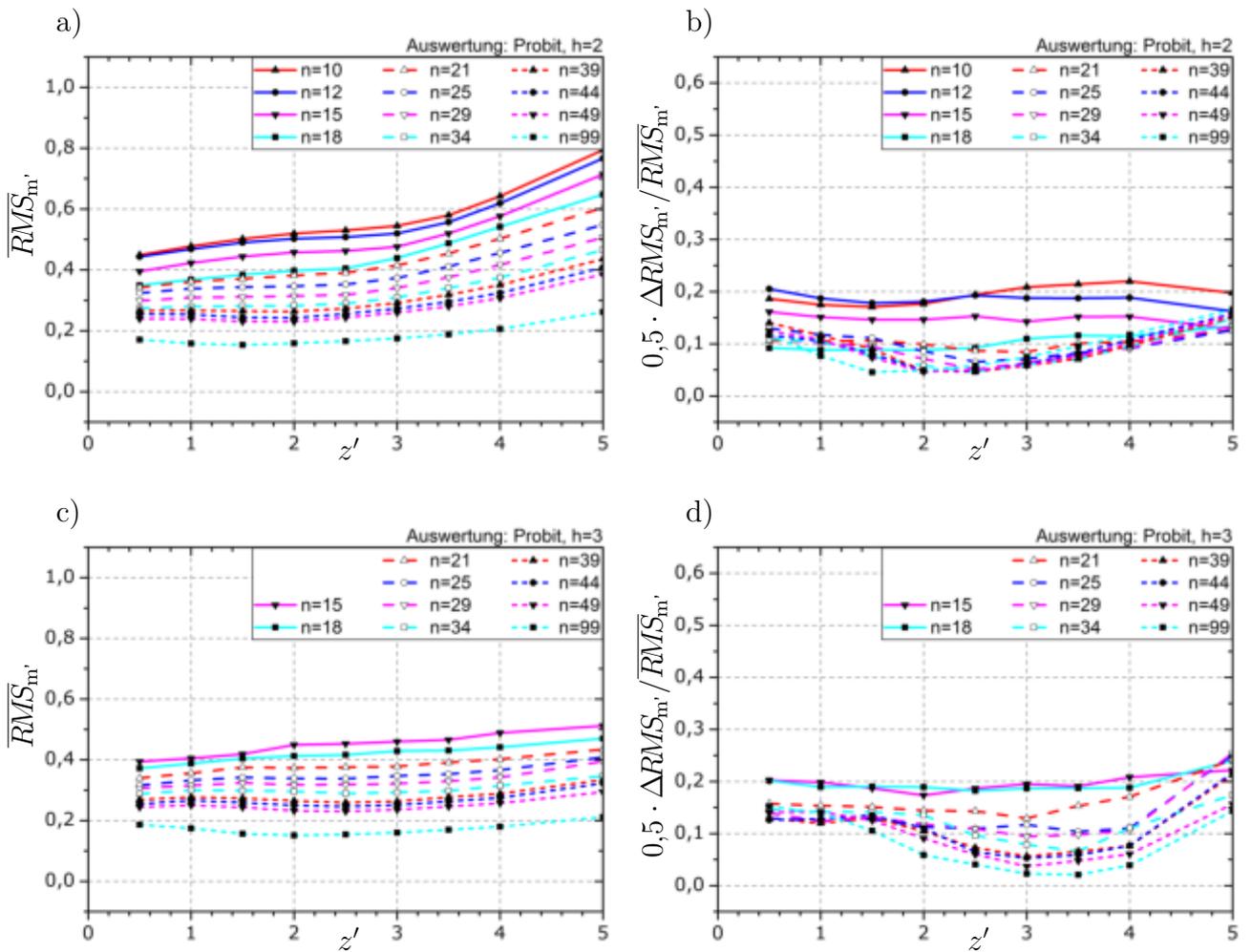


Abb. A.26: Mittlerer RMS-Fehler $\overline{RMS}_{m'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{m'} / \overline{RMS}_{m'}$ des Schätzers m' für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

A Grafische Darstellung der Simulationsergebnisse

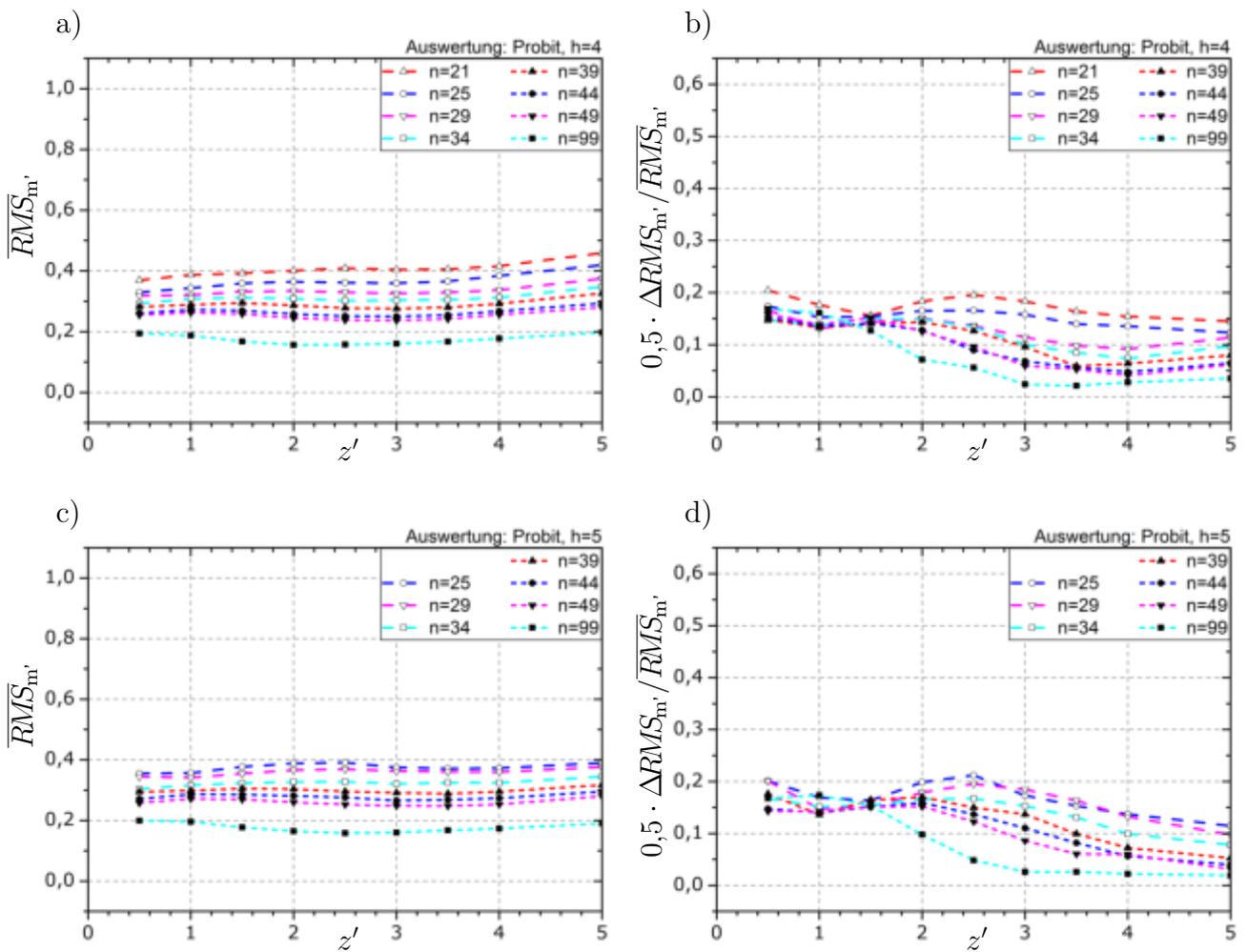


Abb. A.27: Mittlerer RMS-Fehler $\overline{RMS}_{m'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{m'} / \overline{RMS}_{m'}$ des Schätzers m' für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.2.4 Mittlere verallgemeinerte 5%-Quantile des Schätzers m'

Simulationsergebnisse der Treppenstufenversuche

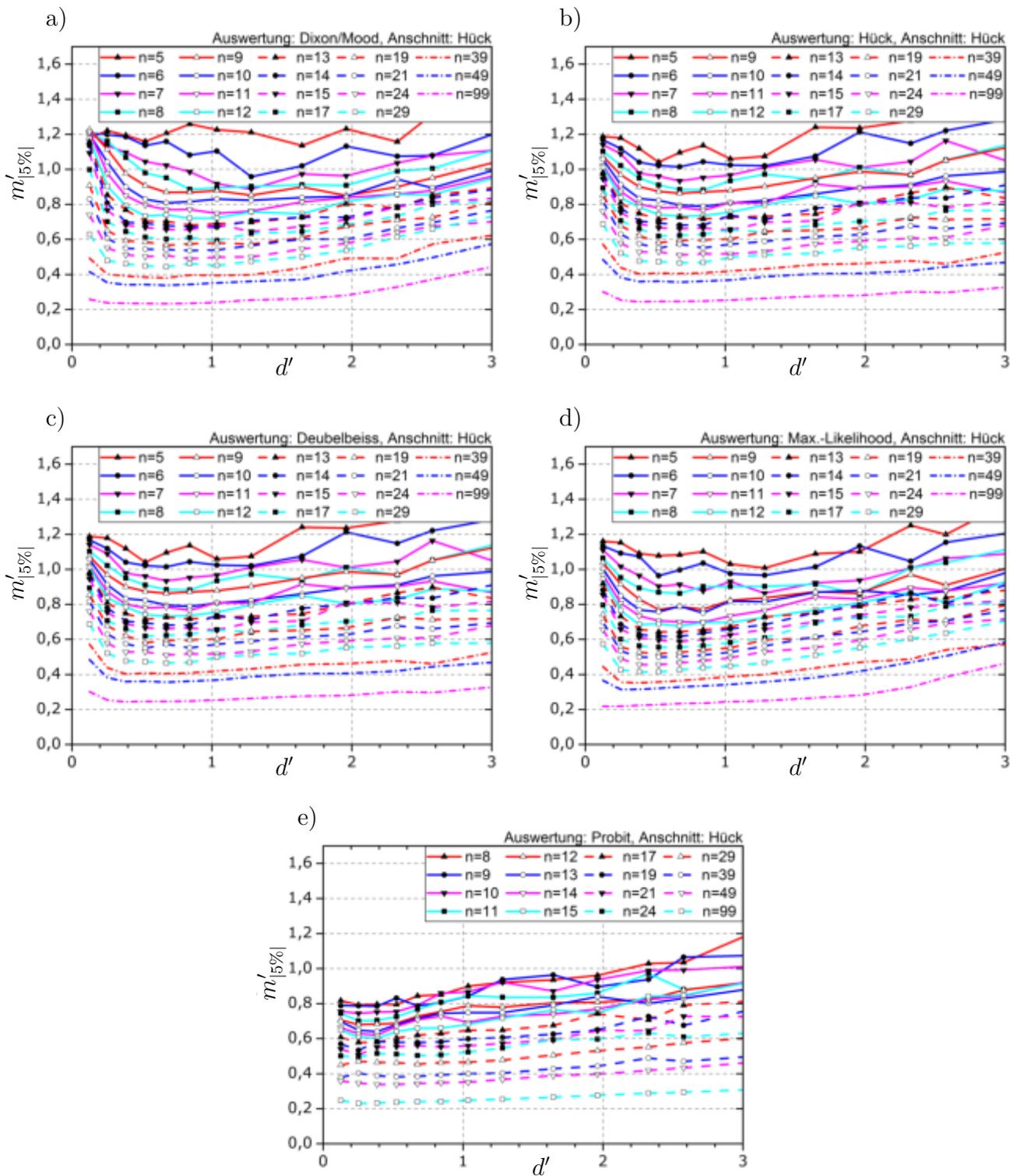


Abb. A.28: Mittlere verallgemeinerte 5%-Quantile $m'_{|5\%|}$ des Schätzers m' für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK. Für die Auswertemethode nach DIXON/MOOD in a), nach HÜCK in b), nach DEUBELBEISS in c), MAX.-LIKELIHOOD-METHODE in d) sowie PROBIT-METHODE in e).

Simulationsergebnisse der Horizontversuche

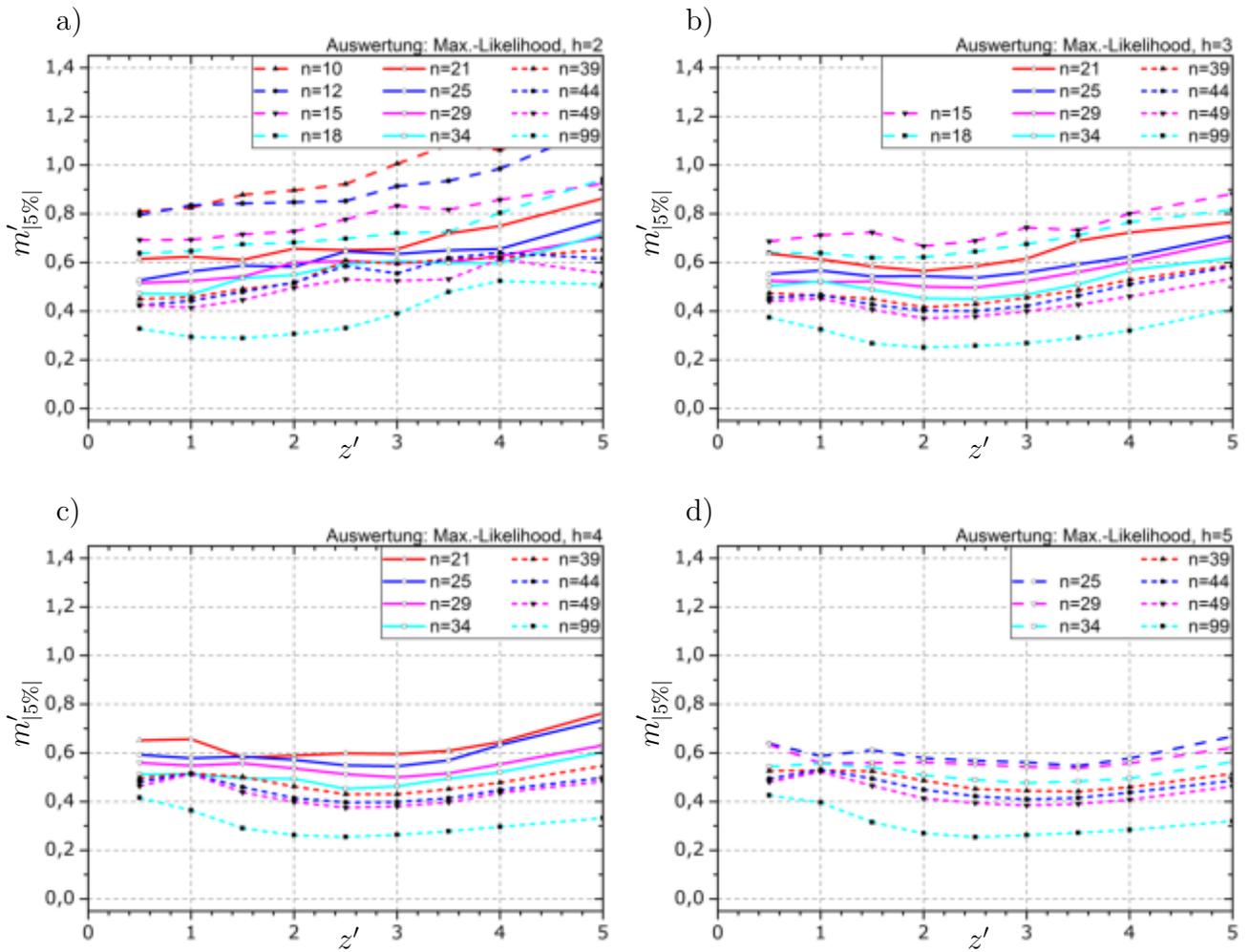


Abb. A.29: Mittlere verallgemeinerte 5%-Quantile $m'_{|5\%}$ des Schätzers m' für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuchen auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

A.2 Erwartungswertschätzung bei bekannter Standardabweichung

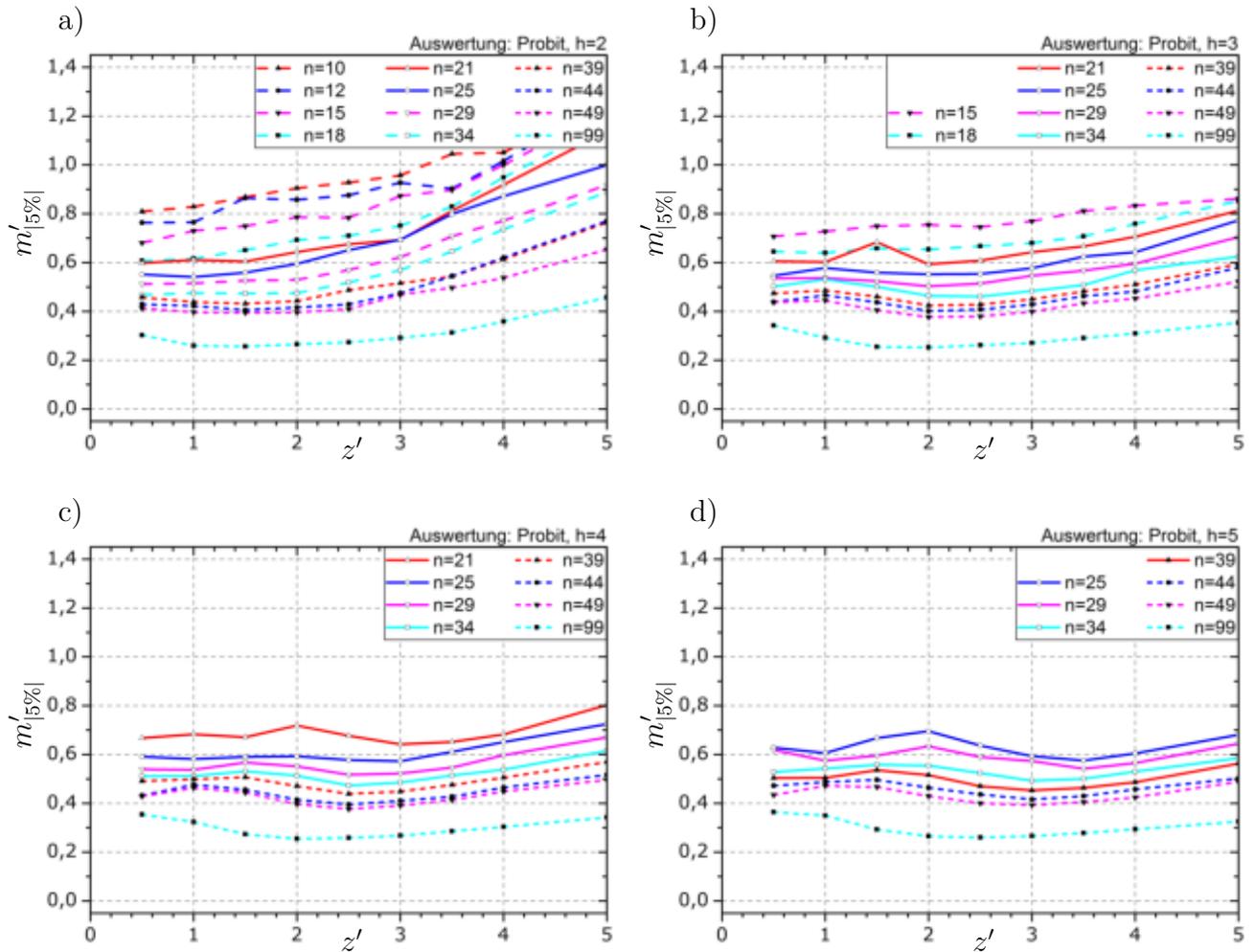


Abb. A.30: Mittlere verallgemeinerte 5%-Quantile $m'_{|5%|}$ des Schätzers m' für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

A.3 Standardabweichung

A.3.1 Mittlere Verzerrungen und Standardfehler des Schätzers s'_{lg}

Simulationsergebnisse der Treppenstufenversuche

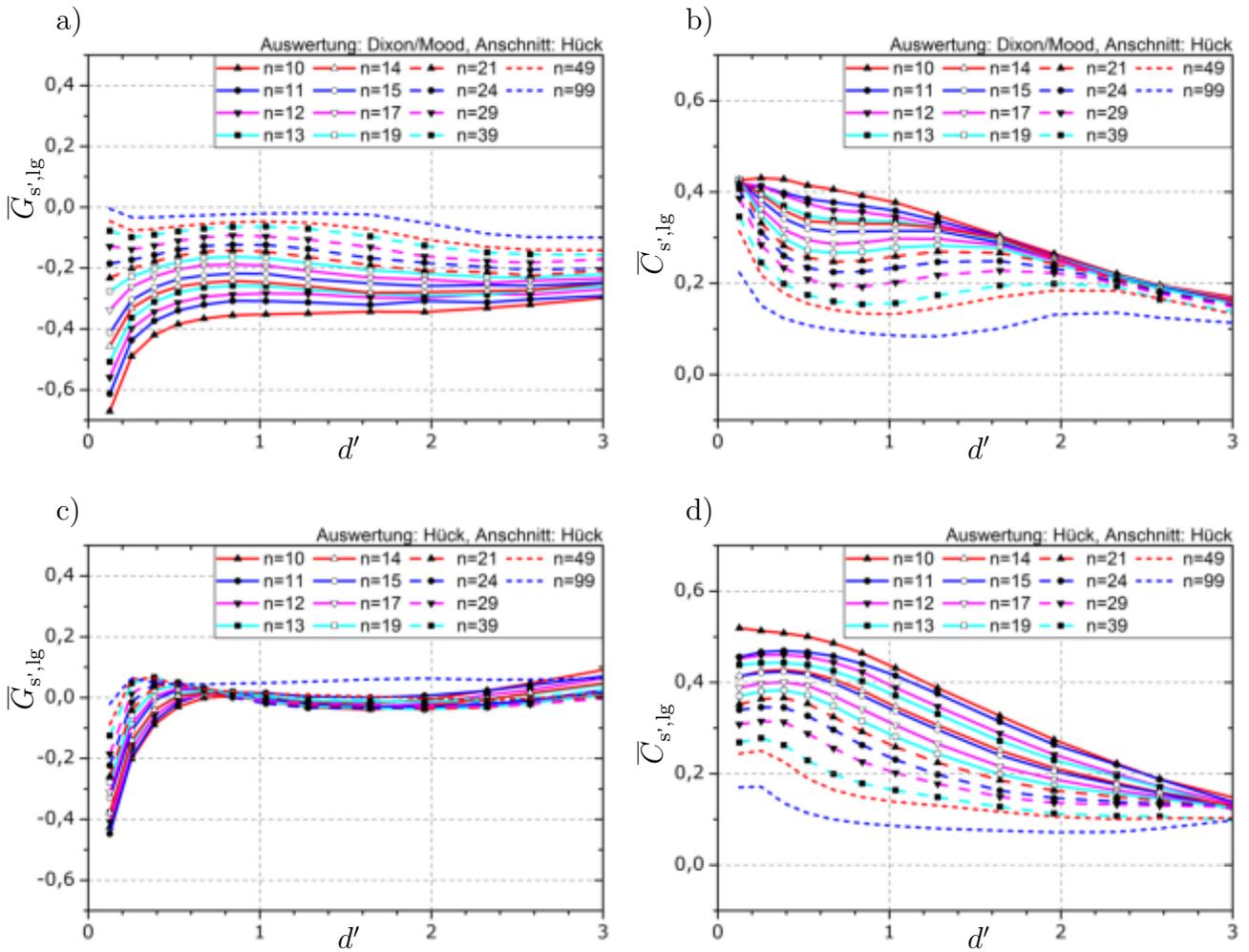


Abb. A.31: Mittlere Verzerrungen und Standardfehler des Schätzers s'_{lg} für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethode nach DIXON/MOOD in a) und b) sowie HÜCK in c) und d).

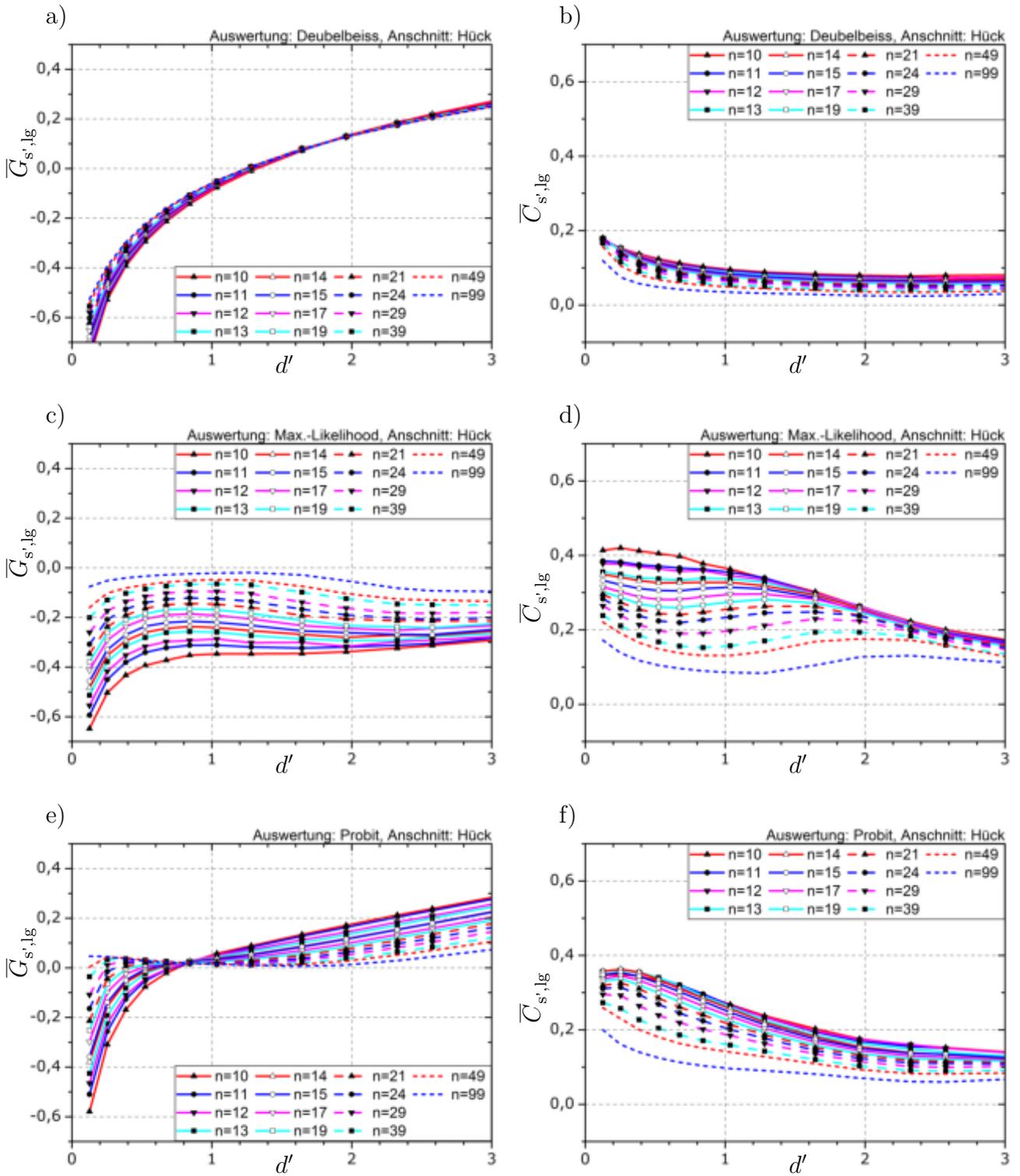


Abb. A.32: Mittlere Verzerrungen und Standardfehler des Schätzers s'_{lg} für Treppenstufenversuche mit Ansnchnittsdefinition nach HÜCK, für Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

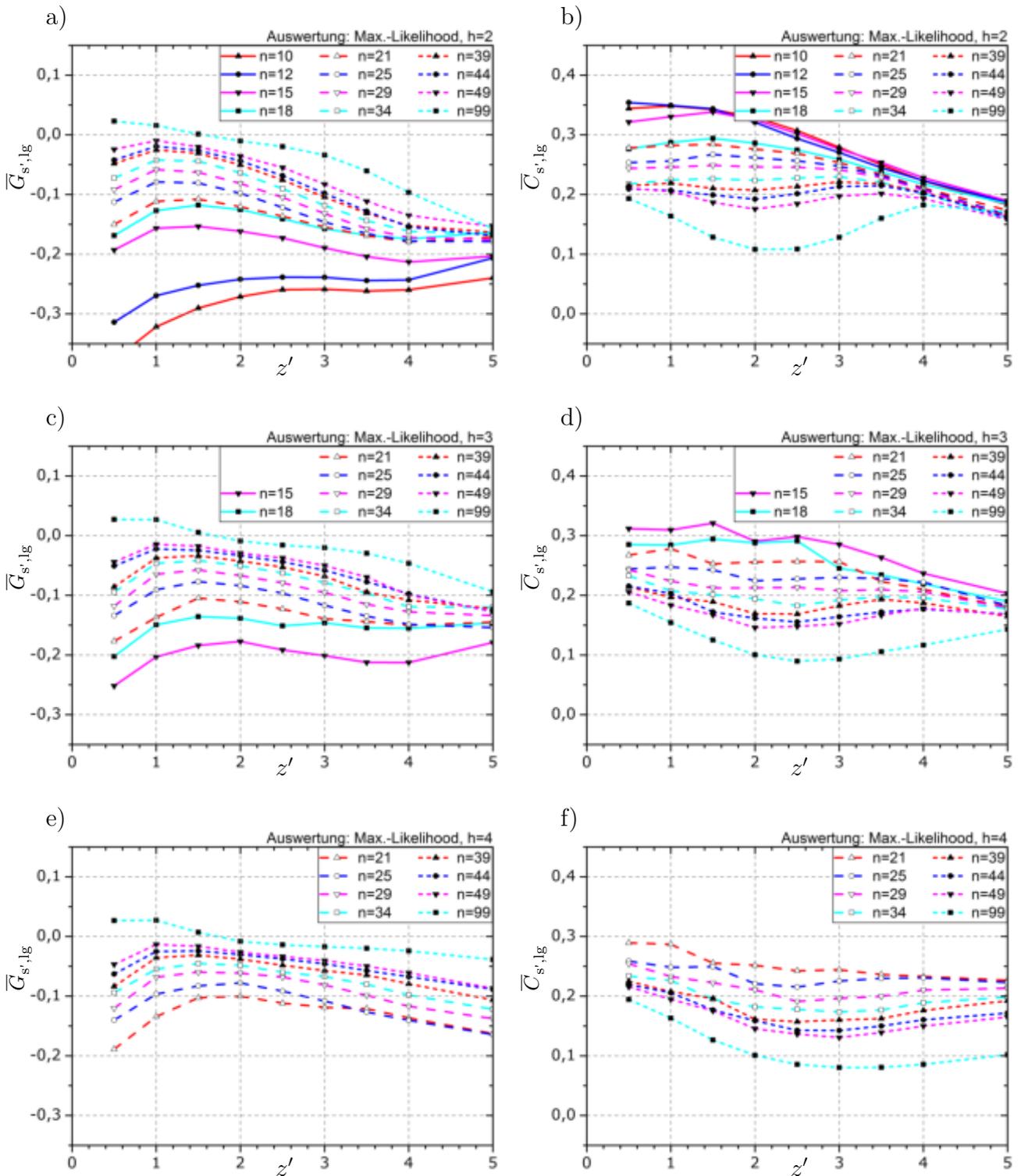


Abb. A.33: Mittlere Verzerrungen und Standardfehler des Schätzers s'_{lg} für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

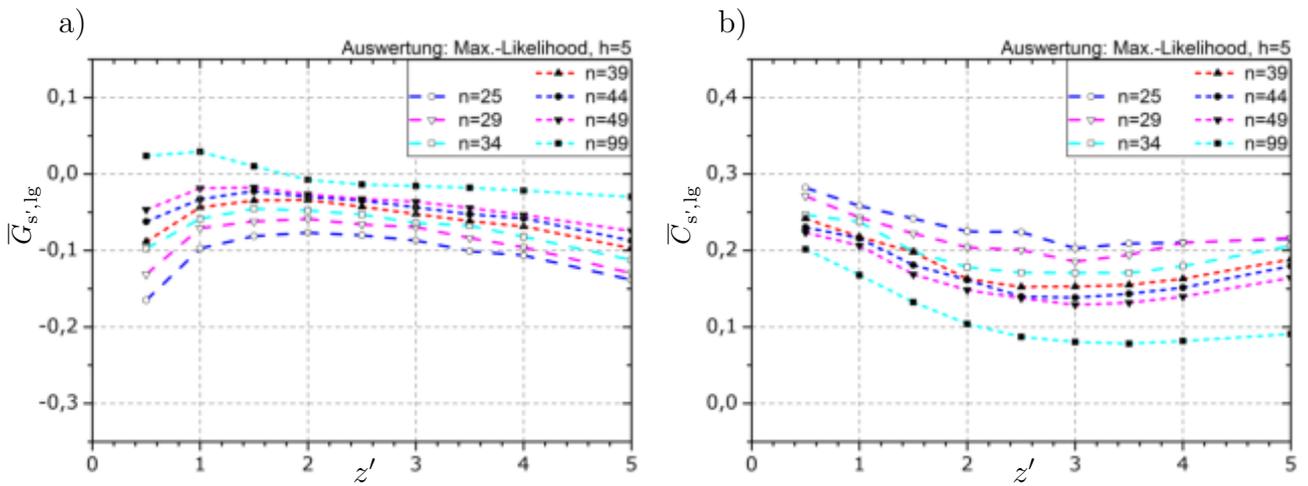


Abb. A.34: Mittlere Verzerrungen und Standardfehler des Schätzers s'_{lg} für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

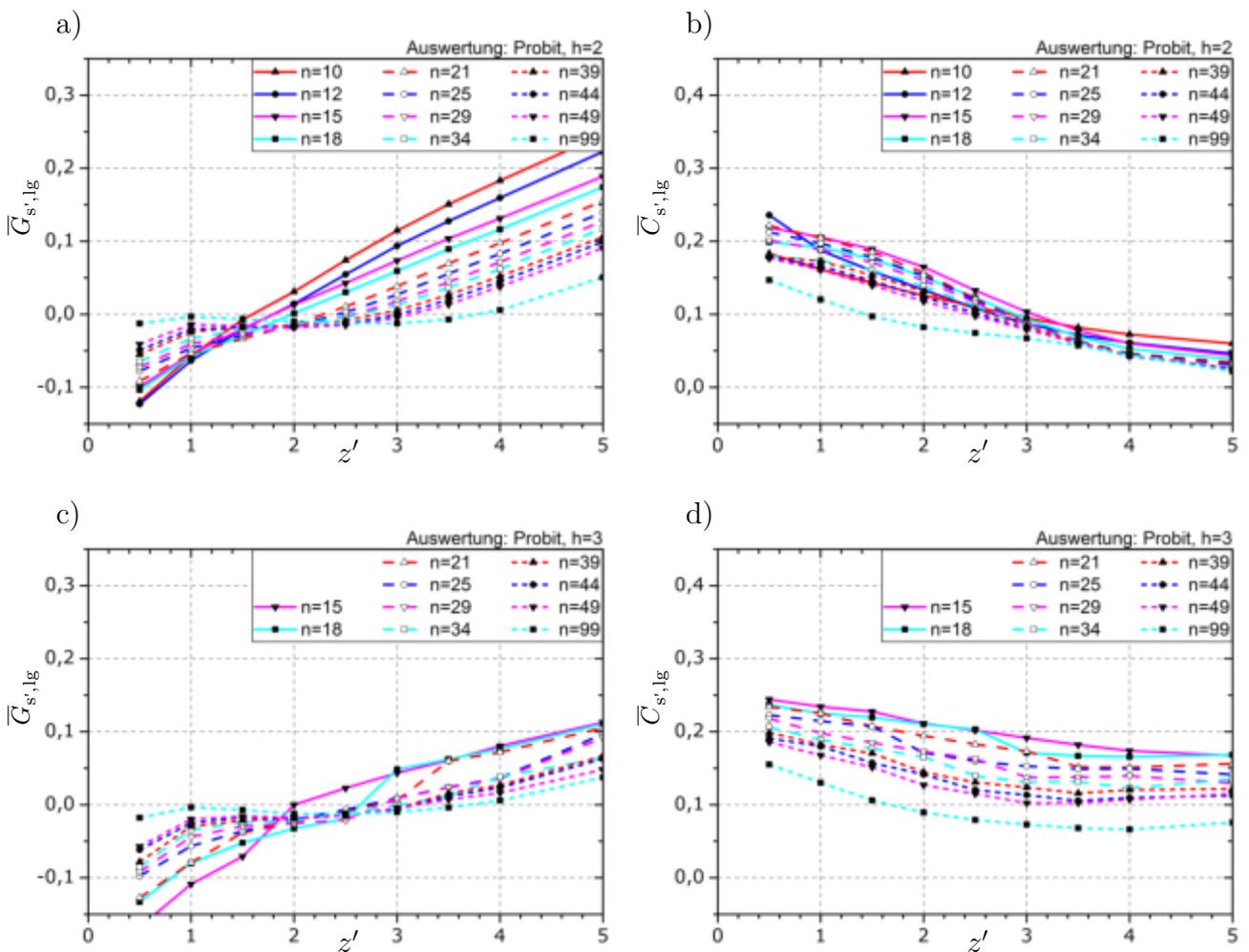


Abb. A.35: Mittlere Verzerrungen und Standardfehler des Schätzers s'_{lg} für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuch auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

A Grafische Darstellung der Simulationsergebnisse

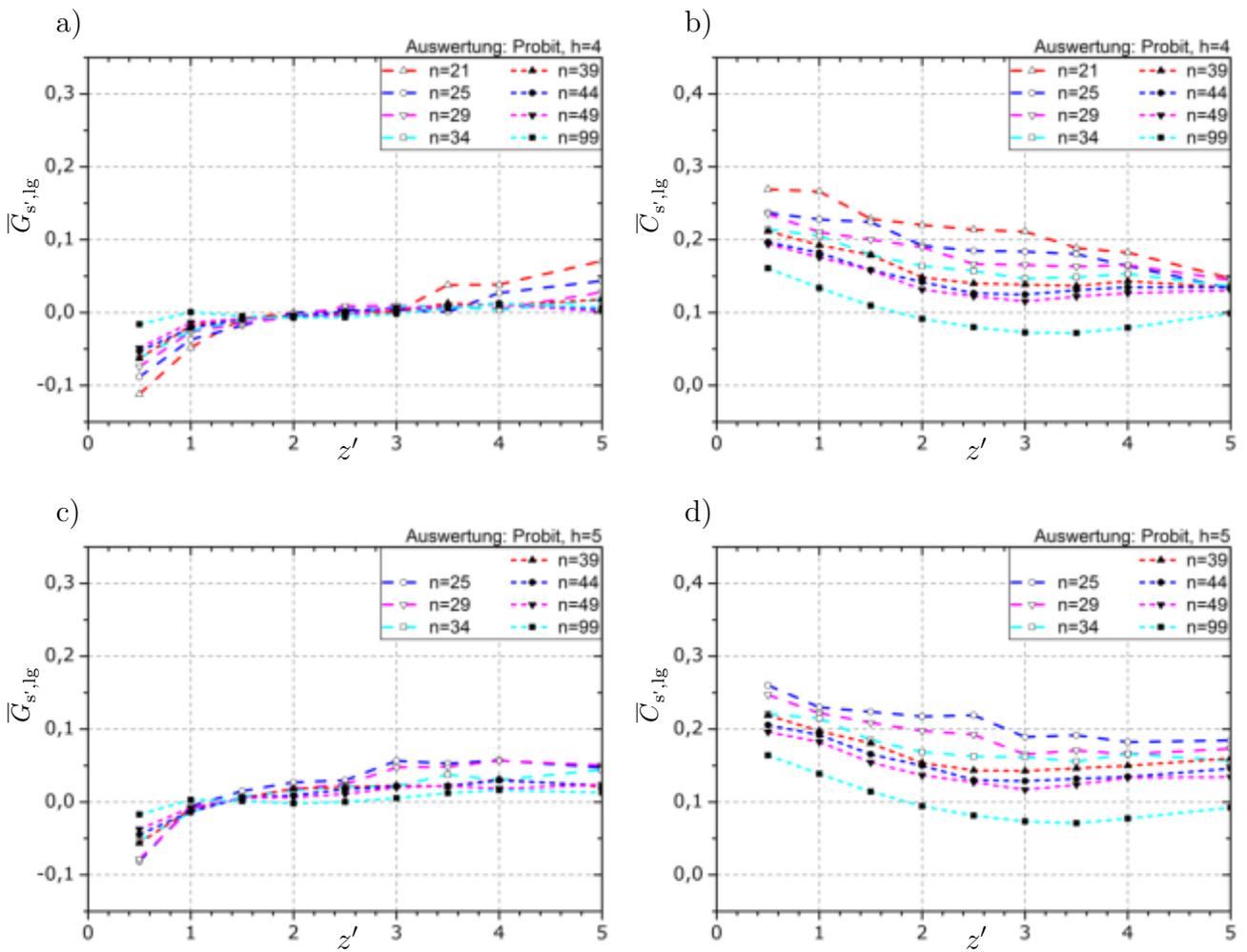


Abb. A.36: Mittlere Verzerrungen und Standardfehler des Schätzers s'_{lg} für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.3.2 RMS-Fehler des Schätzers s'_{lg}

Simulationsergebnisse der Treppenstufenversuche

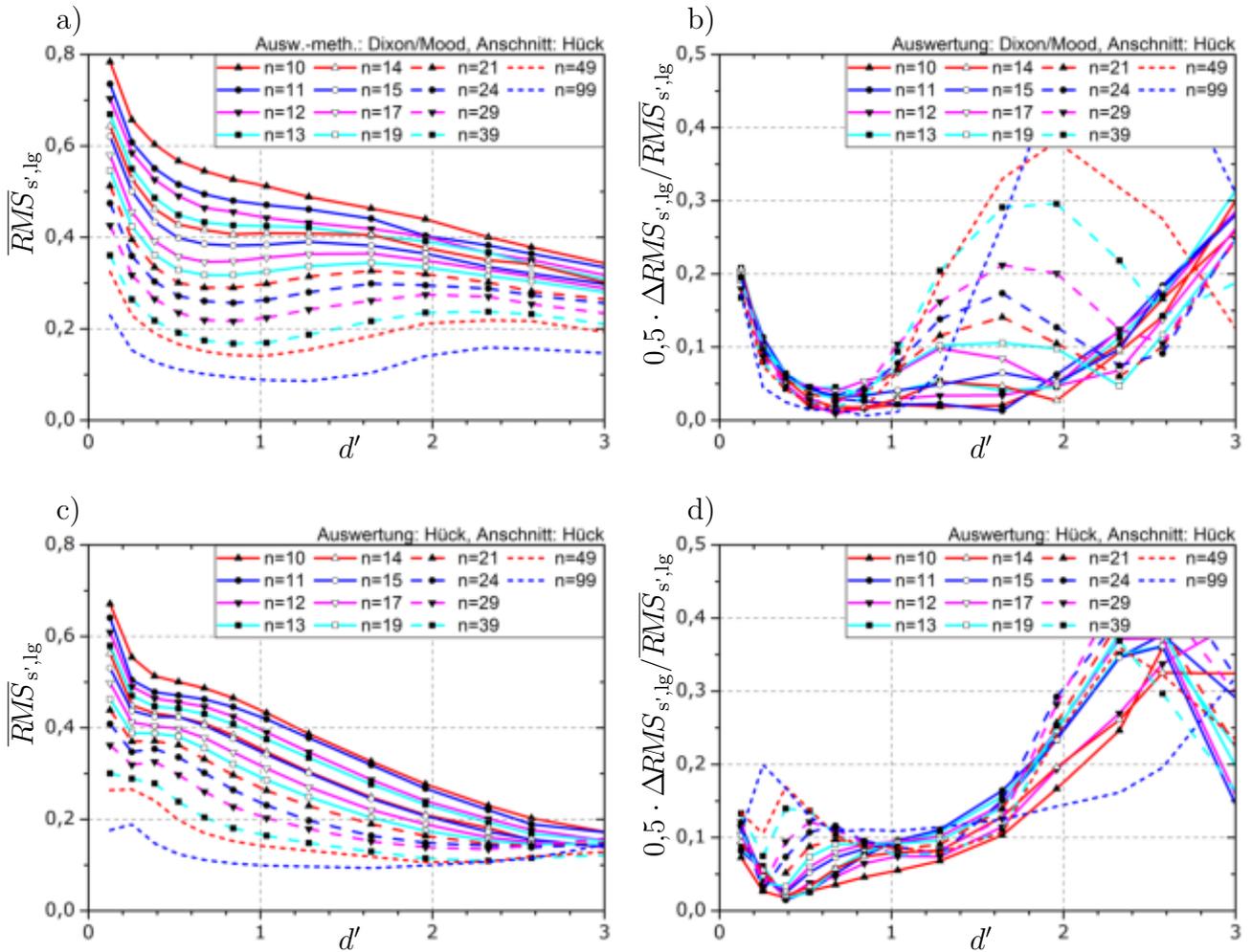


Abb. A.37: Mittlerer RMS-Fehler $\overline{RMS}_{s',lg}$ und relative Abweichung $0,5 \cdot \Delta RMS_{s',lg} / \overline{RMS}_{s',lg}$ des Schätzers s'_{lg} für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für die Auswertemethode nach DIXON/MOOD in a) und b) sowie HÜCK in c) und d)

A Grafische Darstellung der Simulationsergebnisse

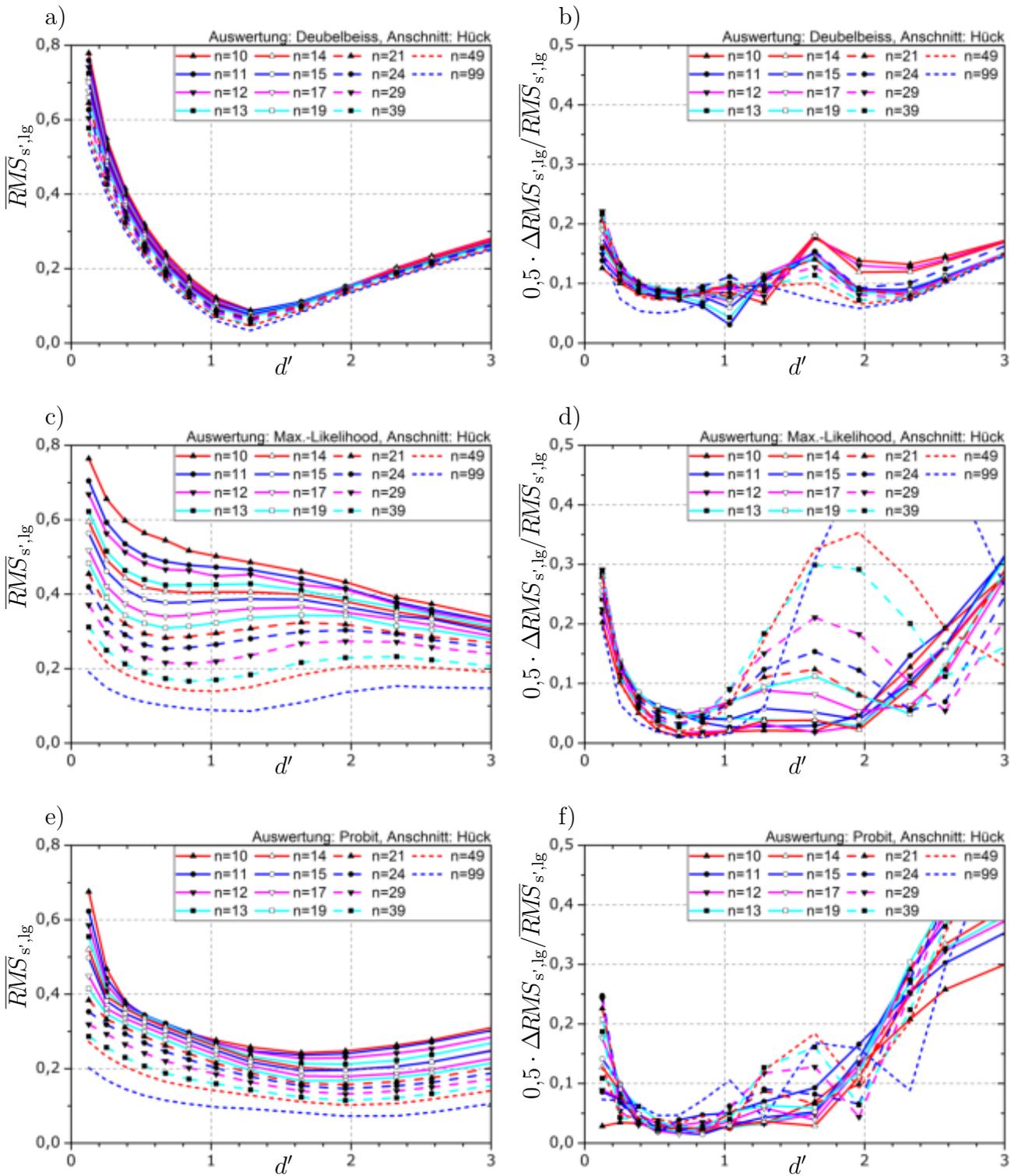


Abb. A.38: Mittlerer RMS-Fehler $\overline{RMS}_{s',lg}$ und relative Abweichung $0,5 \cdot \Delta RMS_{s',lg} / \overline{RMS}_{s',lg}$ des Schätzers s'_{lg} für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für die Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

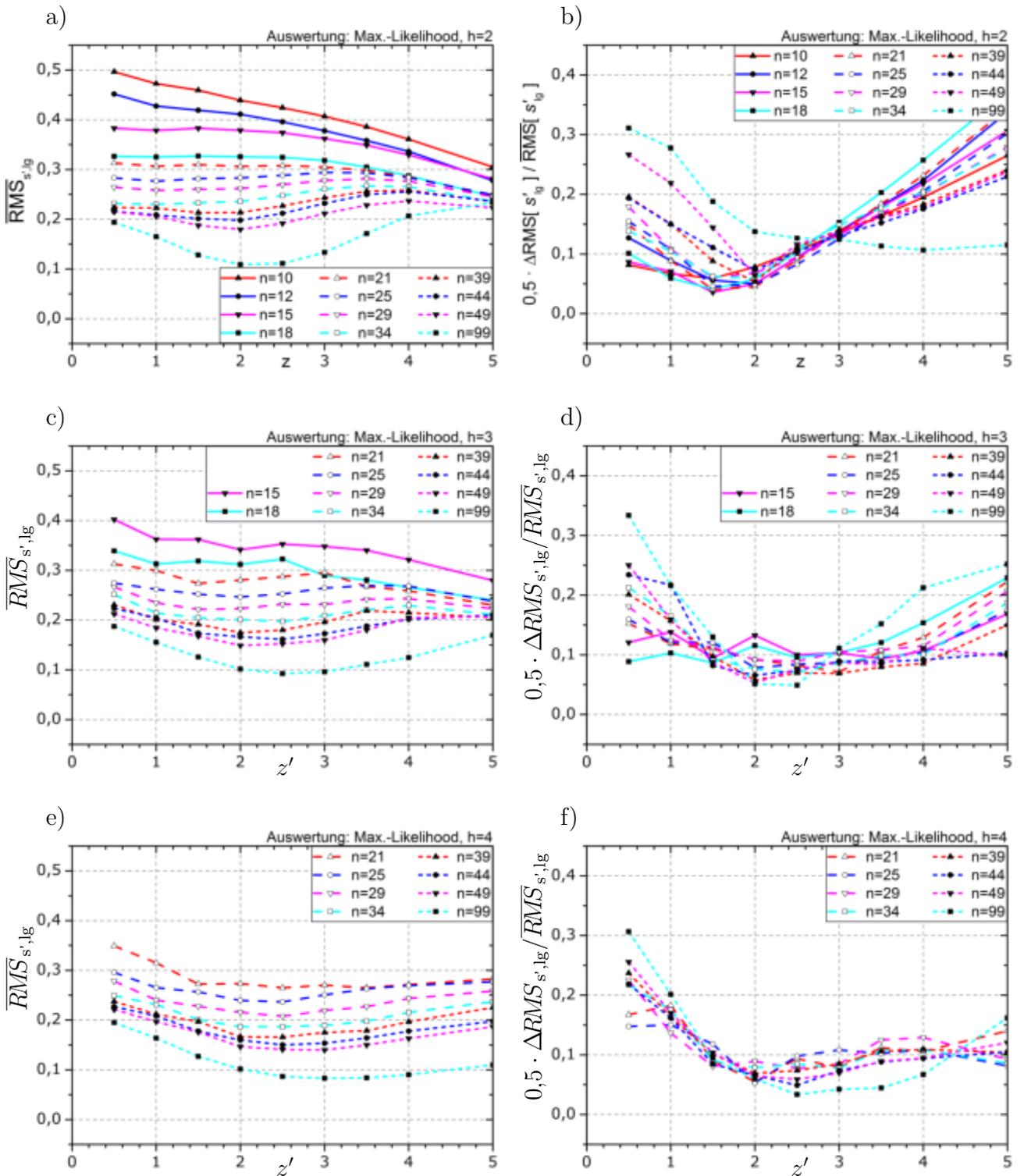


Abb. A.39: Mittlerer RMS-Fehler $\overline{RMS}_{s',lg}$ und relative Abweichung $0,5 \cdot \Delta RMS_{s',lg} / \overline{RMS}_{s',lg}$ des Schätzers s'_{lg} für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

A Grafische Darstellung der Simulationsergebnisse

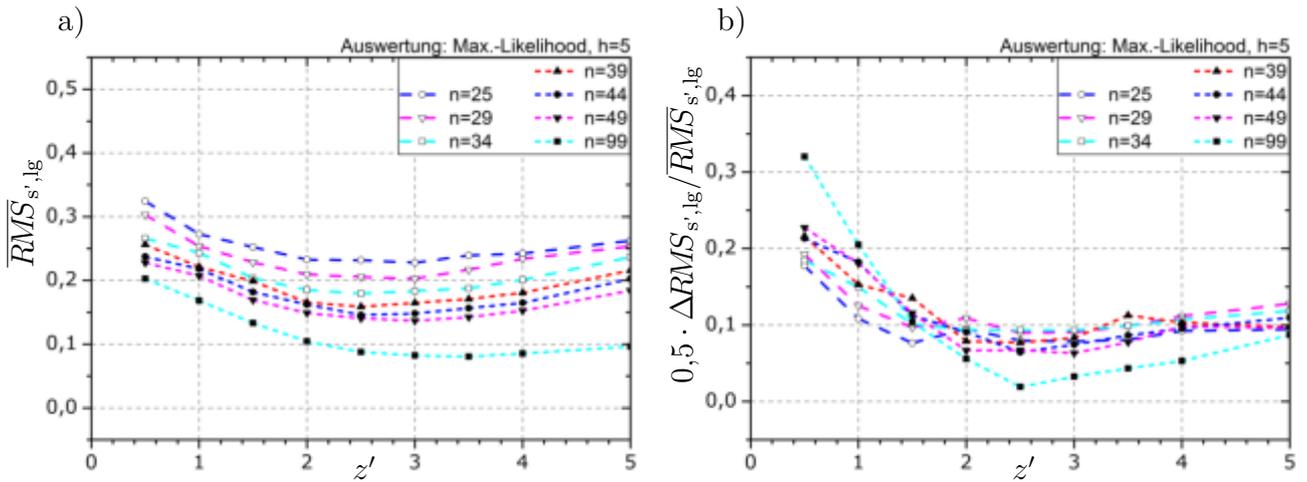


Abb. A.40: Mittlerer RMS-Fehler $\overline{RMS}_{s',lg}$ und relative Abweichung $0,5 \cdot \Delta \overline{RMS}_{s',lg} / \overline{RMS}_{s',lg}$ des Schätzers s'_{lg} für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

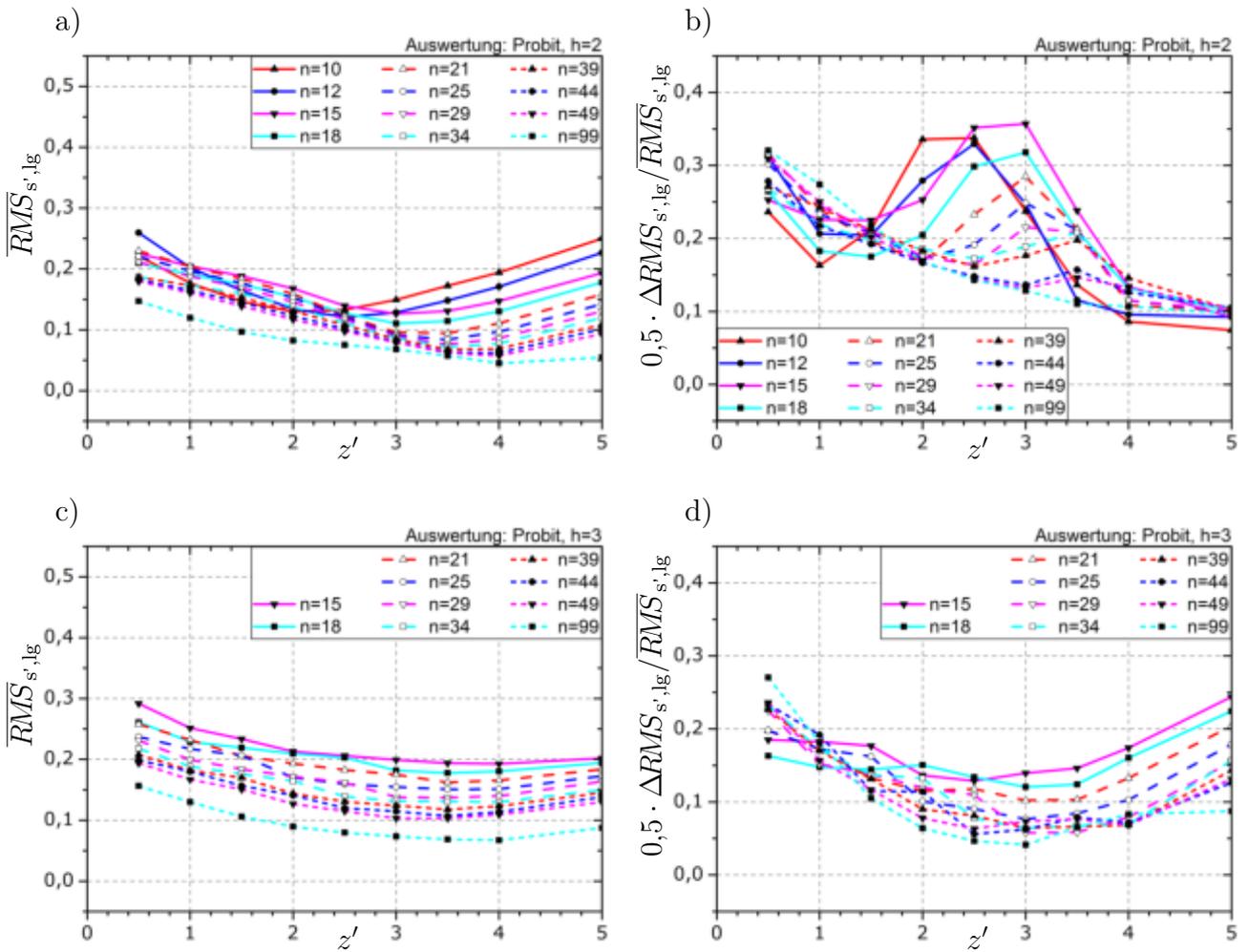


Abb. A.41: Mittlerer RMS-Fehler $\overline{RMS}_{s',lg}$ und relative Abweichung $0,5 \cdot \Delta \overline{RMS}_{s',lg} / \overline{RMS}_{s',lg}$ des Schätzers s'_{lg} für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

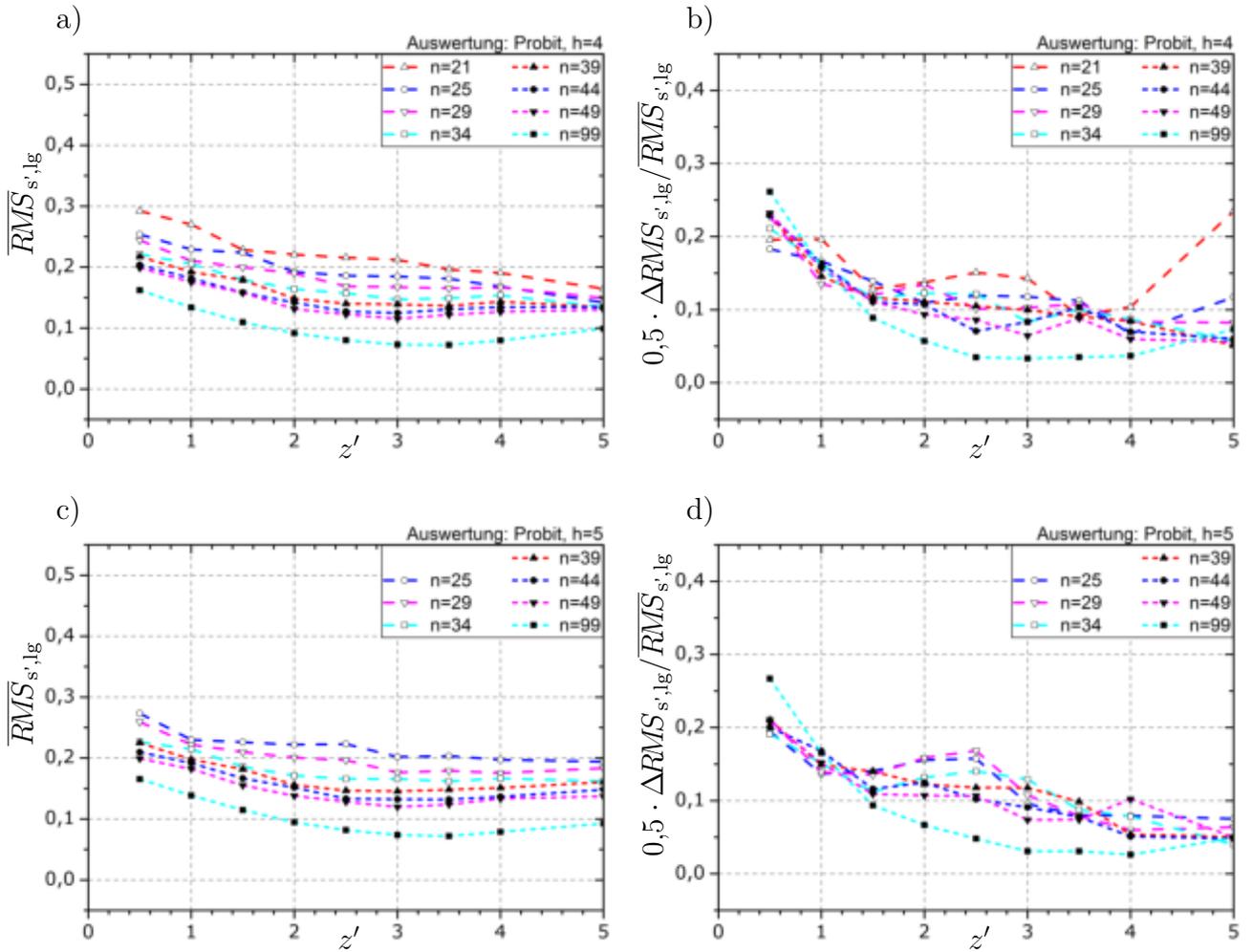


Abb. A.42: Mittlerer RMS-Fehler $\overline{RMS}_{s',lg}$ und relative Abweichung $0,5 \cdot \Delta RMS_{s',lg} / \overline{RMS}_{s',lg}$ des Schätzers s'_{lg} für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.3.3 Mittlere 5%-Quantile $\overline{s'}_{lg,5\%}$ des Schätzers s'_{lg}

Simulationsergebnisse der Treppenstufenversuche

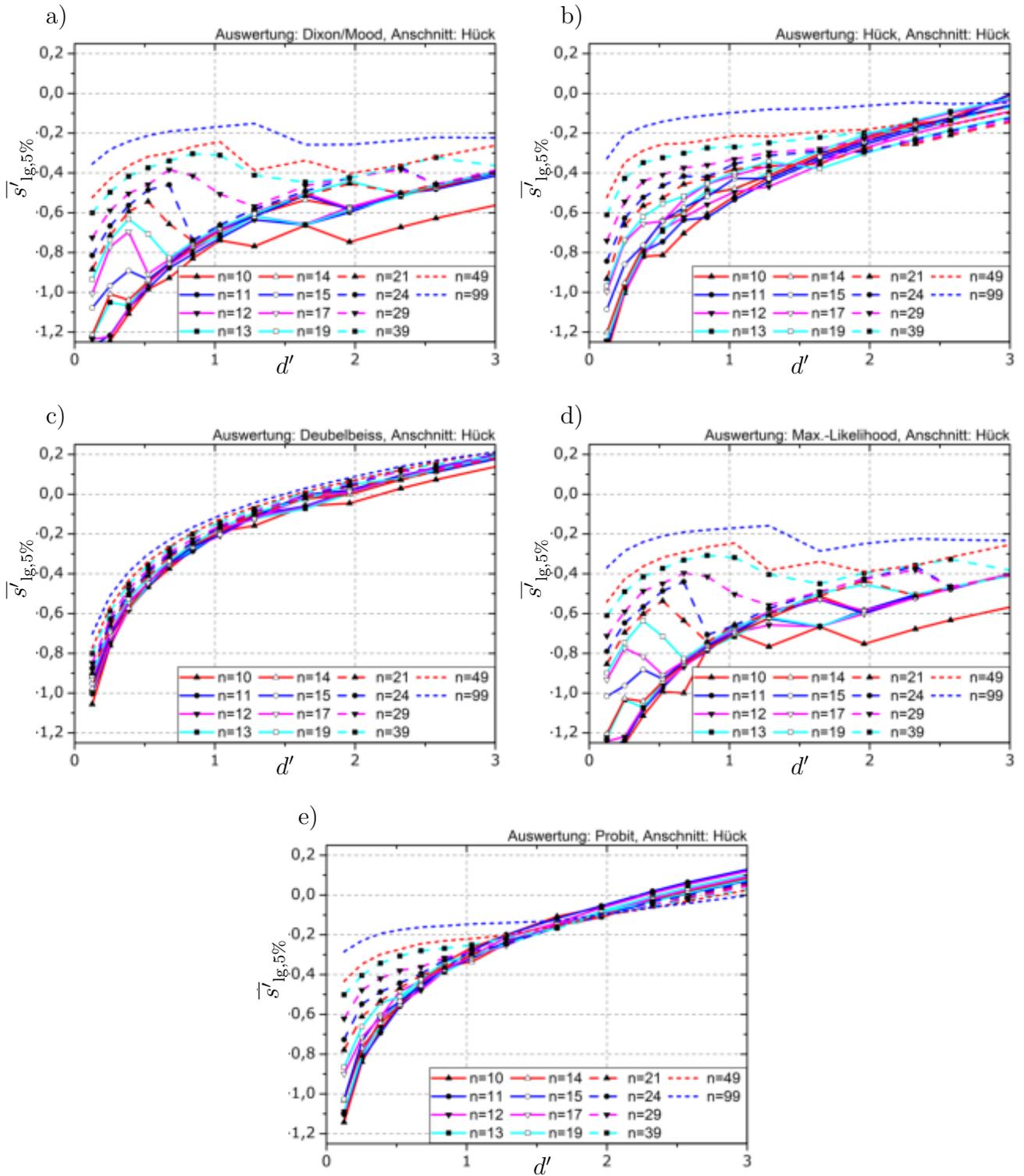


Abb. A.43: Mittlere 5%-Quantile $\overline{s'}_{lg,5\%}$ des Schätzers s'_{lg} aus Treppenstufenversuchen mit Anschnittsdefinition nach HÜCK. Für die Auswertemethode nach DIXON/MOOD in a), nach HÜCK in b), nach DEUBELBEISS in c), MAX.-LIKELIHOOD-METHODE in d) sowie PROBIT-METHODE in e).

Simulationsergebnisse der Horizontversuche

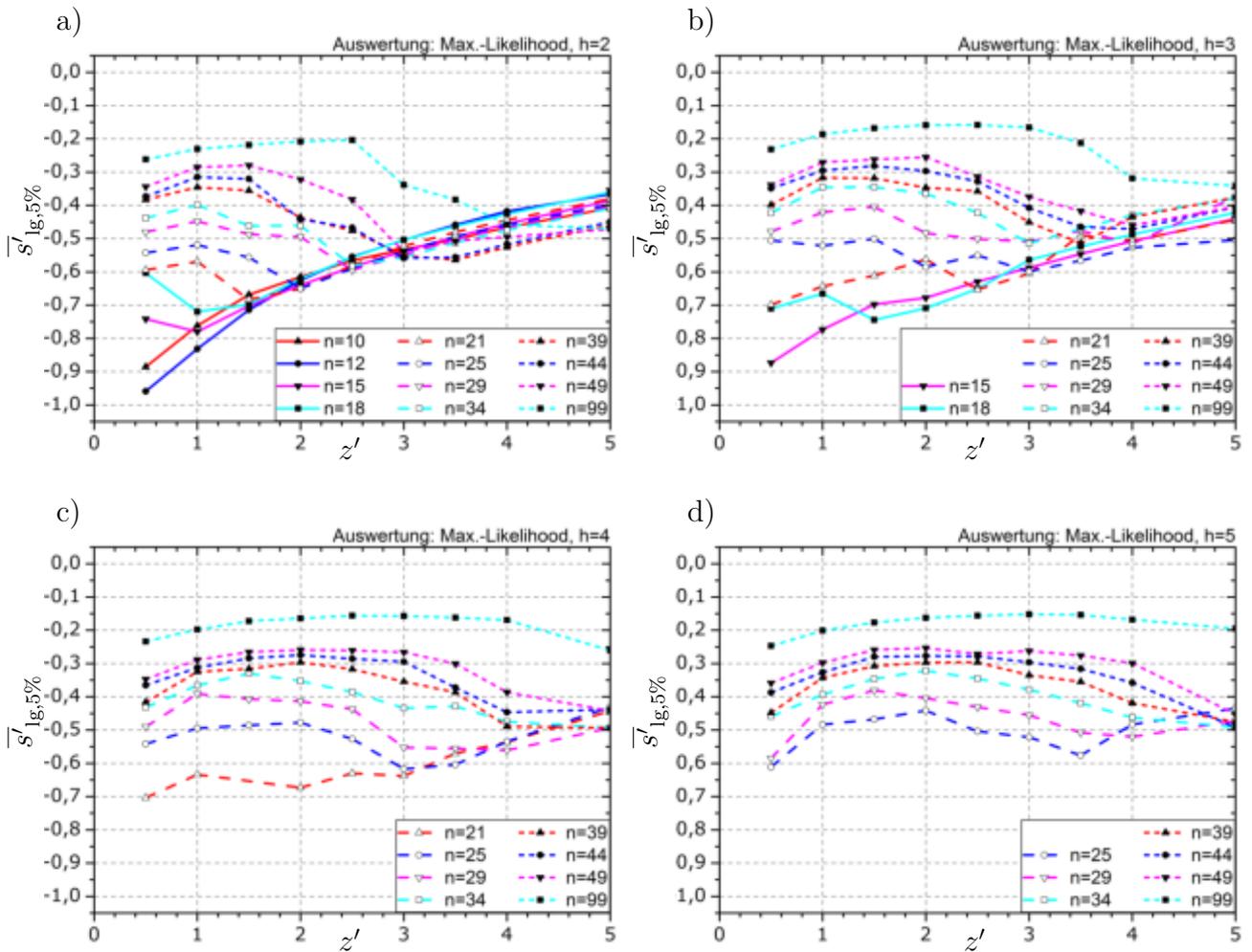


Abb. A.44: Mittlere 5%-Quantile $\overline{s'_{lg,5\%}}$ des Schätzers s'_{lg} aus Horizontversuchen für die Auswertung mittels MAX-LIKELIHOOD-METHODE und Versuchen auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

A Grafische Darstellung der Simulationsergebnisse

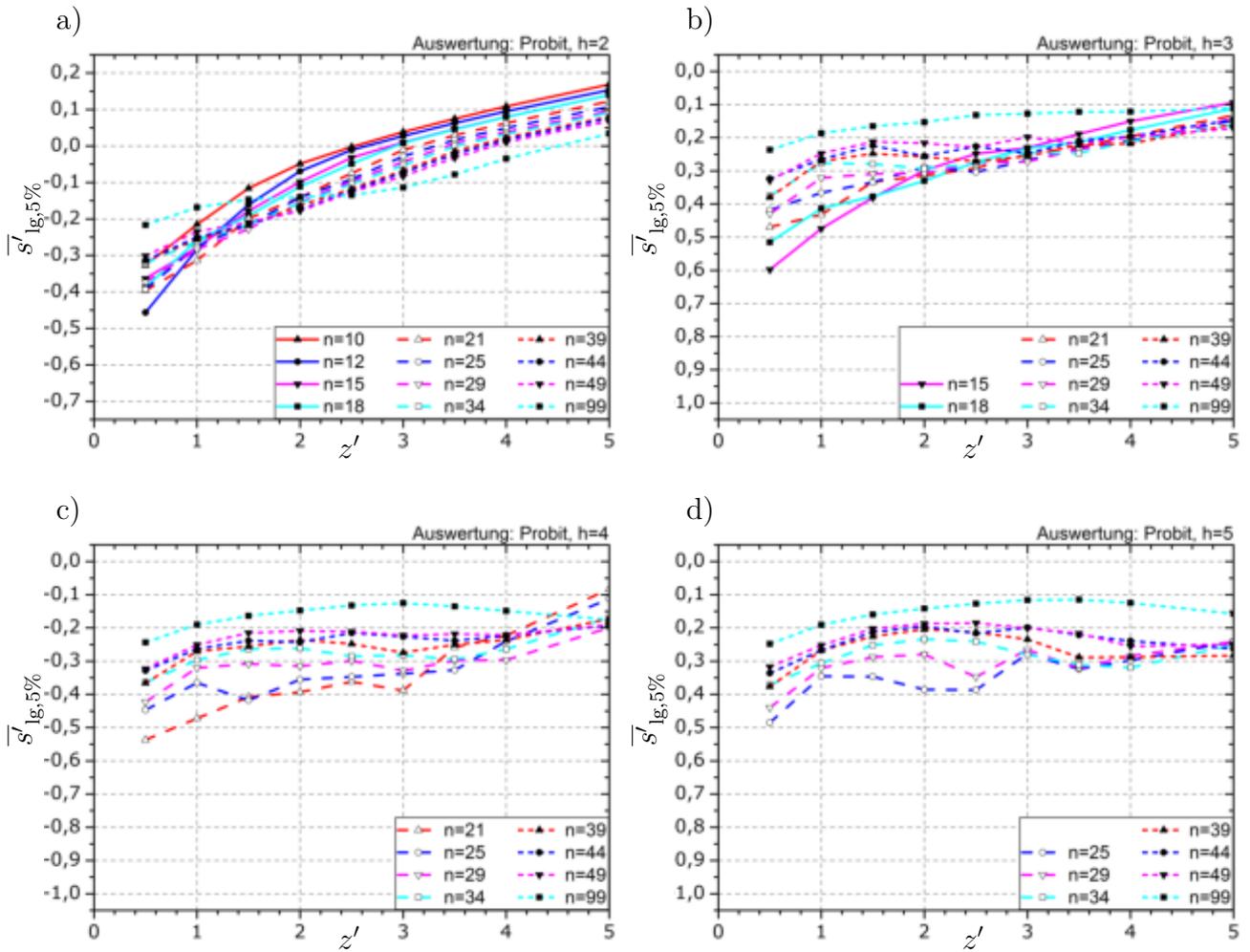


Abb. A.45: Mittlere 5%-Quantile $\overline{s'_{lg,5\%}}$ des Schätzers s'_{lg} aus Horizontversuchen für die Auswertung mittels PROBIT-METHODE und Versuchen auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

A.4 Erwartungswert bei unbekannter Standardabweichung

A.4.1 Mittlere Verzerrungen und Standardfehler des Schätzers m^*

Simulationsergebnisse der Treppenstufenversuche

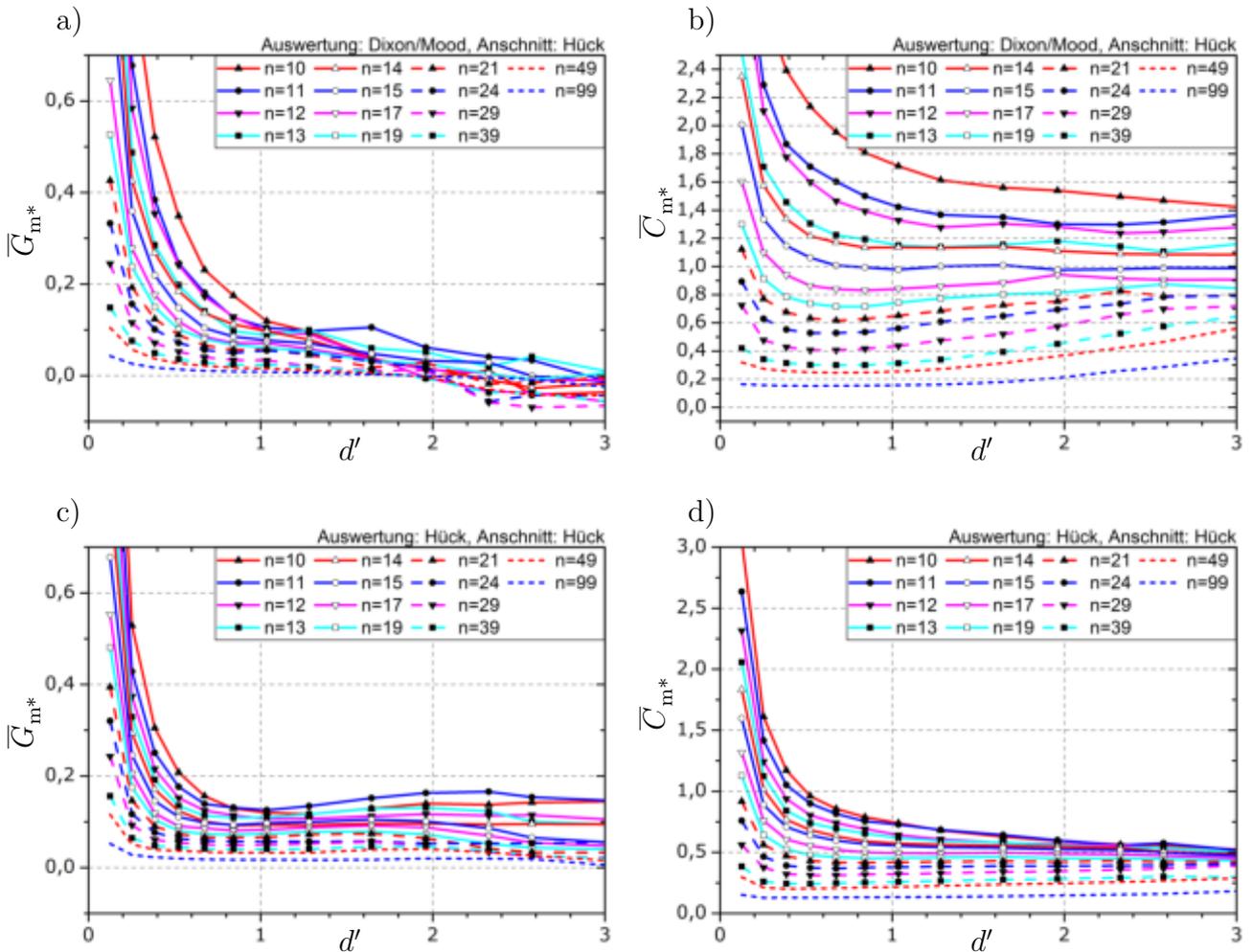


Abb. A.46: Mittlere Verzerrungen und Standardfehler des Schätzers m^* für Treppenstufenversuche mit Ansnchnittsdefinition nach HÜCK, für Auswertemethode nach DIXON/MOOD in a) und b) sowie HÜCK in c) und d).

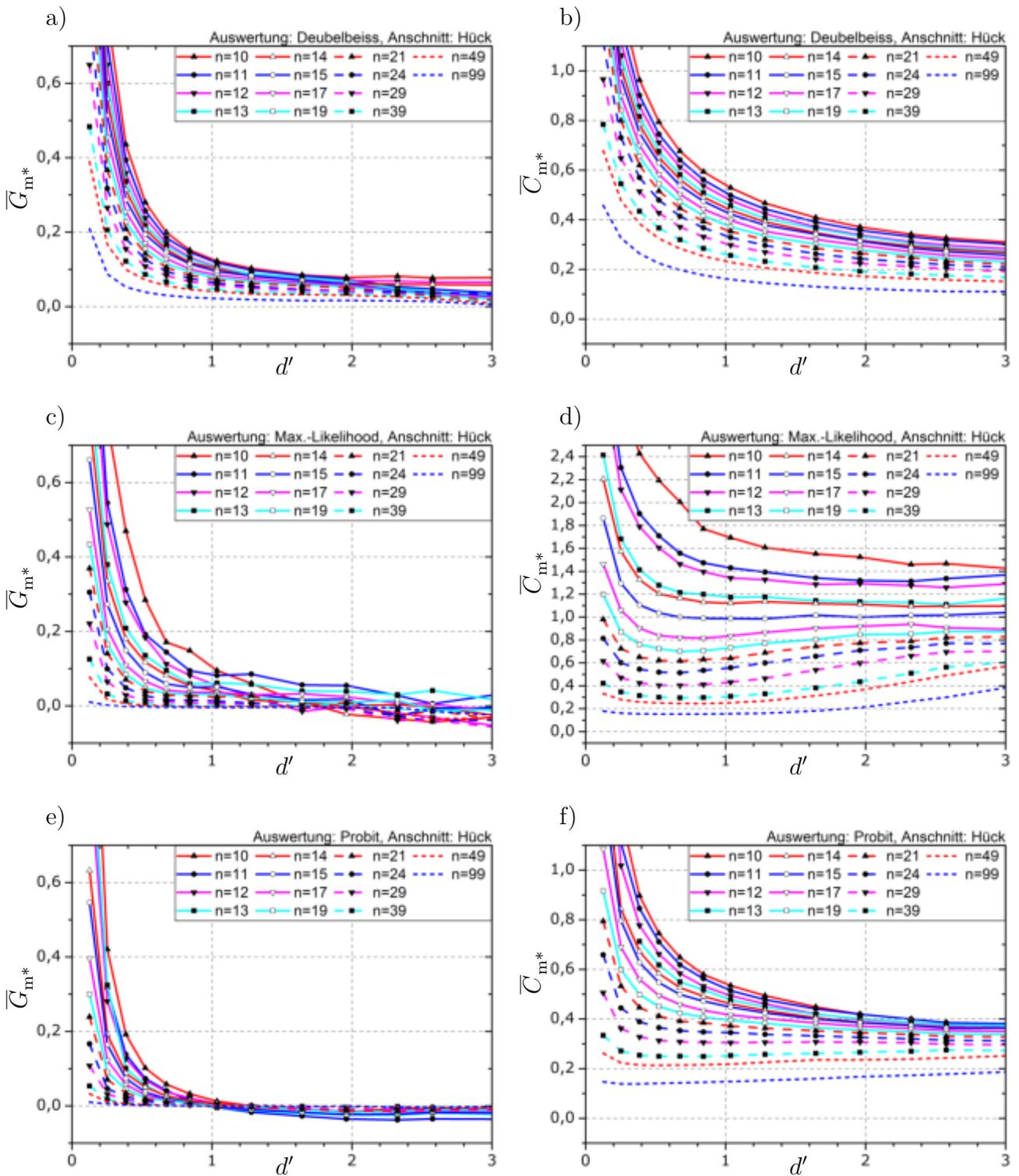


Abb. A.47: Mittlere Verzerrungen und Standardfehler des Schätzers m^* für Treppenstufenversuche mit Ansnchnittsdefinition nach HÜCK, für Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

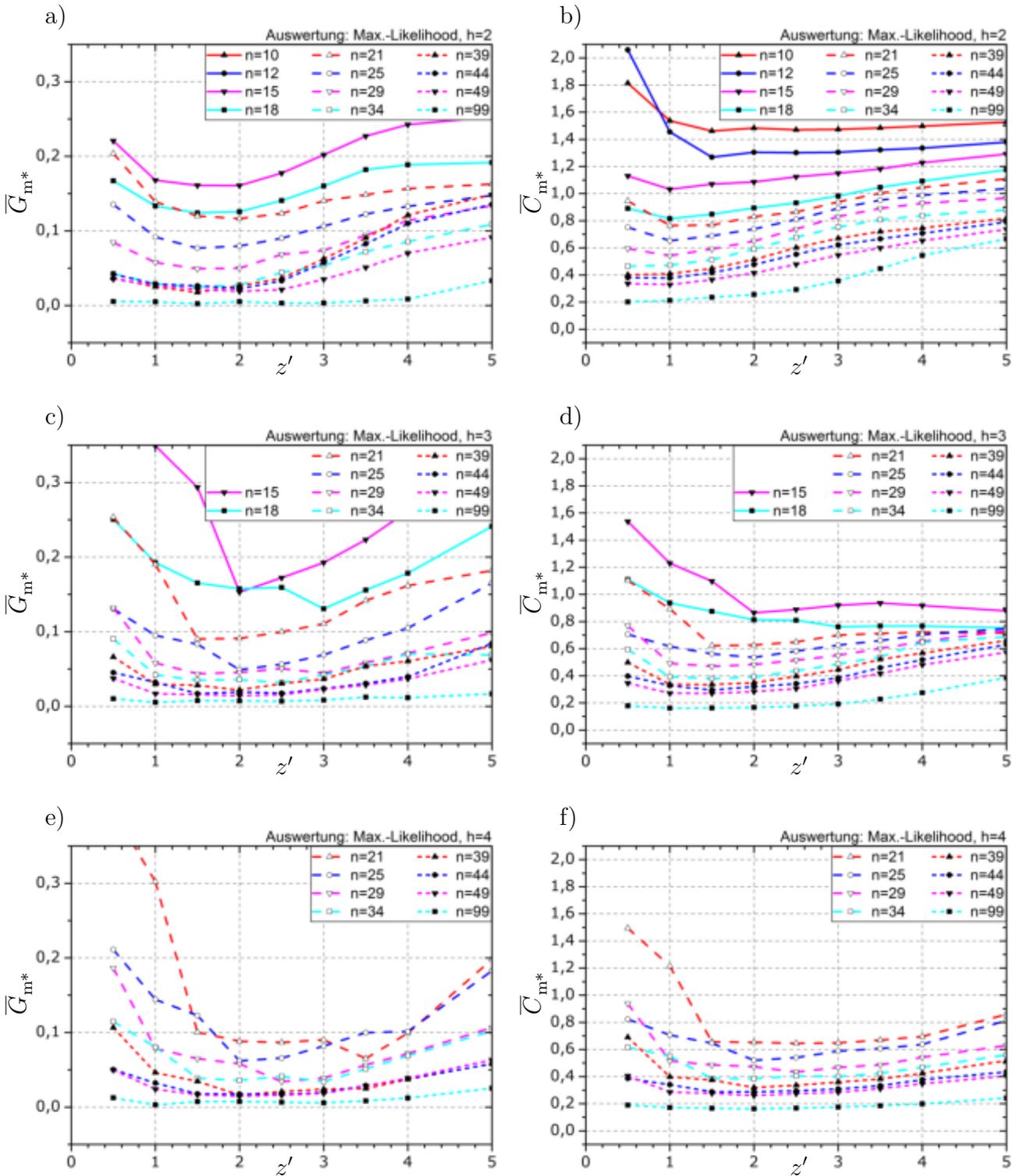


Abb. A.48: Mittlere Verzerrungen und Standardfehler des Schätzers m^* für Horizontversuche mit Auswertung nach MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

A Grafische Darstellung der Simulationsergebnisse

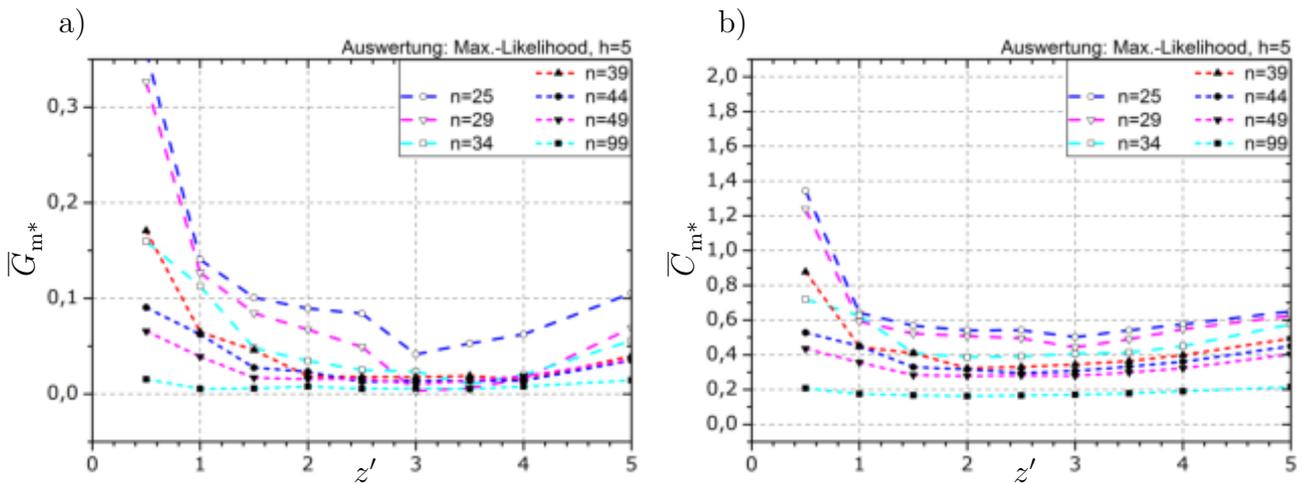


Abb. A.49: Mittlere Verzerrungen und Standardfehler des Schätzers m^* für Horizontversuche mit Auswertung nach MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

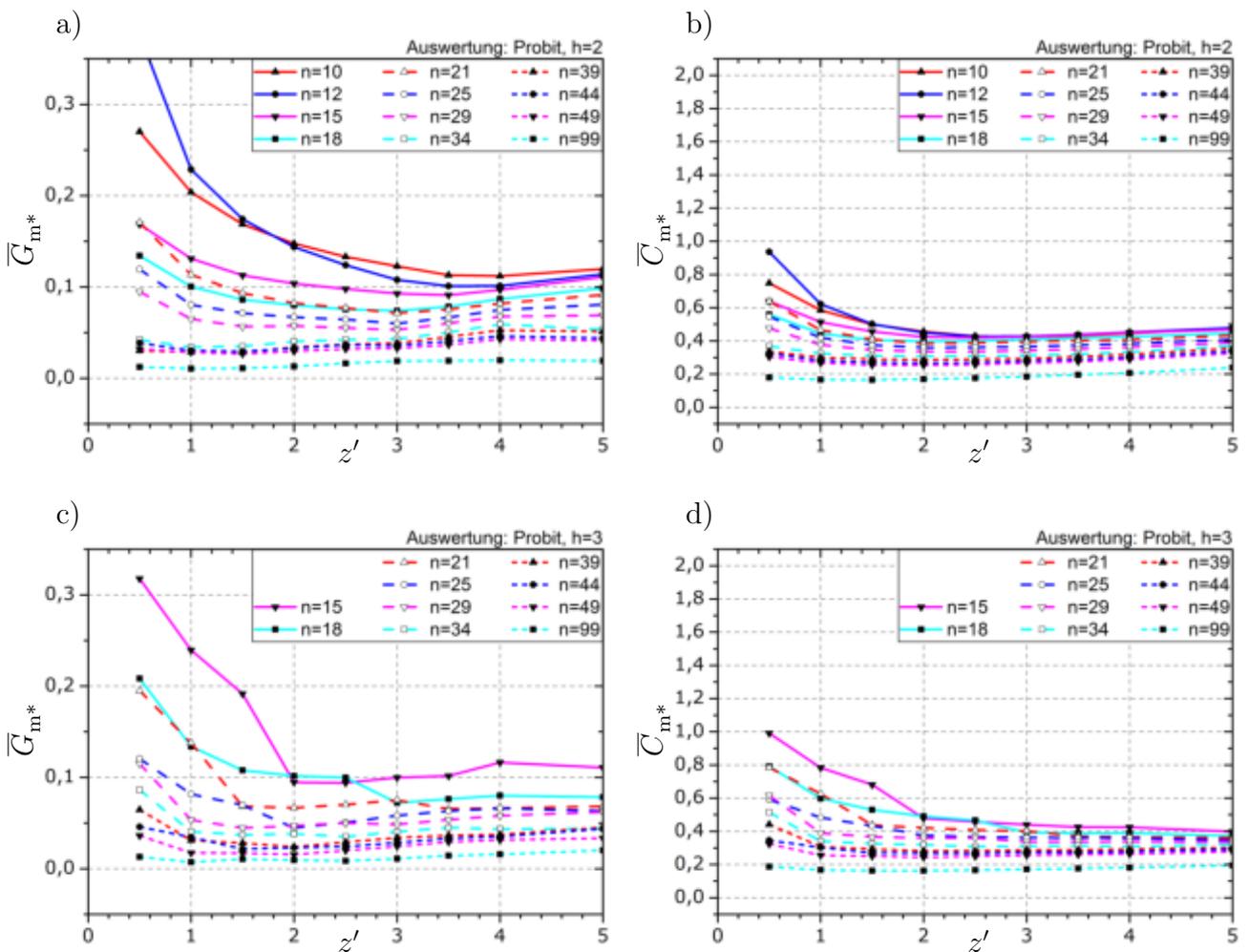


Abb. A.50: Mittlere Verzerrungen und Standardfehler des Schätzers m^* für Horizontversuche mit Auswertung nach PROBIT-METHODE, für Versuche auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

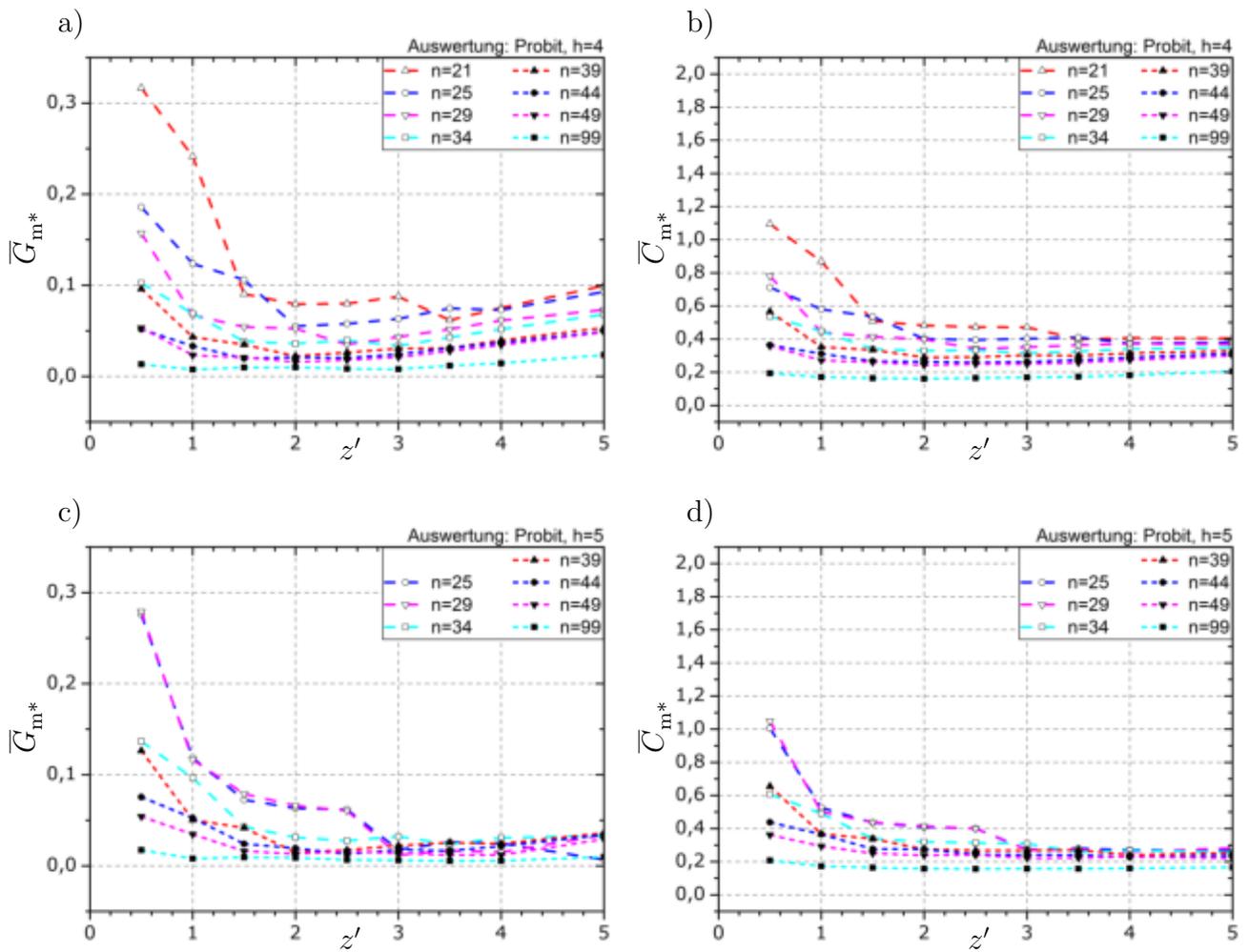


Abb. A.51: Mittlere Verzerrungen und Standardfehler des Schätzers m^* für Horizontversuche mit Auswertung nach PROBIT-METHODE, für Versuche auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.4.2 Verallgemeinerte Standardfehler des Schätzers m^*

Simulationsergebnisse der Treppenstufenversuche

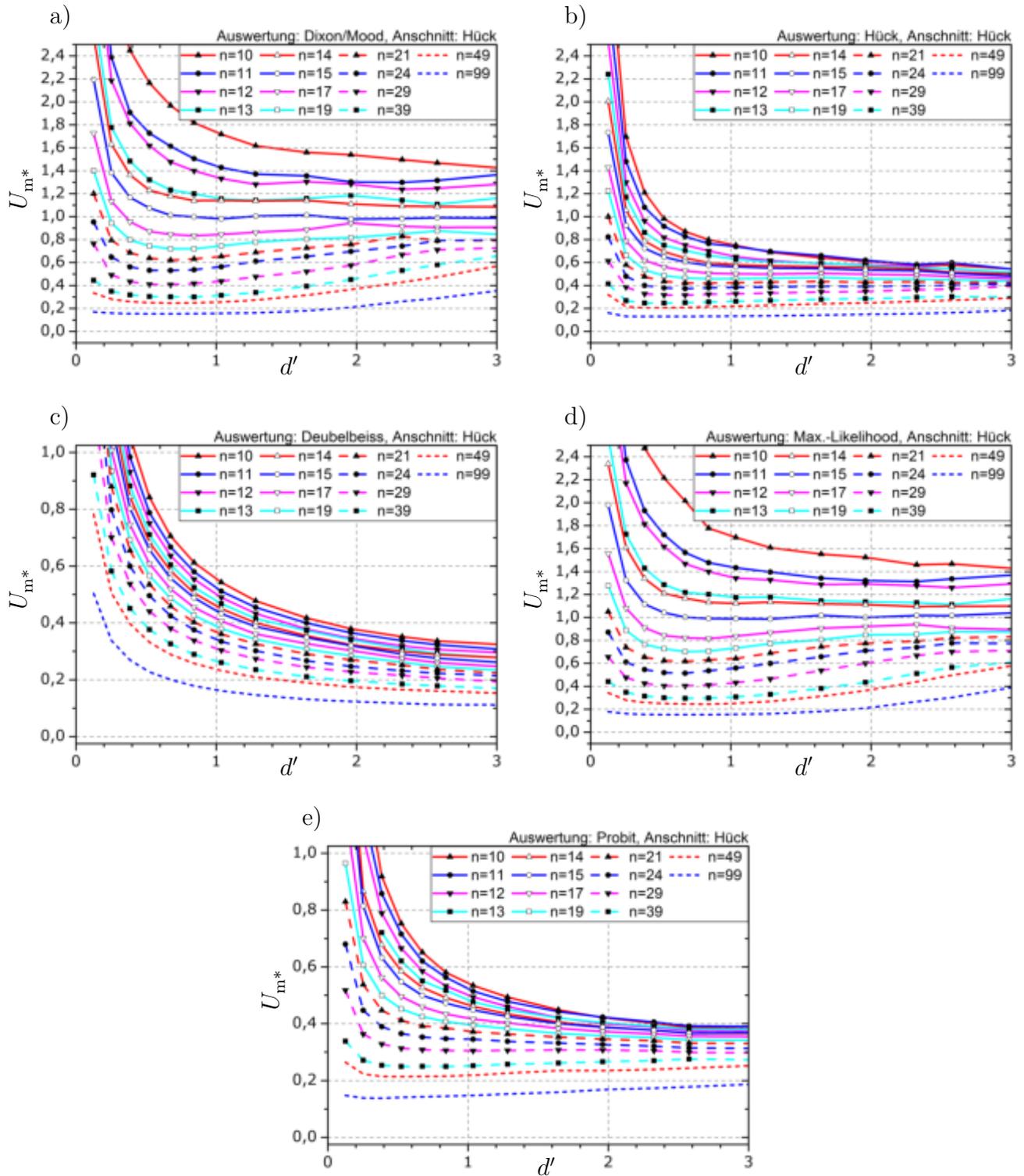


Abb. A.52: Verallgemeinerte Standardfehler U_{m^*} des Schätzers m^* für Treppenstufenversuche mit der Anschnittsdefinition nach HÜCK, für die Auswertemethode nach DIXON/MOOD in a), nach HÜCK in b), nach DEUBELBEISS in c), MAX.-LIKELIHOOD-METHODE in d) sowie PROBIT-METHODE in e).

Simulationsergebnisse der Horizontversuche

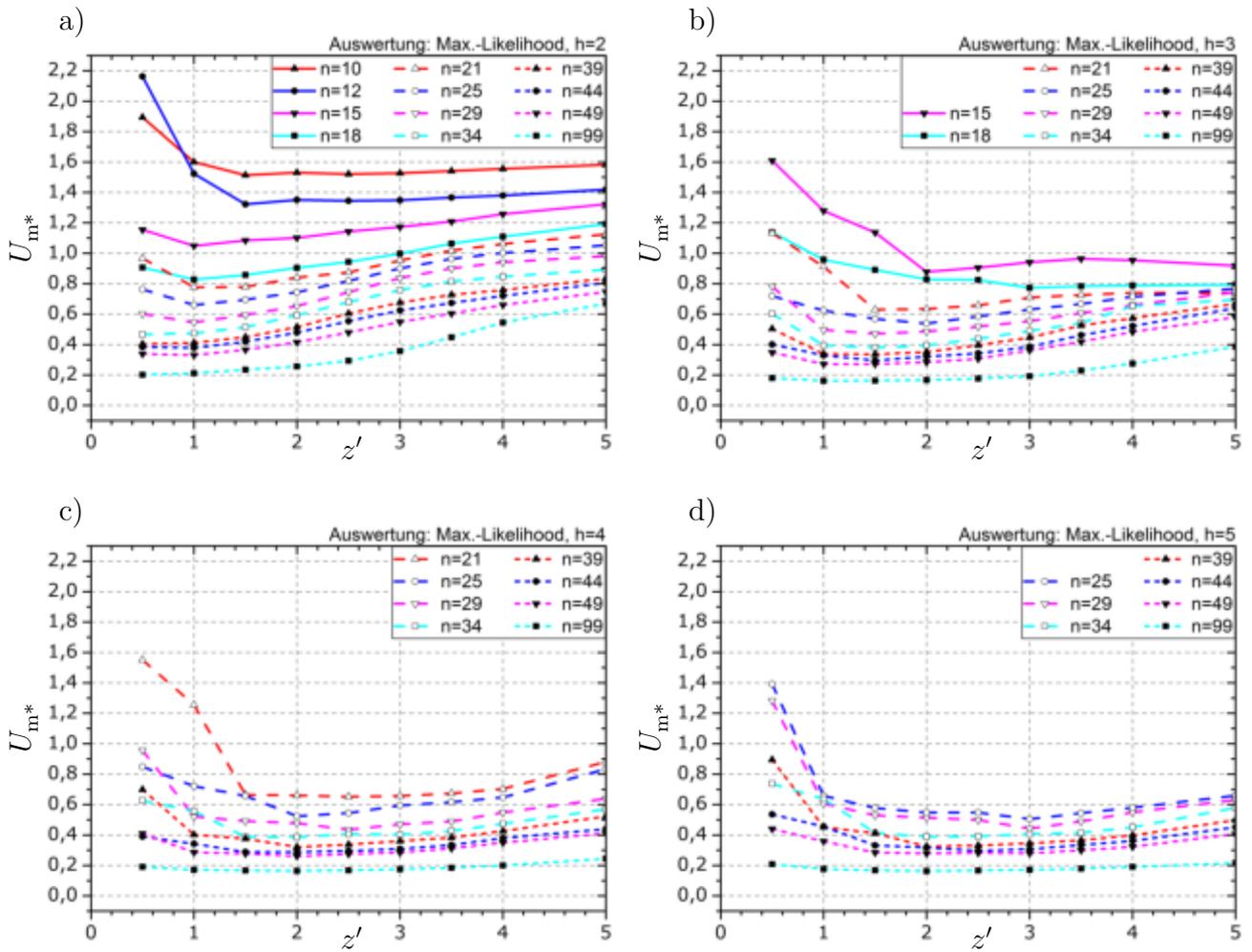


Abb. A.53: Verallgemeinerte Standardfehler U_{m^*} des Schätzers m^* aus Horizontversuchen. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

A Grafische Darstellung der Simulationsergebnisse

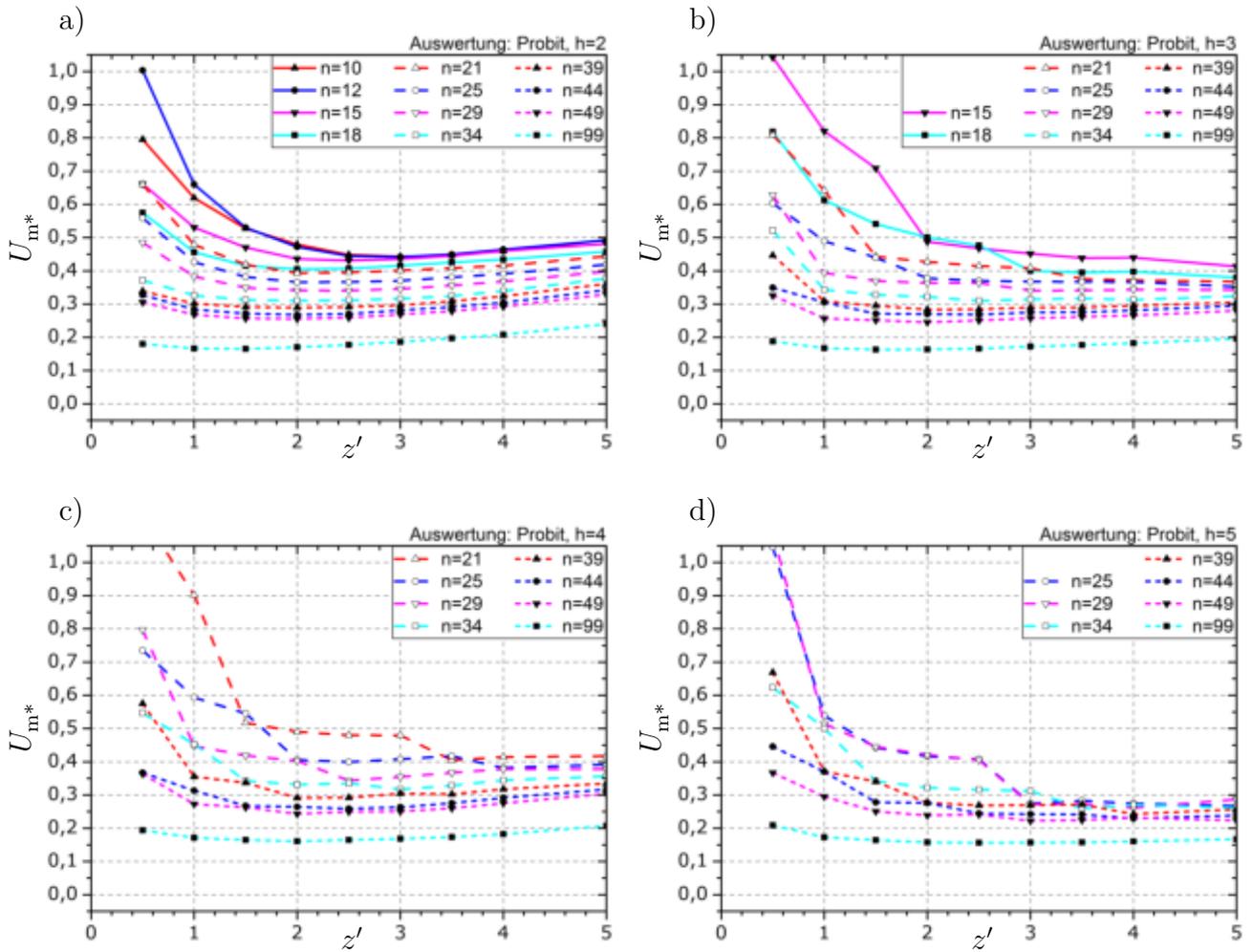


Abb. A.54: Verallgemeinerte Standardfehler U_{m^*} des Schätzers m^* aus Horizontversuchen. Auswertung mittels PROBIT-METHODE, für Versuche auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

A.4.3 RMS-Schätzfehler des Schätzers m^*

Simulationsergebnisse der Treppenstufenversuche

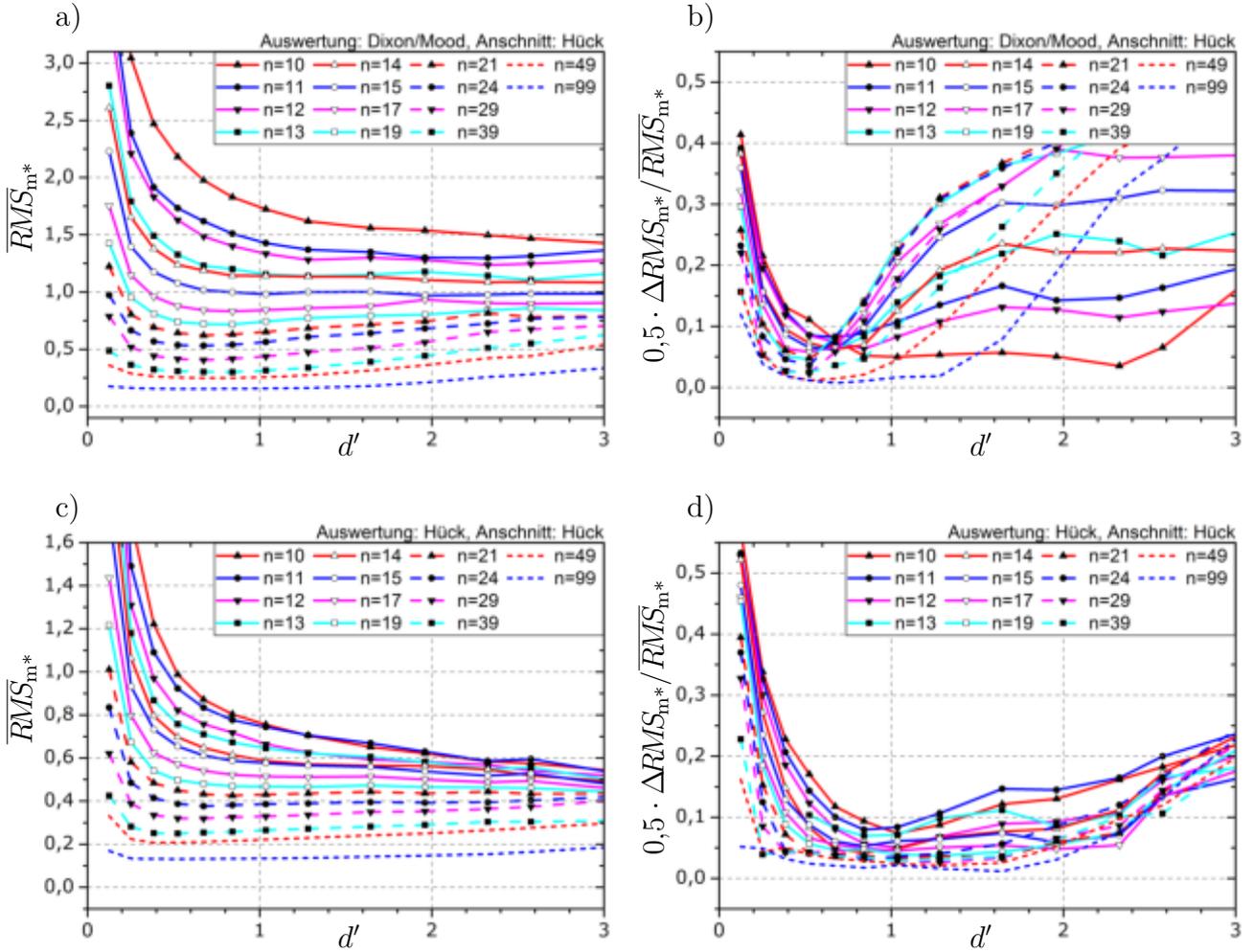


Abb. A.55: Mittlerer RMS-Fehler \overline{RMS}_{m^*} und relative Abweichung $0,5 \cdot \Delta RMS_{m^*} / \overline{RMS}_{m^*}$ des Schätzers m^* für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für die Auswertemethode nach DIXON/MOOD in a) und b) sowie HÜCK in c) und d).

A Grafische Darstellung der Simulationsergebnisse

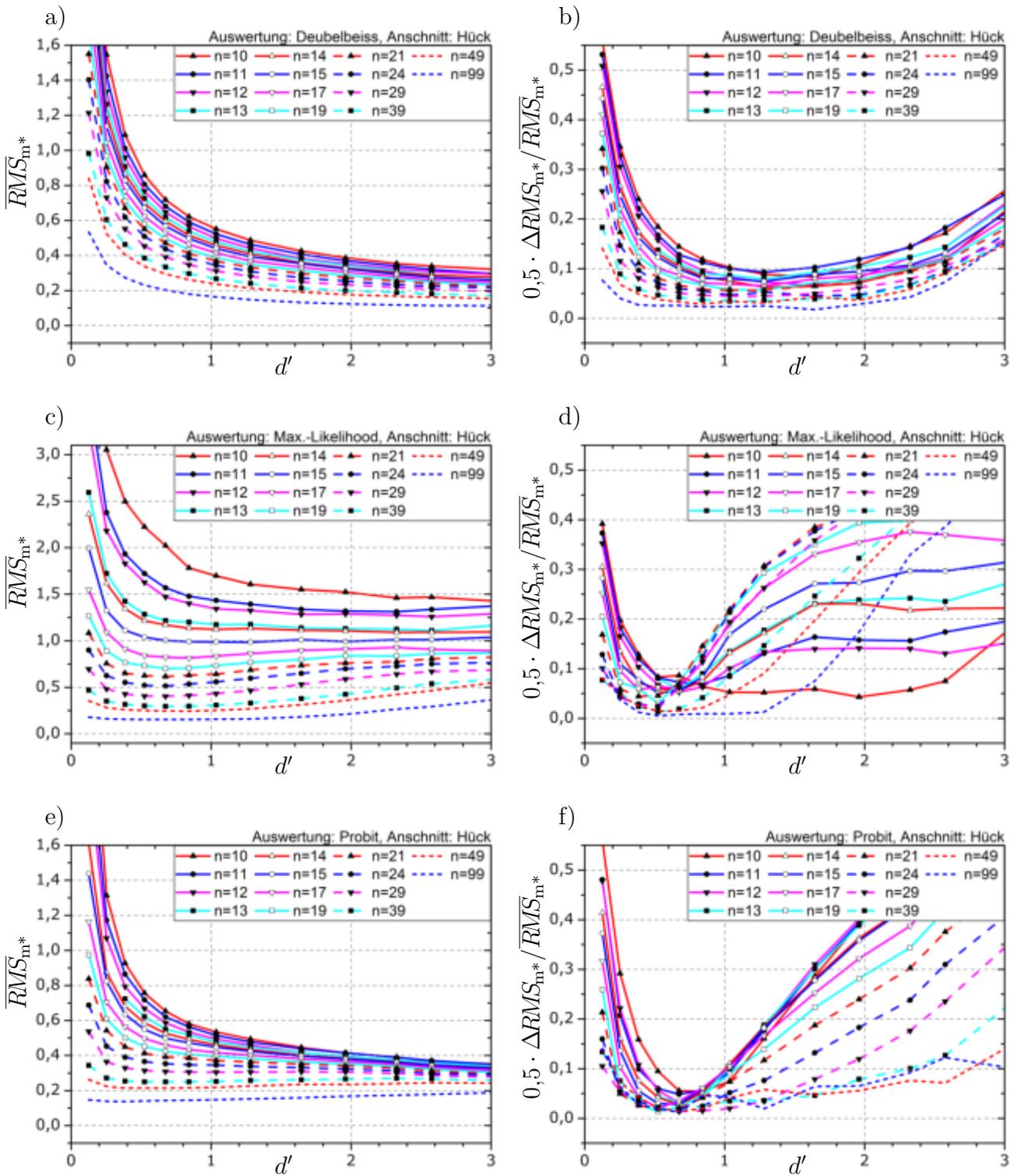


Abb. A.56: Mittlerer RMS-Fehler \overline{RMS}_{m^*} und relative Abweichung $0,5 \cdot \Delta RMS_{m^*} / \overline{RMS}_{m^*}$ des Schätzers m^* für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für die Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

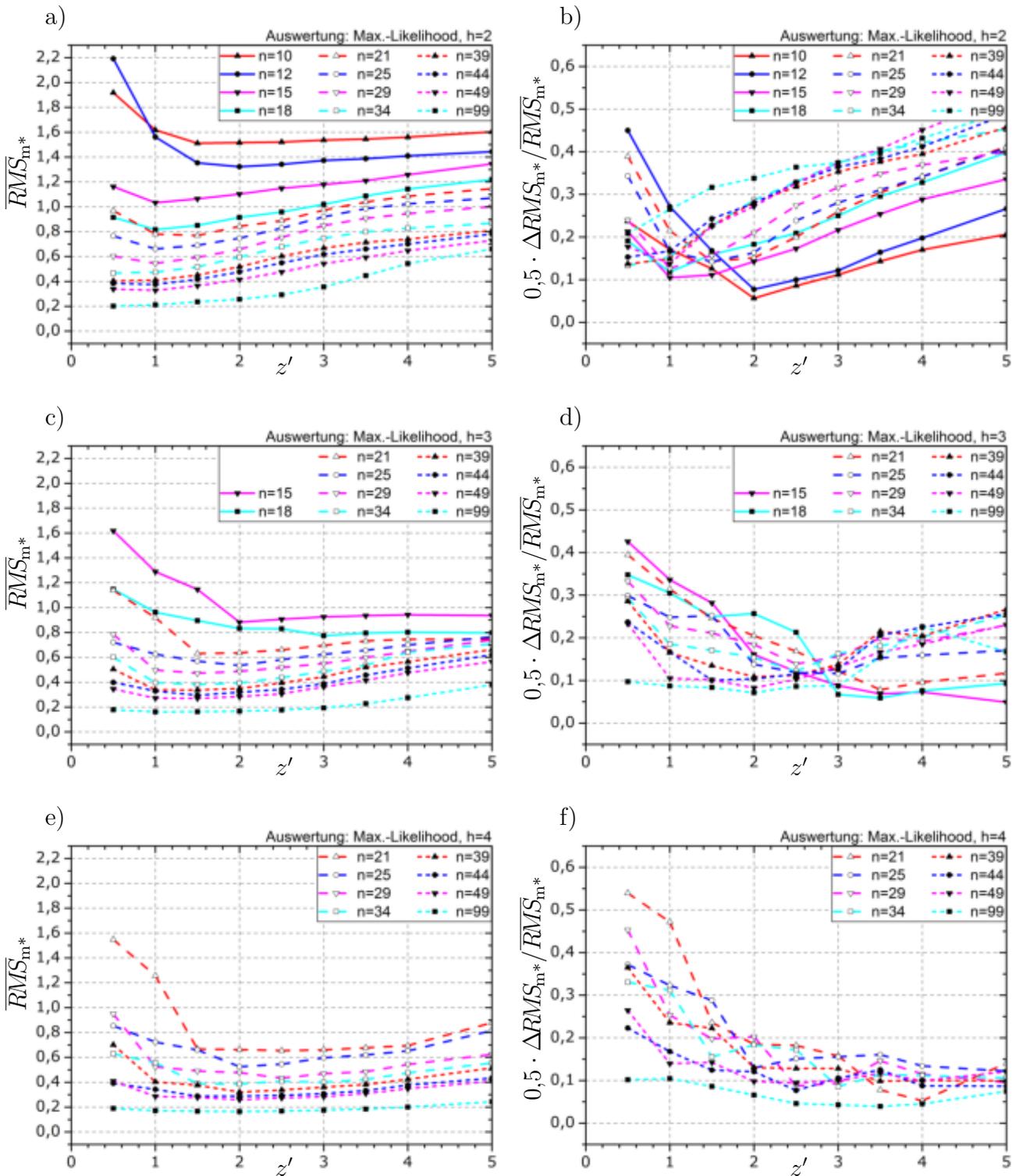


Abb. A.57: Mittlerer RMS-Fehler \overline{RMS}_{m^*} und relative Abweichung $0,5 \cdot \Delta RMS_{m^*} / \overline{RMS}_{m^*}$ des Schätzers m^* für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

A Grafische Darstellung der Simulationsergebnisse

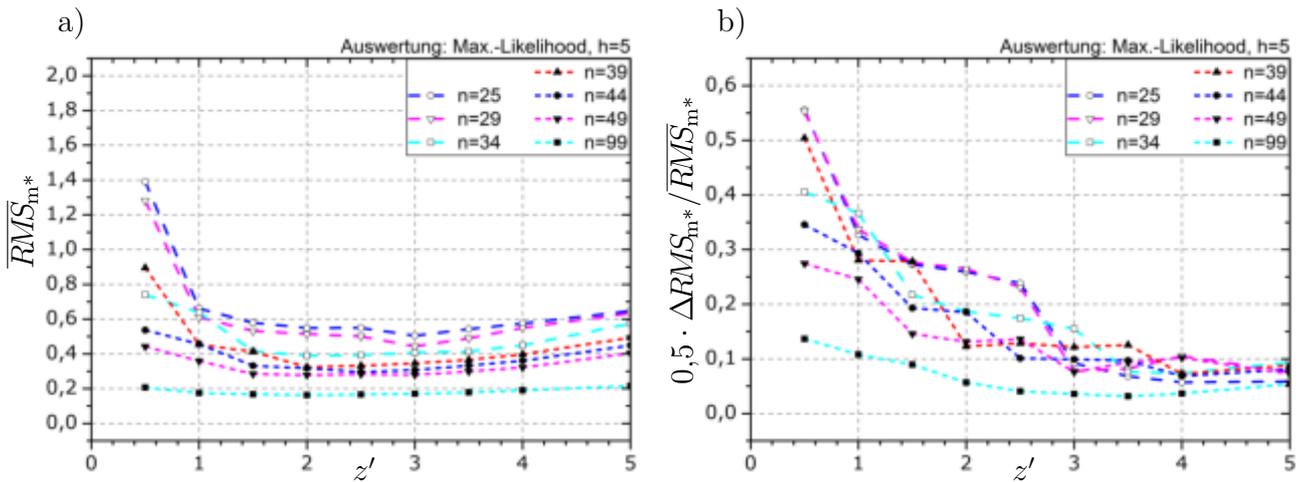


Abb. A.58: Mittlerer RMS-Fehler \overline{RMS}_{m^*} und relative Abweichung $0,5 \cdot \Delta RMS_{m^*} / \overline{RMS}_{m^*}$ des Schätzers m^* für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

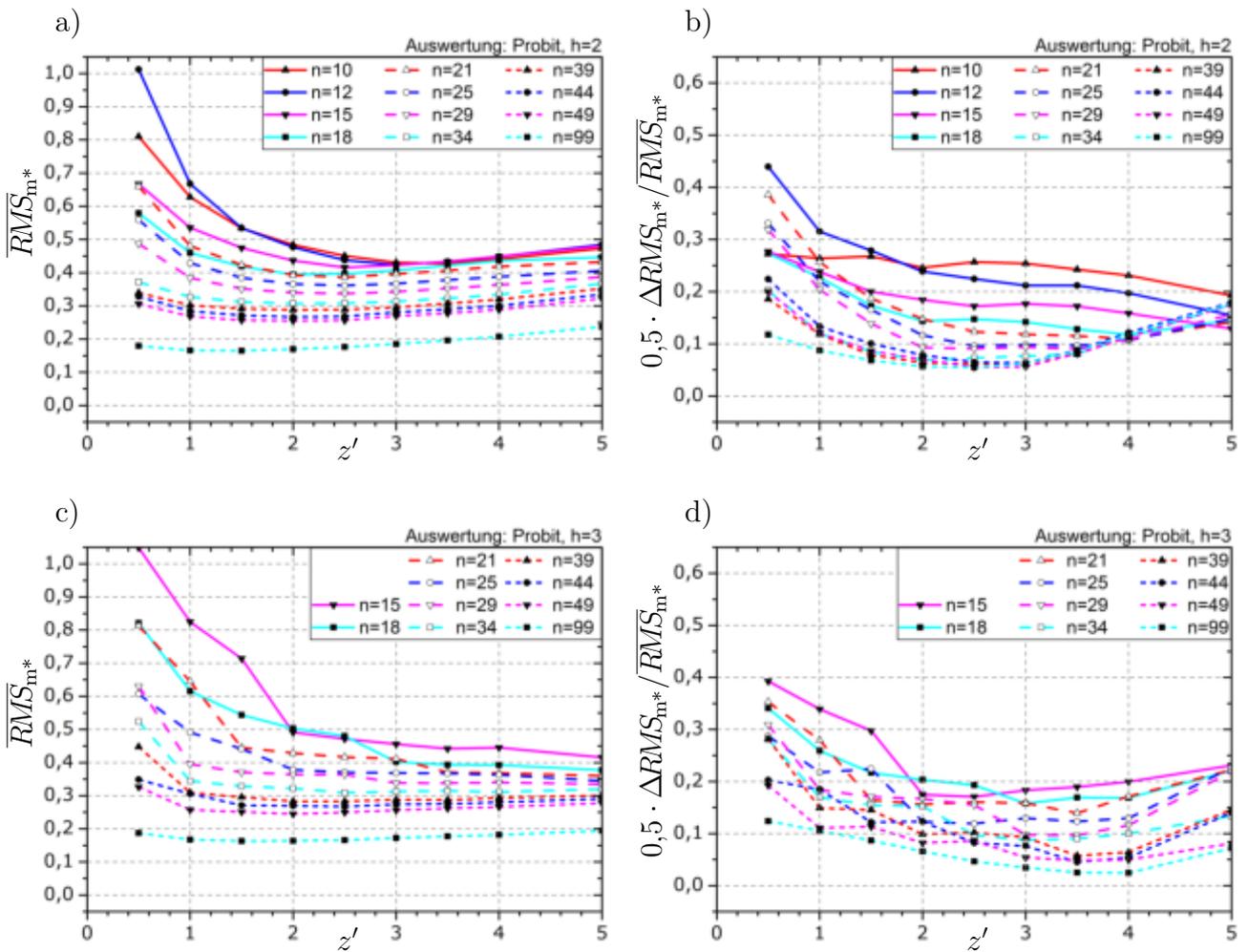


Abb. A.59: Mittlerer RMS-Fehler \overline{RMS}_{m^*} und relative Abweichung $0,5 \cdot \Delta RMS_{m^*} / \overline{RMS}_{m^*}$ des Schätzers m^* für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

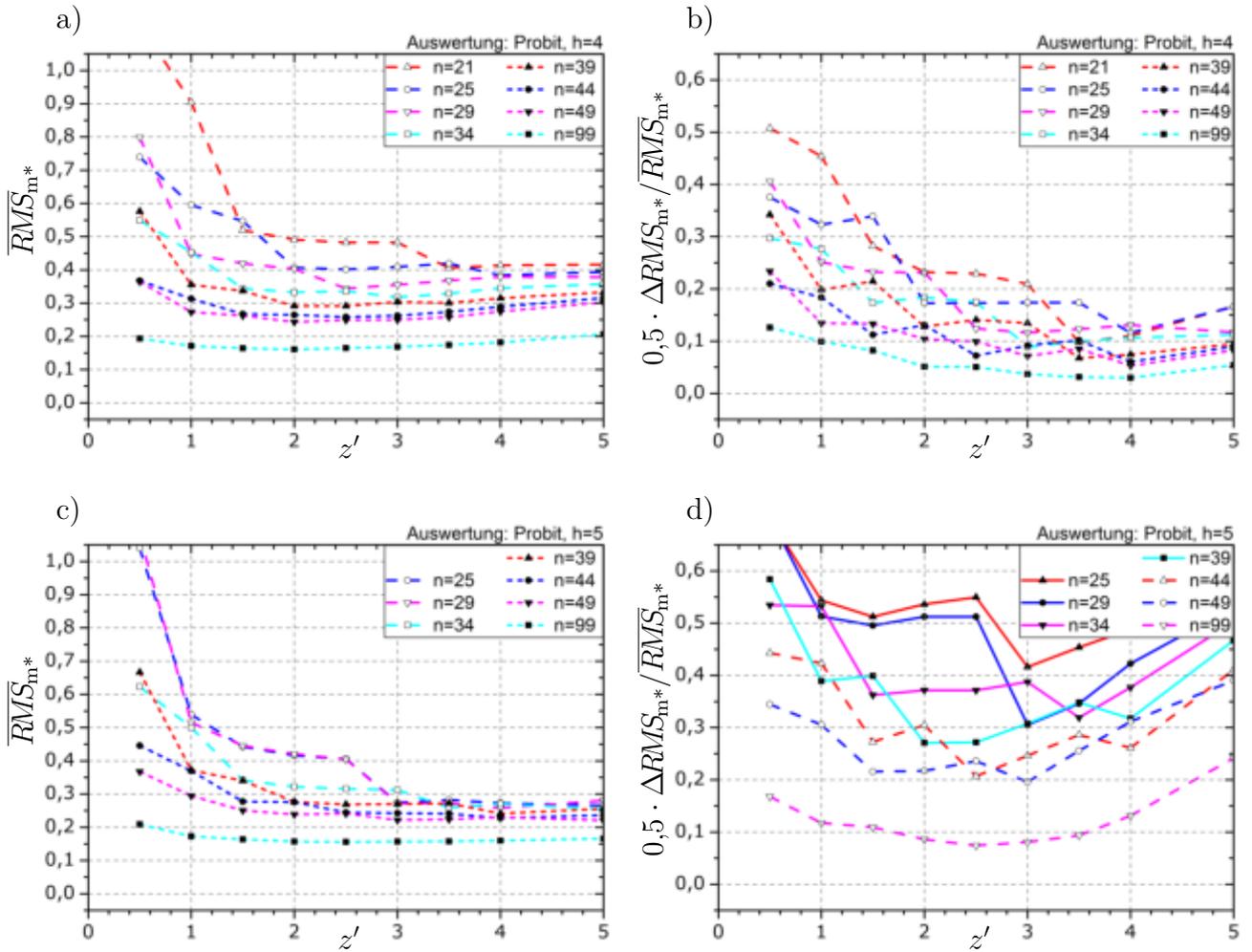


Abb. A.60: Mittlerer RMS-Fehler \overline{RMS}_{m^*} und relative Abweichung $0,5 \cdot \Delta RMS_{m^*} / \overline{RMS}_{m^*}$ des Schätzers m^* für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.4.4 Mittlere verallgemeinerte 5%-Quantile des Schätzers m^*

Simulationsergebnisse der Treppenstufenversuche

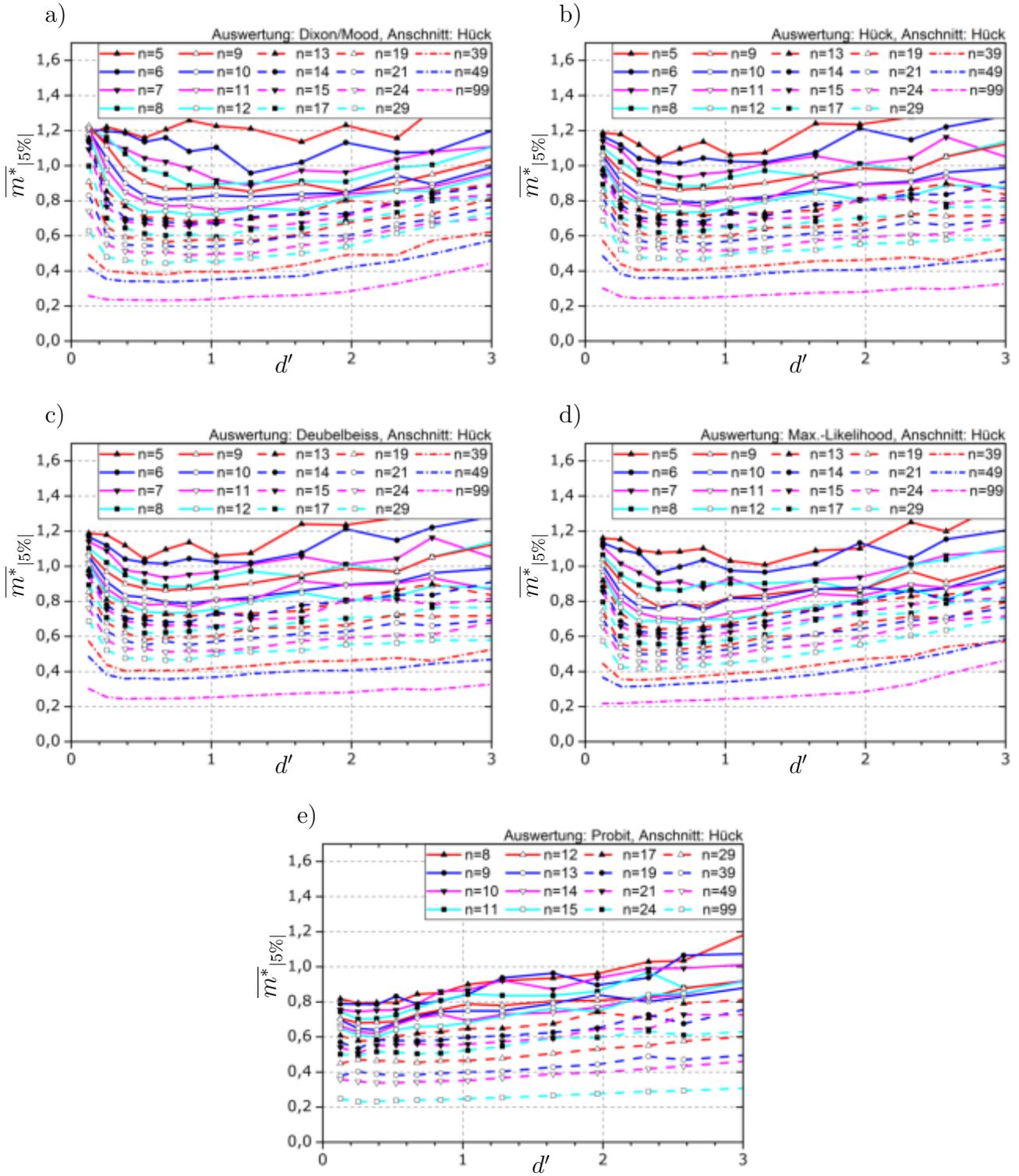


Abb. A.61: Mittlere verallgemeinerte 5%-Quantile $\overline{m^*}_{|5\%|}$ des Schätzers m^* für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK. Für die Auswertemethode nach DIXON/MOOD in a), nach HÜCK in b), nach DEUBELBEISS in c), MAX.-LIKELIHOOD-METHODE in d) sowie PROBIT-METHODE in e).

Simulationsergebnisse der Horizontversuche

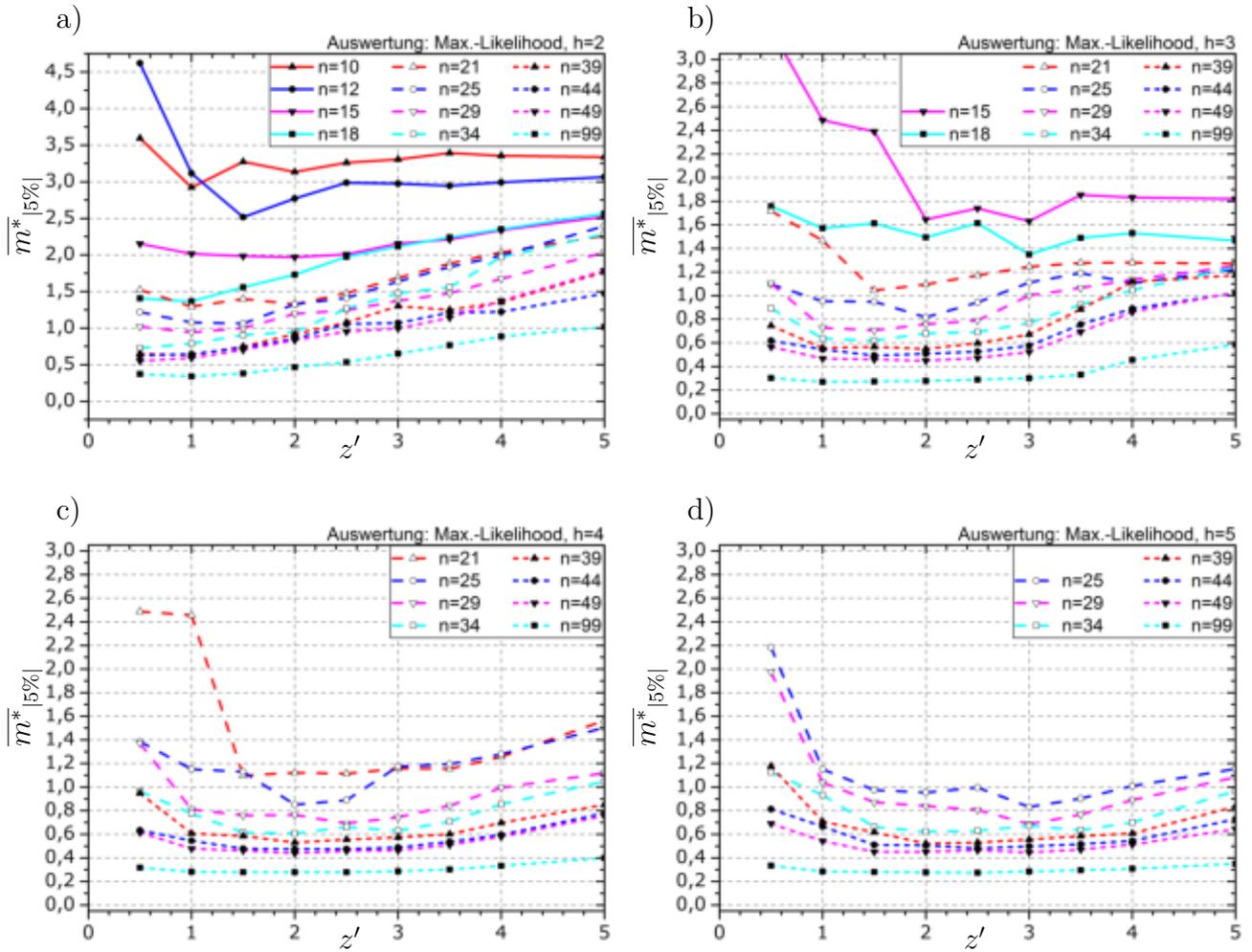


Abb. A.62: Mittlere verallgemeinerte 5%-Quantile $\overline{m^*}_{|5\%|}$ des Schätzers m^* aus Horizontversuchen für die Auswertung mittels MAX.-LIKELIHOOD-METHODE und Versuchen auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

A Grafische Darstellung der Simulationsergebnisse

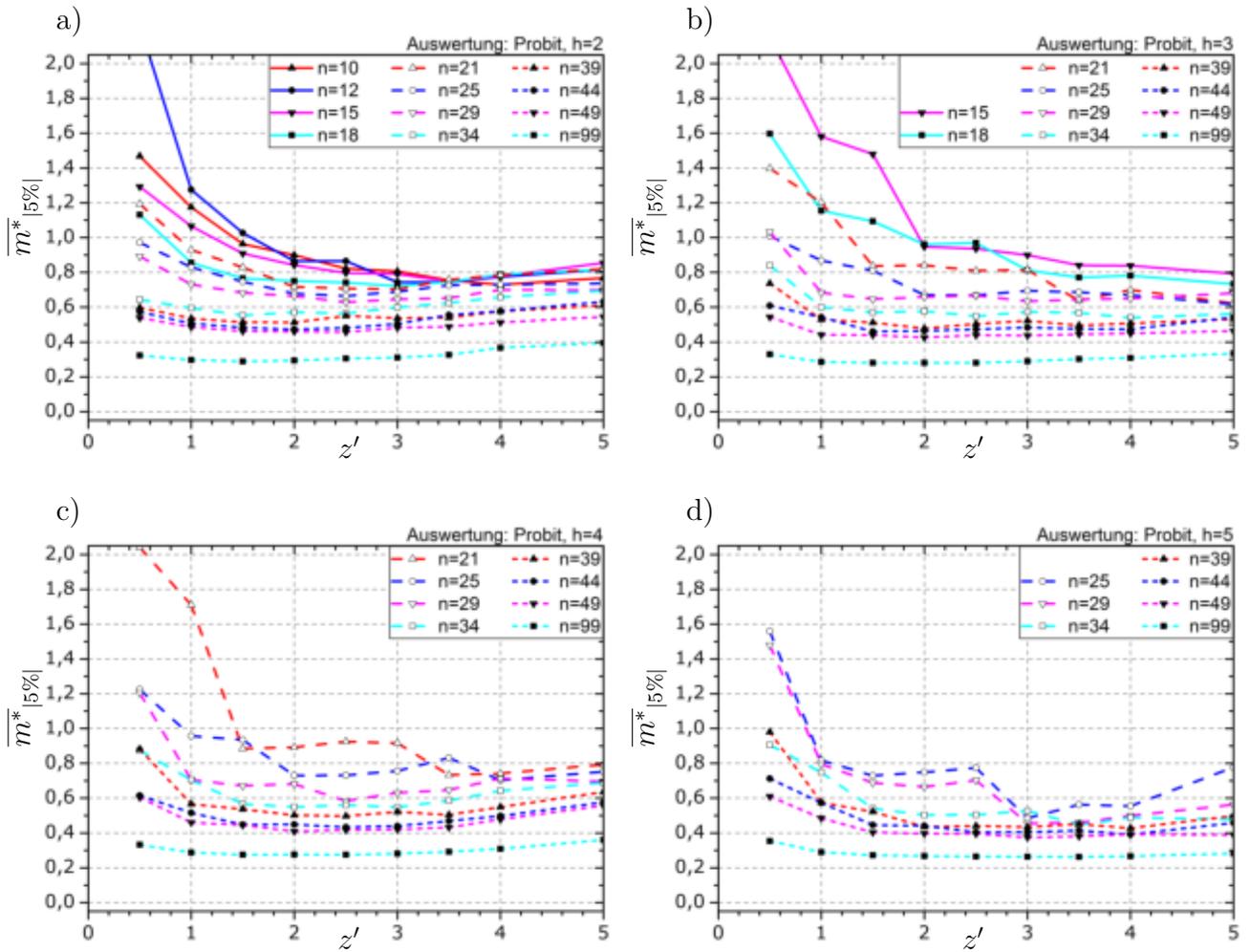


Abb. A.63: Mittlere verallgemeinerte 5%-Quantile $\overline{m^*}_{|5\%|}$ des Schätzers m^* aus Horizontversuchen für die Auswertung mittels PROBIT-METHODE und Versuchen auf 2 Horzonten in a), 3 Horzonten in b), 4 Horzonten in c) sowie 5 Horzonten in d).

A.5 Schätzung des Quantils 2,5%-Quantils bei bekannter Standardabweichung

A.5.1 Mittlere Verzerrungen und Standardfehler des Schätzers x'

Simulationsergebnisse der Treppenstufenversuche

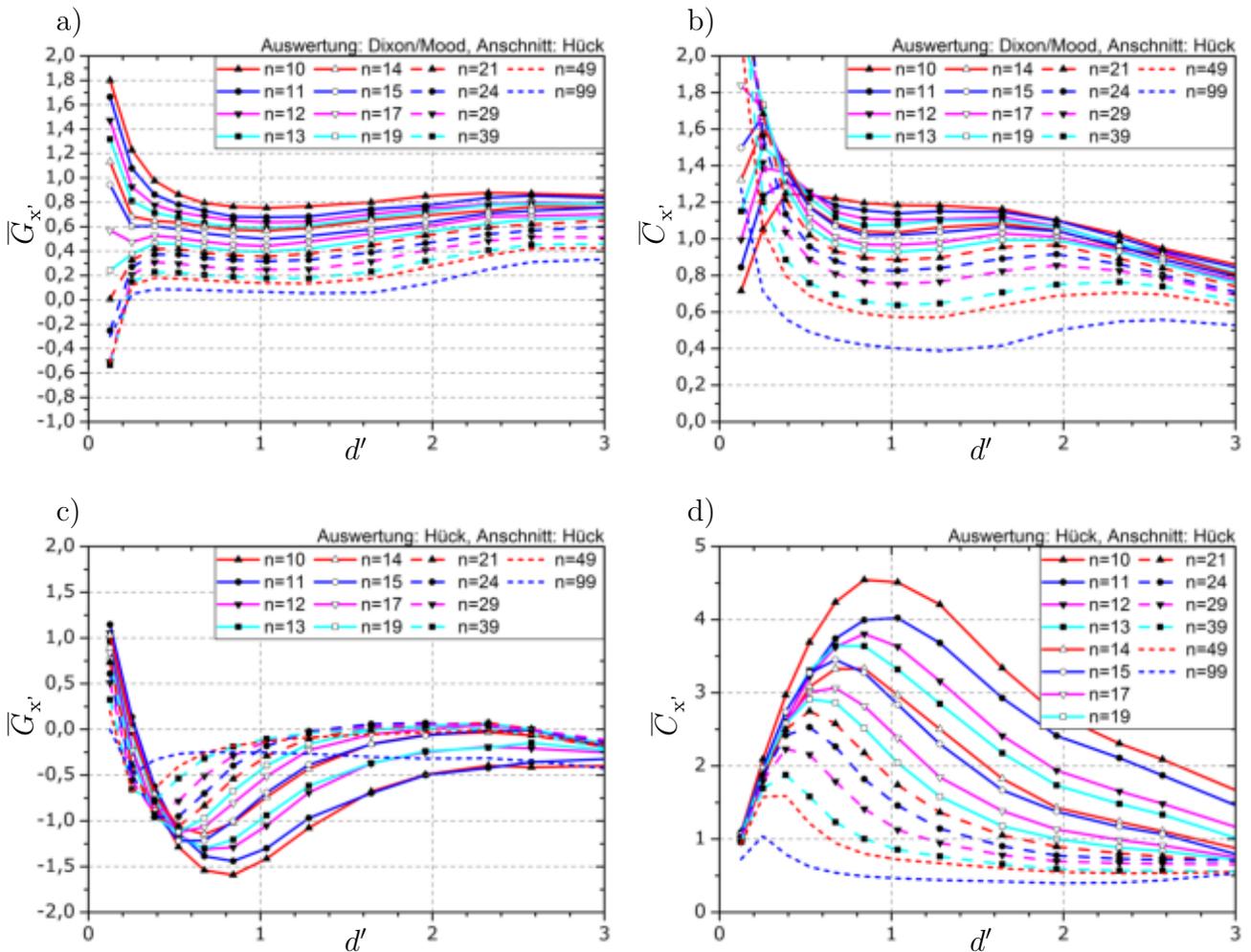


Abb. A.64: Mittlere Verzerrungen und Standardfehler des Schätzers x' für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethode nach DIXON/MOOD in a) und b) sowie HÜCK in c) und d).

A Grafische Darstellung der Simulationsergebnisse

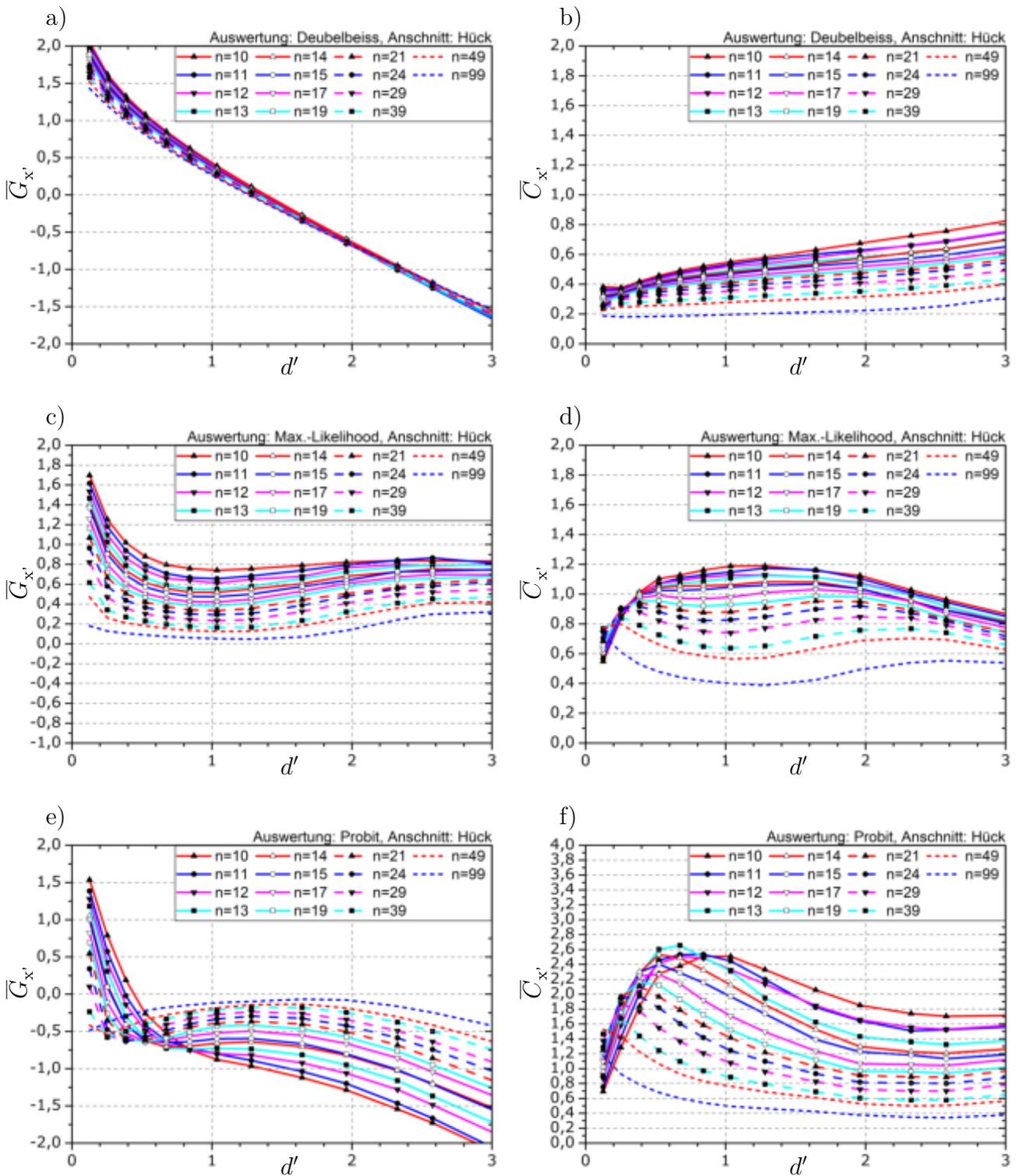


Abb. A.65: Mittlere Verzerrungen und Standardfehler des Schätzers x' für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

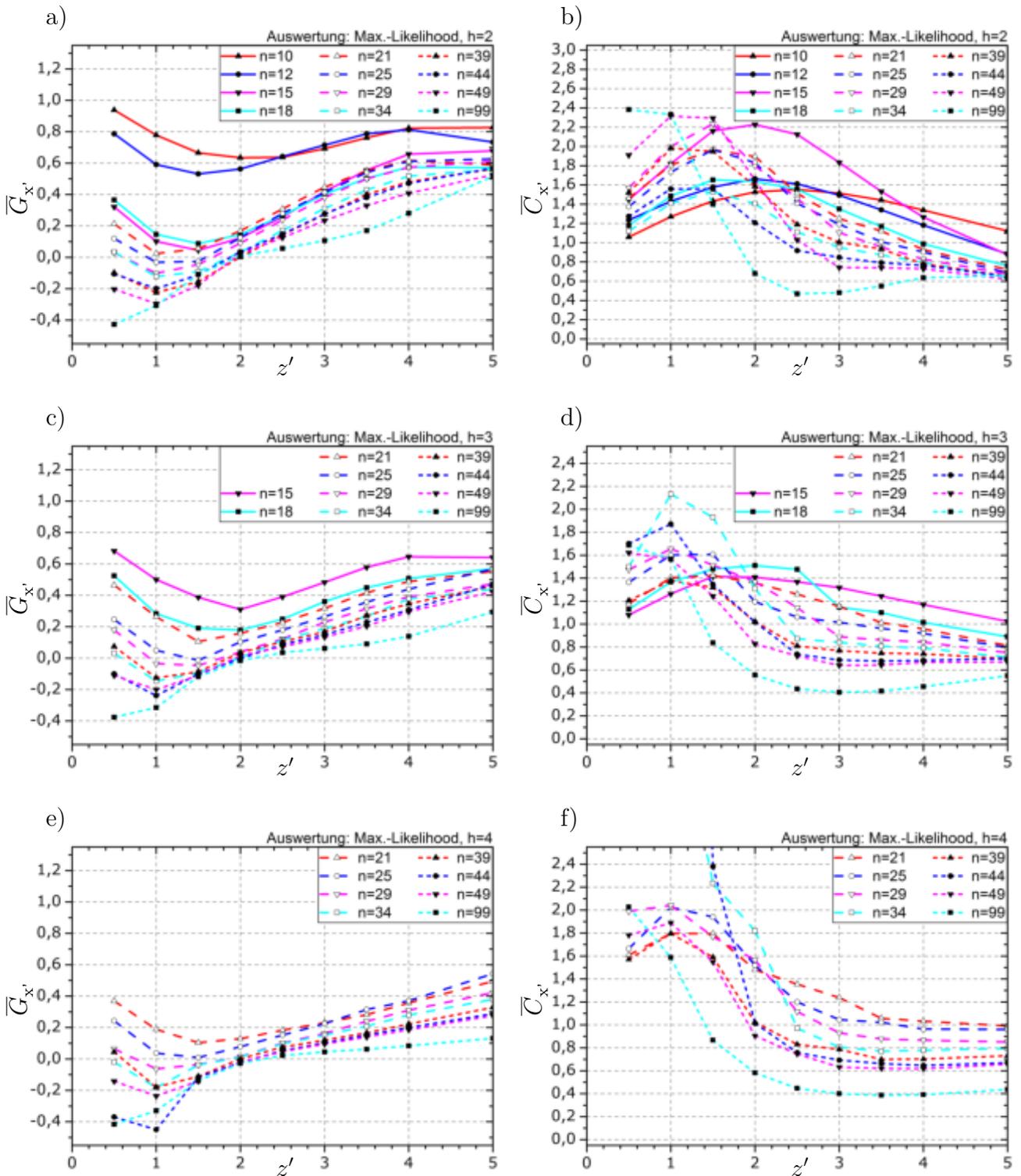


Abb. A.66: Mittlere Verzerrungen und Standardfehler des Schätzers x' für Horizontversuche mit Auswertung nach MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

A Grafische Darstellung der Simulationsergebnisse

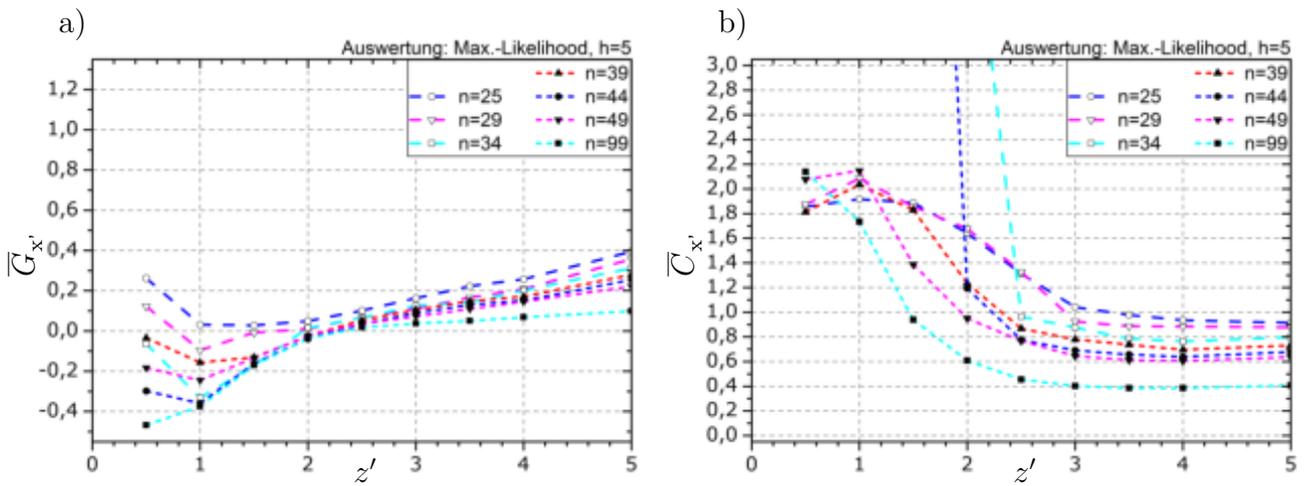


Abb. A.67: Mittlere Verzerrungen und Standardfehler des Schätzers x' für Horizontversuche mit Auswertung nach MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

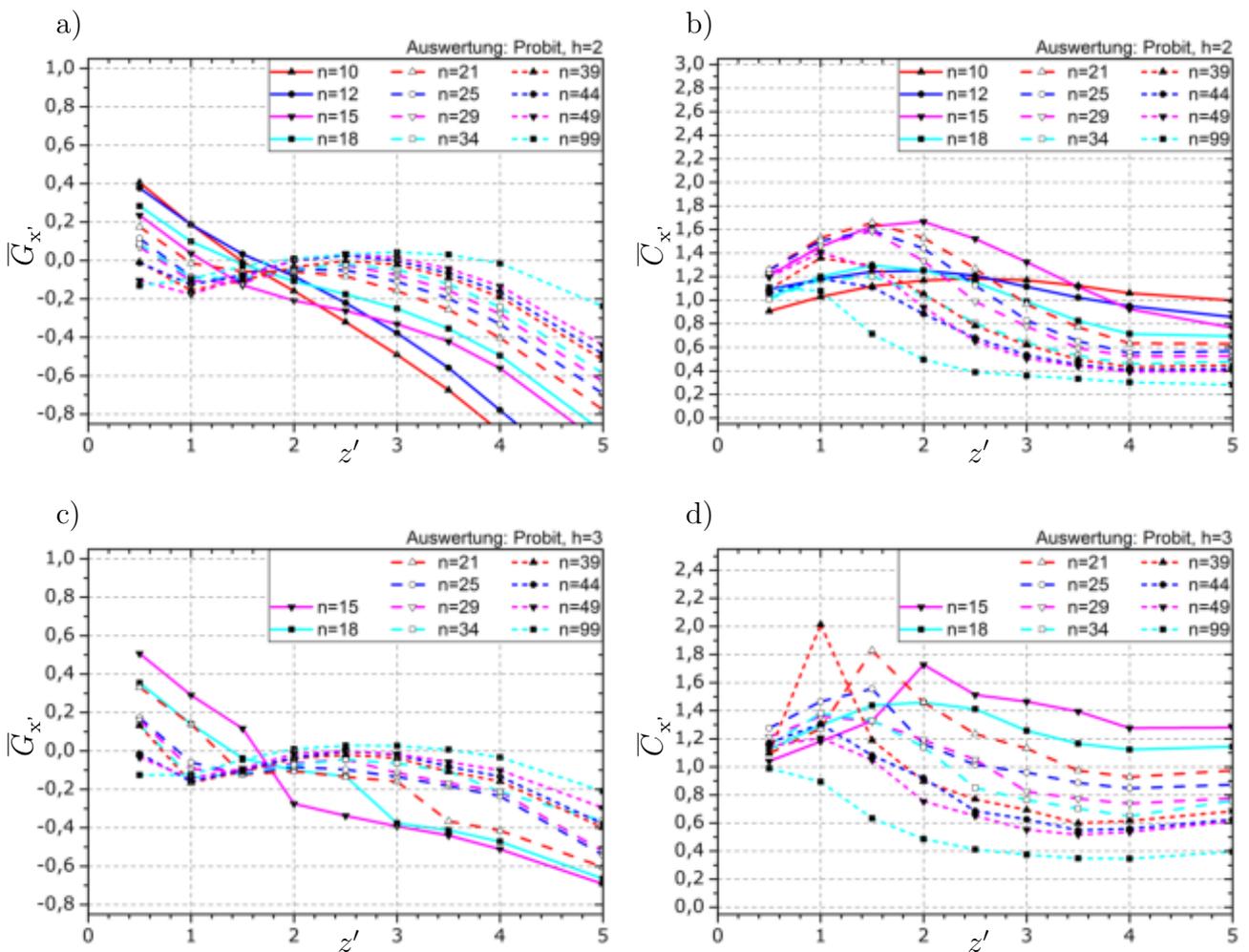


Abb. A.68: Mittlere Verzerrungen und Standardfehler des Schätzers x' für Horizontversuche mit Auswertung nach PROBIT-METHODE, für Versuche auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

A.5 Schätzung des 2,5%-Quantils bei bekannter Standardabweichung

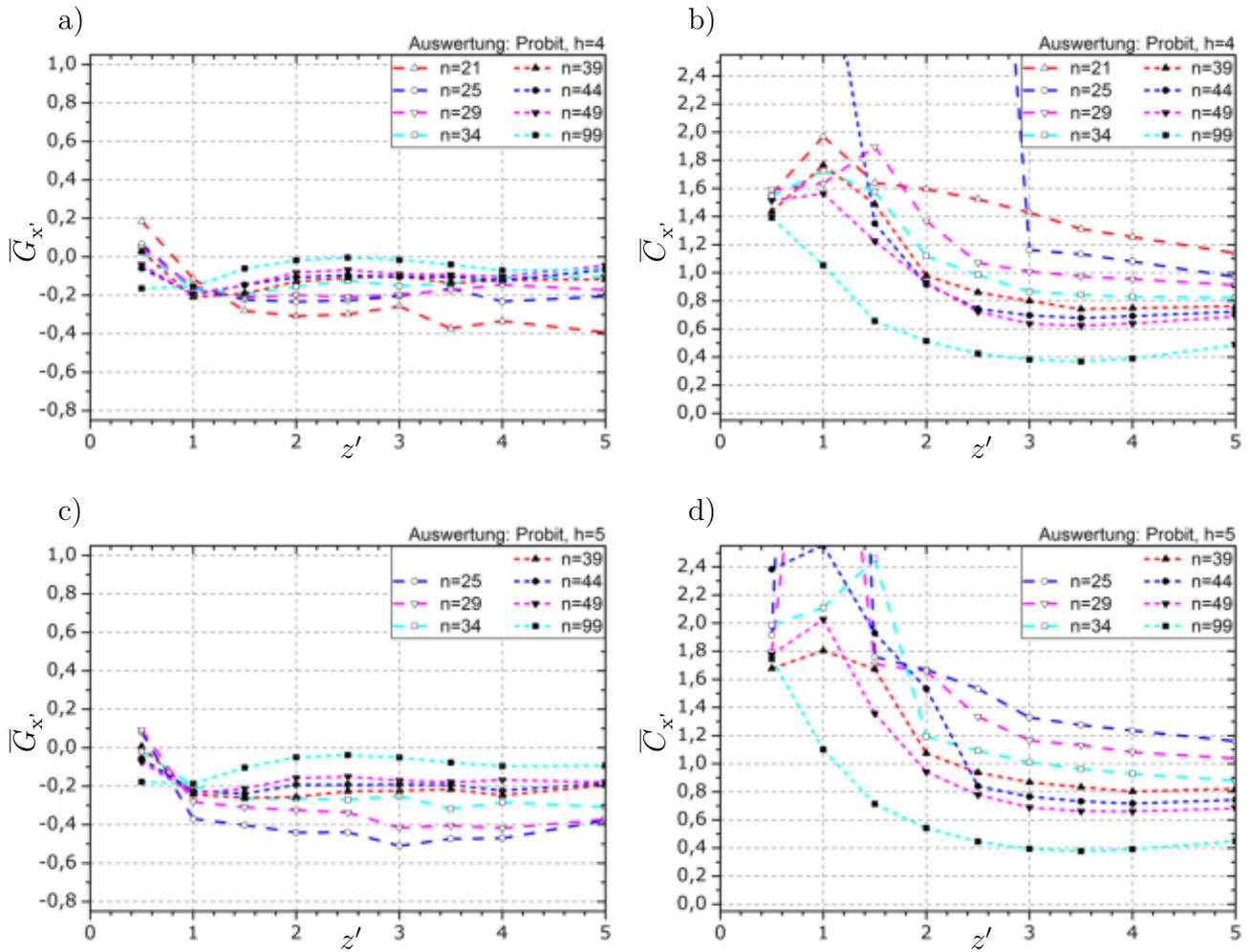


Abb. A.69: Mittlere Verzerrungen und Standardfehler des Schätzers x' für Horizontversuche mit Auswertung nach PROBIT-METHODE, für Versuche auf auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.5.2 RMS-Fehler des Schätzers x'

Simulationsergebnisse der Treppenstufenversuche

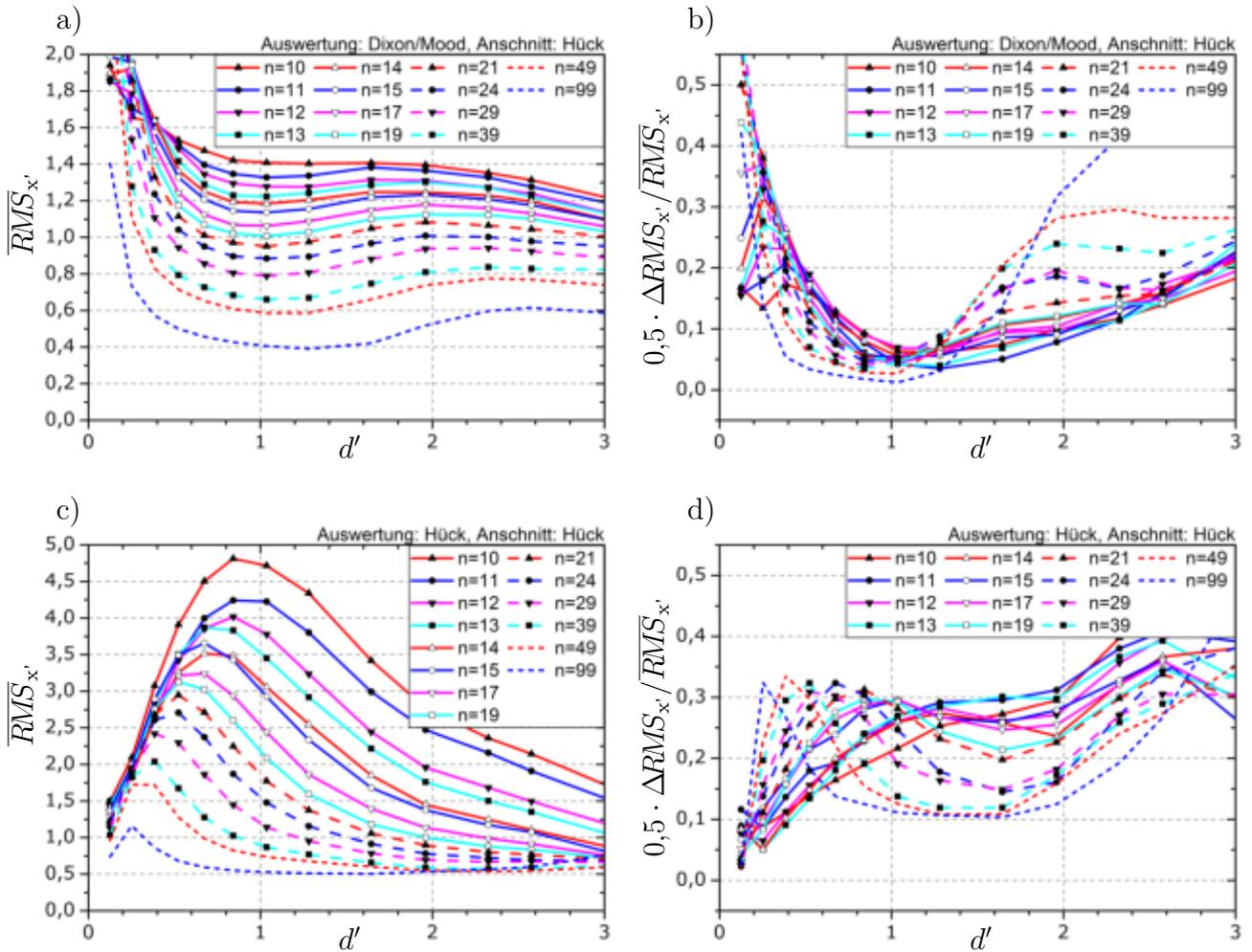


Abb. A.70: Mittlerer RMS-Fehler $\overline{RMS}_{x'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{x'} / \overline{RMS}_{x'}$ des Schätzers x' für Treppenstufenversuche mit Ansnchnittsdefinition nach HÜCK, für die Auswertemethode nach DIXON/MOOD in a) und b) sowie HÜCK in c) und d).

A.5 Schätzung des 2,5%-Quantils bei bekannter Standardabweichung

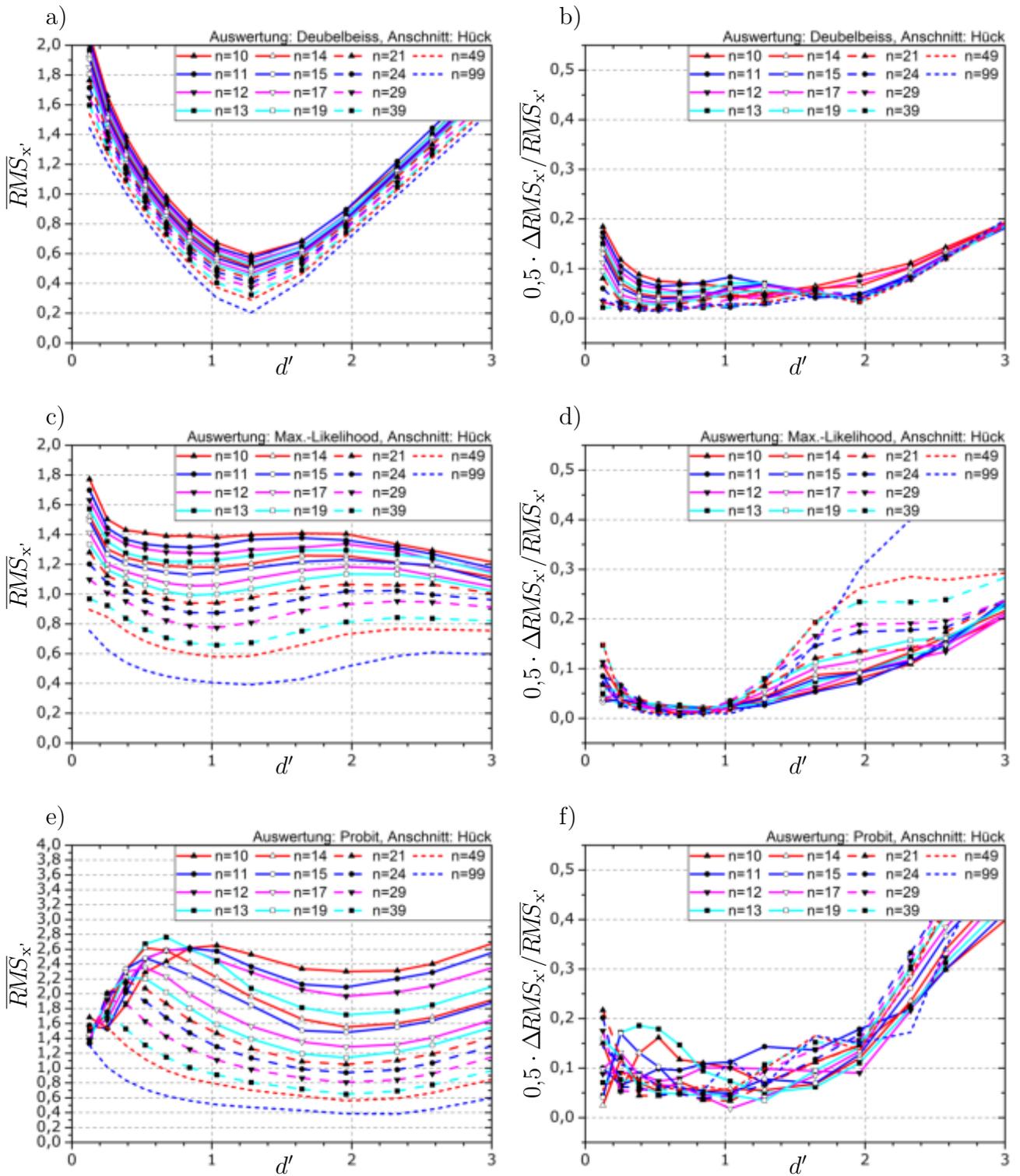


Abb. A.71: Mittlerer RMS-Fehler $\overline{RMS}_{x'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{x'} / \overline{RMS}_{x'}$ des Schätzers x' für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

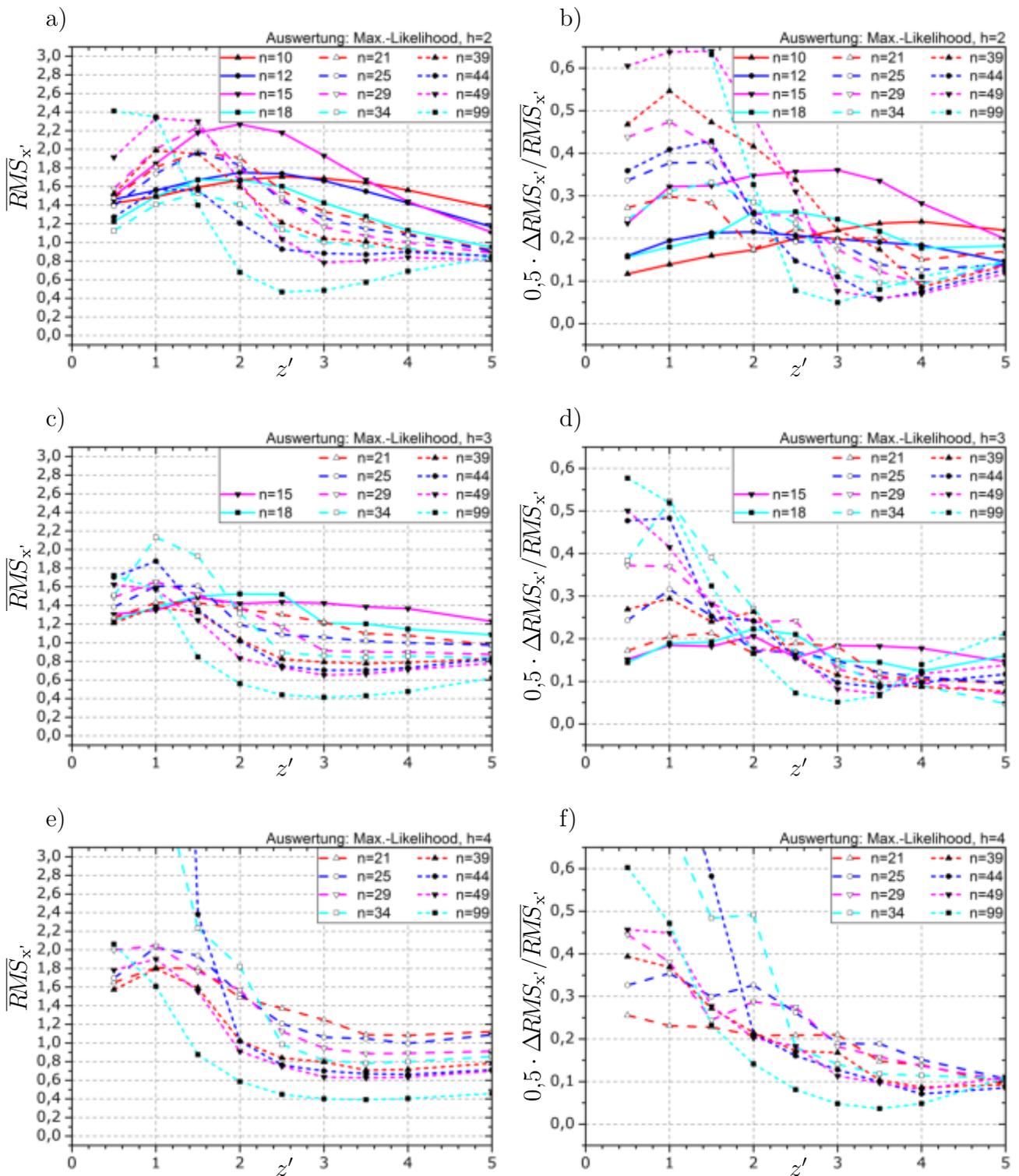


Abb. A.72: Mittlerer RMS-Fehler $\overline{RMS}_{x'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{x'} / \overline{RMS}_{x'}$ des Schätzers x' für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

A.5 Schätzung des 2,5%-Quantils bei bekannter Standardabweichung

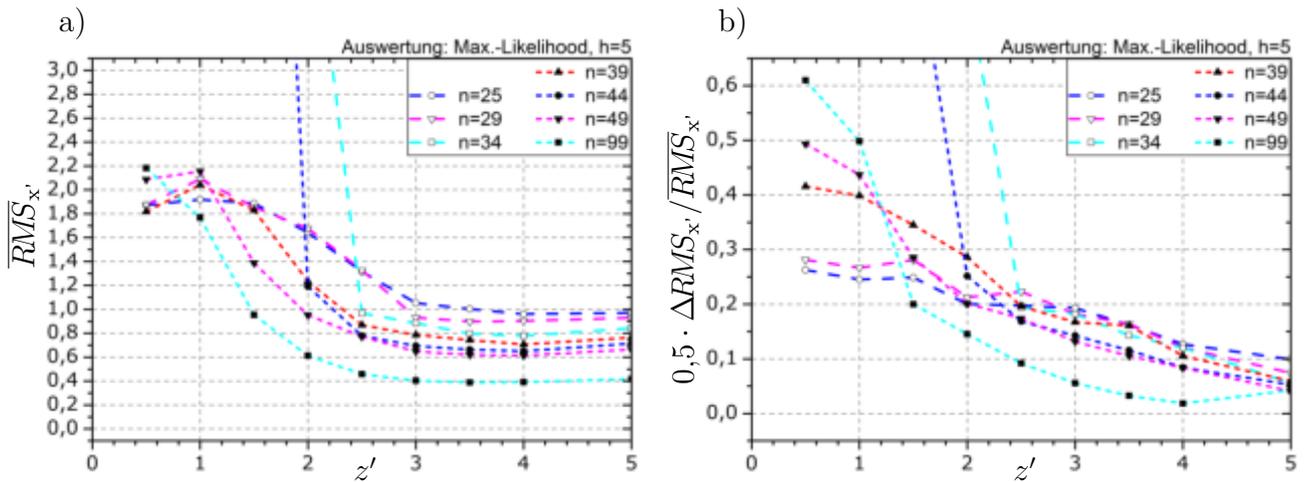


Abb. A.73: Mittlerer RMS-Fehler $\overline{RMS}_{x'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{x'} / \overline{RMS}_{x'}$ des Schätzers x' für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

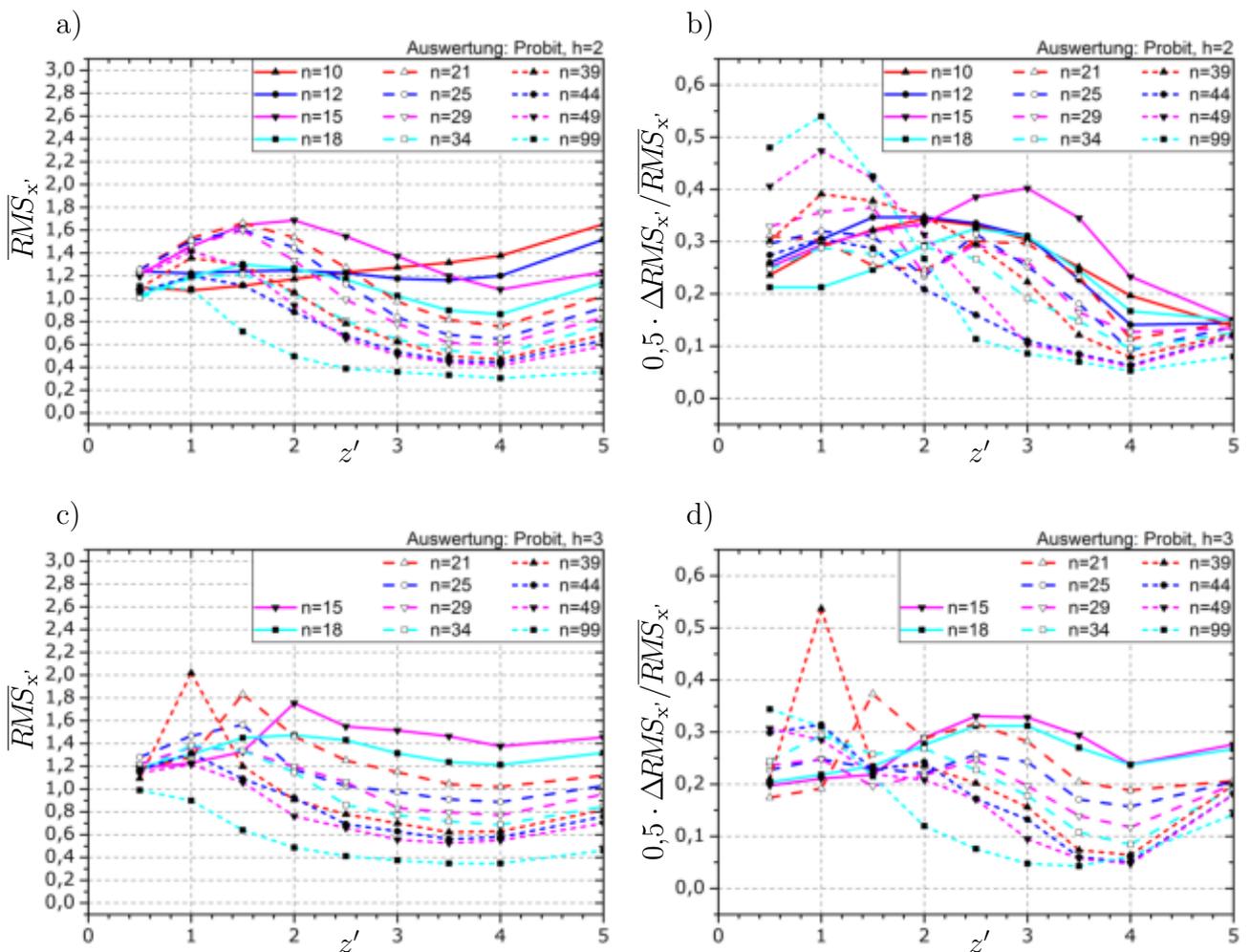


Abb. A.74: Mittlerer RMS-Fehler $\overline{RMS}_{x'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{x'} / \overline{RMS}_{x'}$ des Schätzers x' für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

A Grafische Darstellung der Simulationsergebnisse

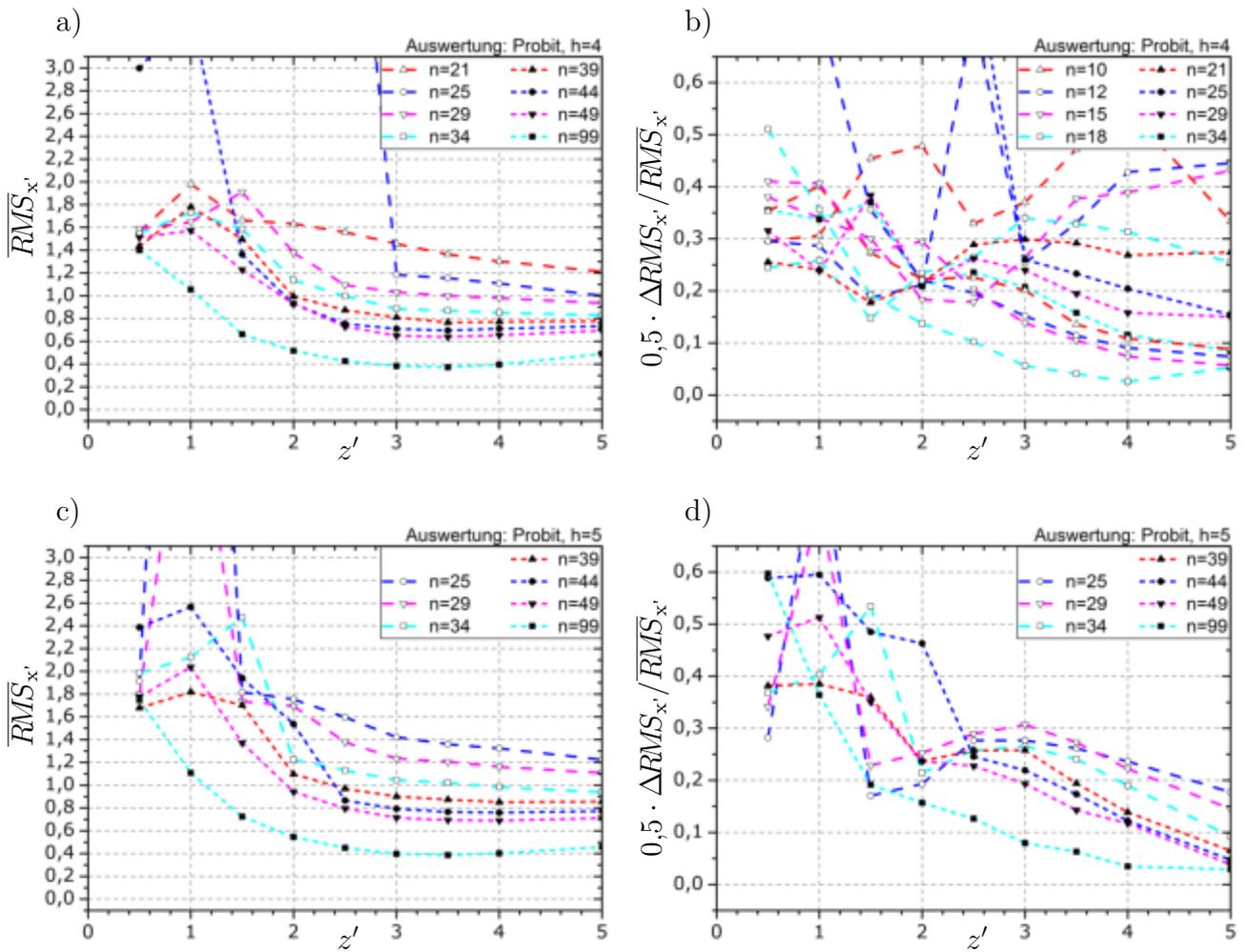


Abb. A.75: Mittlerer RMS-Fehler $\overline{RMS}_{x'}$ und relative Abweichung $0,5 \cdot \Delta RMS_{x'} / \overline{RMS}_{x'}$ des Schätzers x' für Horizontversuche. Auswertung mittels PROBIT-METHODE, für Versuche auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.5.3 Mittlere 95%-Quantile des Schätzers x'

Simulationsergebnisse der Treppenstufenversuche

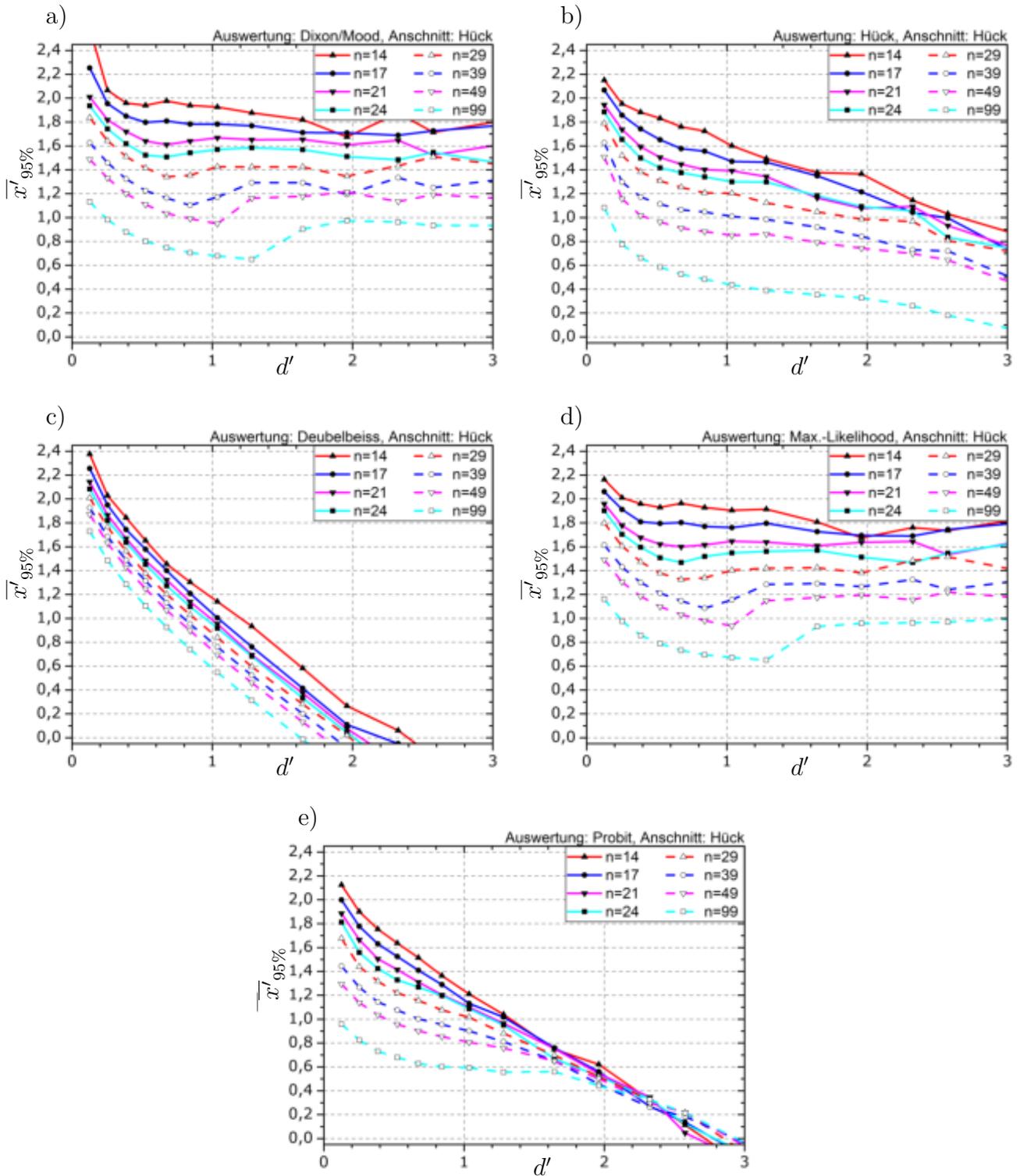


Abb. A.76: Mittlere 95%-Quantile $\overline{x'_{95\%}}$ des Schätzers x' für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK. Für die Auswertemethode nach DIXON/MOOD in a), nach HÜCK in b), nach DEUBELBEISS in c), MAX.-LIKELIHOOD-METHODE in d) sowie PROBIT-METHODE in e).

Simulationsergebnisse der Horizontversuche

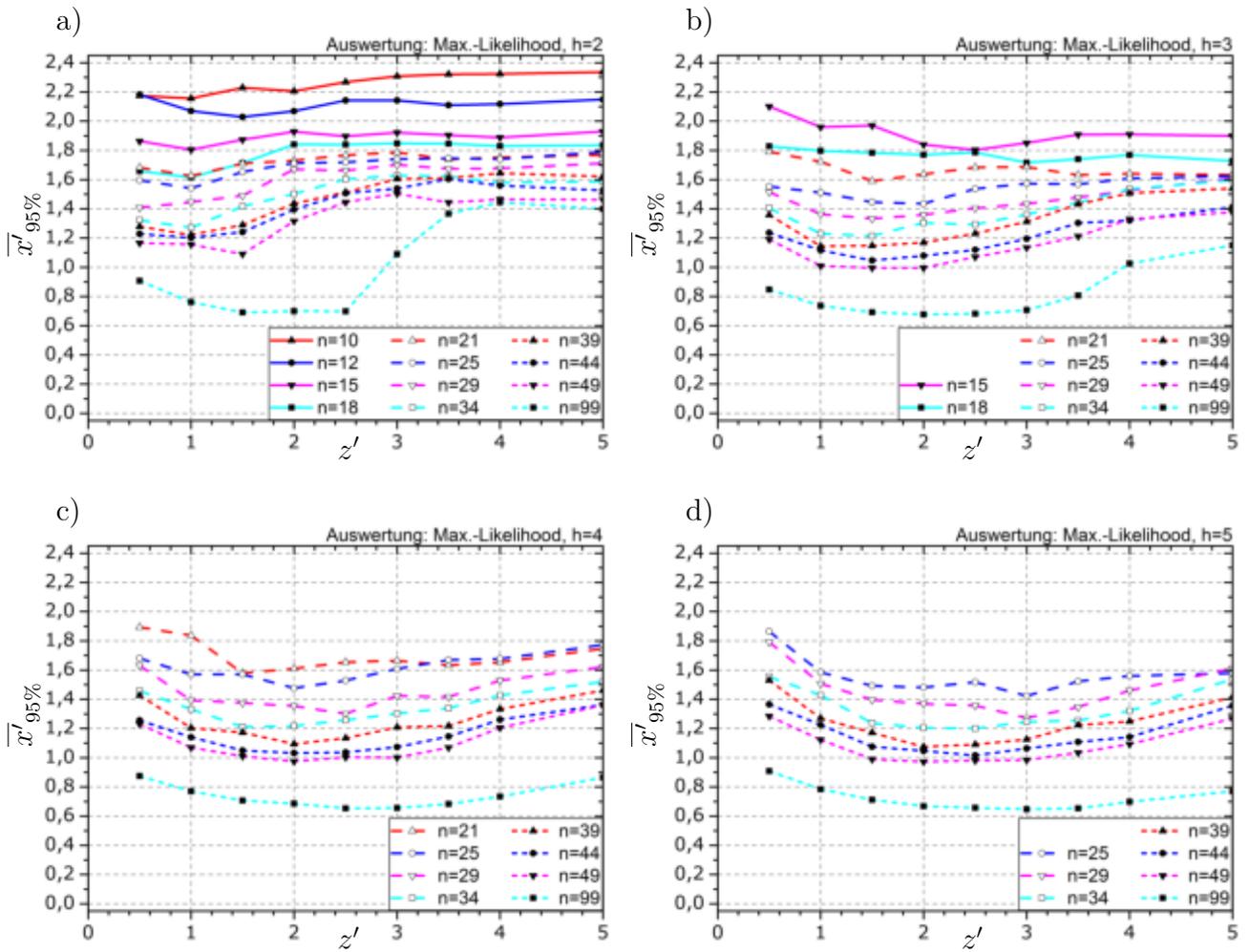


Abb. A.77: Mittlere 95%-Quantile $\bar{x}'_{95\%}$ des Schätzers x' aus Horizontversuchen für die Auswertung mittels MAX.-LIKELIHOOD-METHODE und Versuchen auf 2 Horizonten in a), 3 Horizonten in b), 4 Horizonten in c) sowie 5 Horizonten in d).

A.5 Schätzung des 2,5%-Quantils bei bekannter Standardabweichung

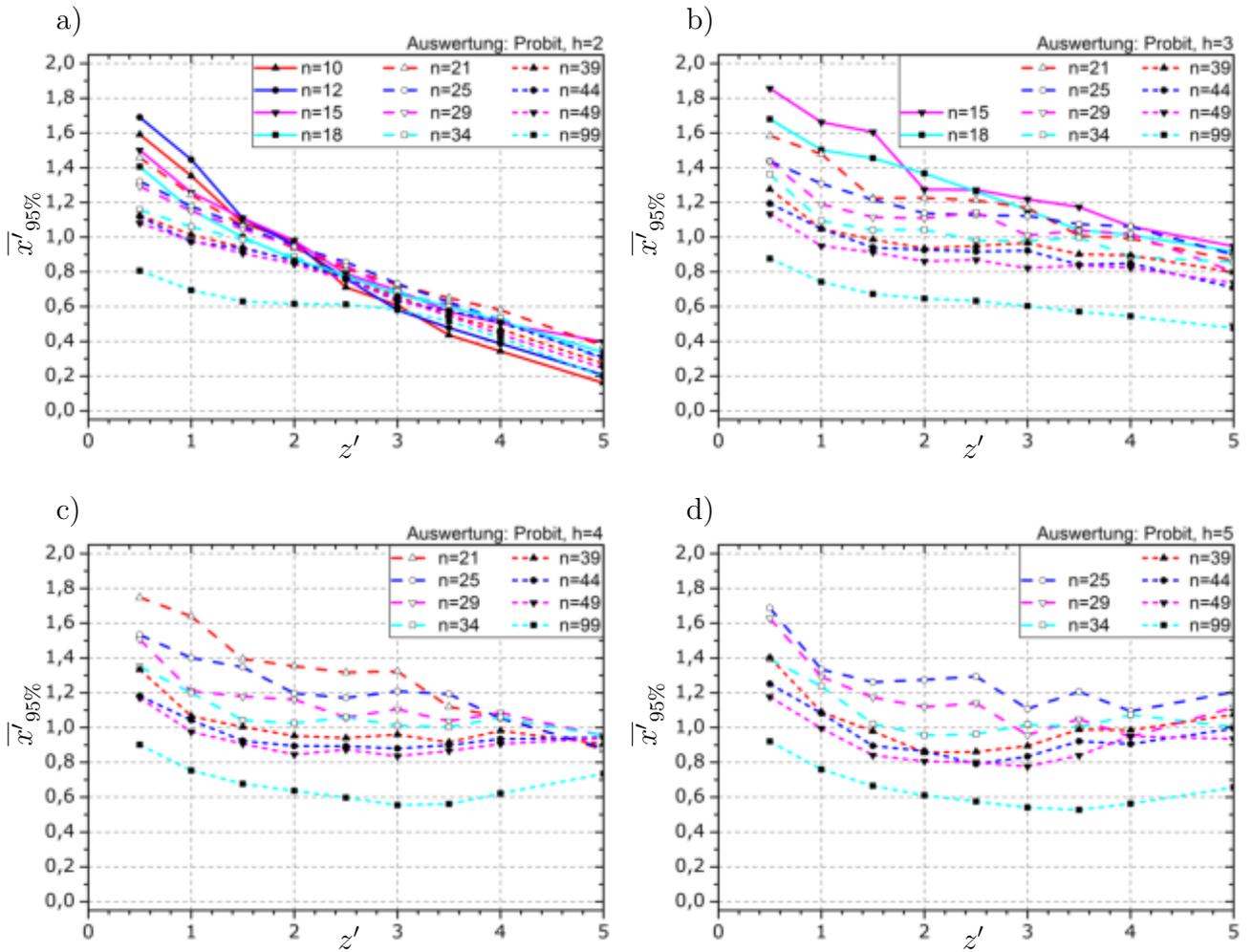


Abb. A.78: Mittlere 95%-Quantile $\bar{x}'_{95\%}$ des Schätzers x' aus Horizontversuchen für die Auswertung mittels PROBIT-METHODE und Versuchen auf 2 Horzonten in a), 3 Horzonten in b), 4 Horzonten in c) sowie 5 Horzonten in d).

A.6 Schätzung des 2,5%-Quantils bei unbekannter Standardabweichung

A.6.1 Mittlere Verzerrungen und Standardfehler des Schätzers x^*

Simulationsergebnisse der Treppenstufenversuche

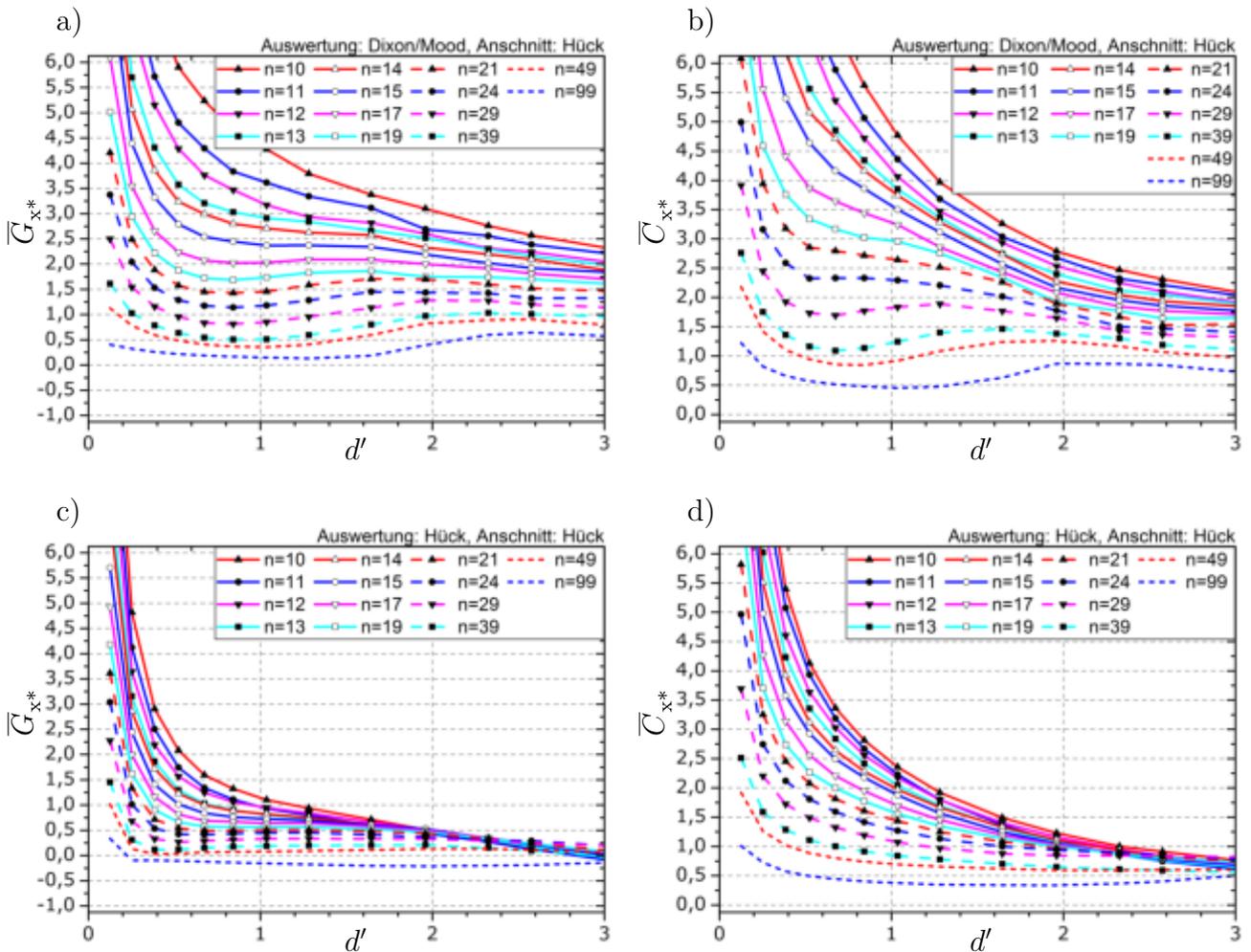


Abb. A.79: Mittlere Verzerrungen und Standardfehler des Schätzers x^* für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

A.6 Schätzung des 2,5%-Quantils bei unbekannter Standardabweichung

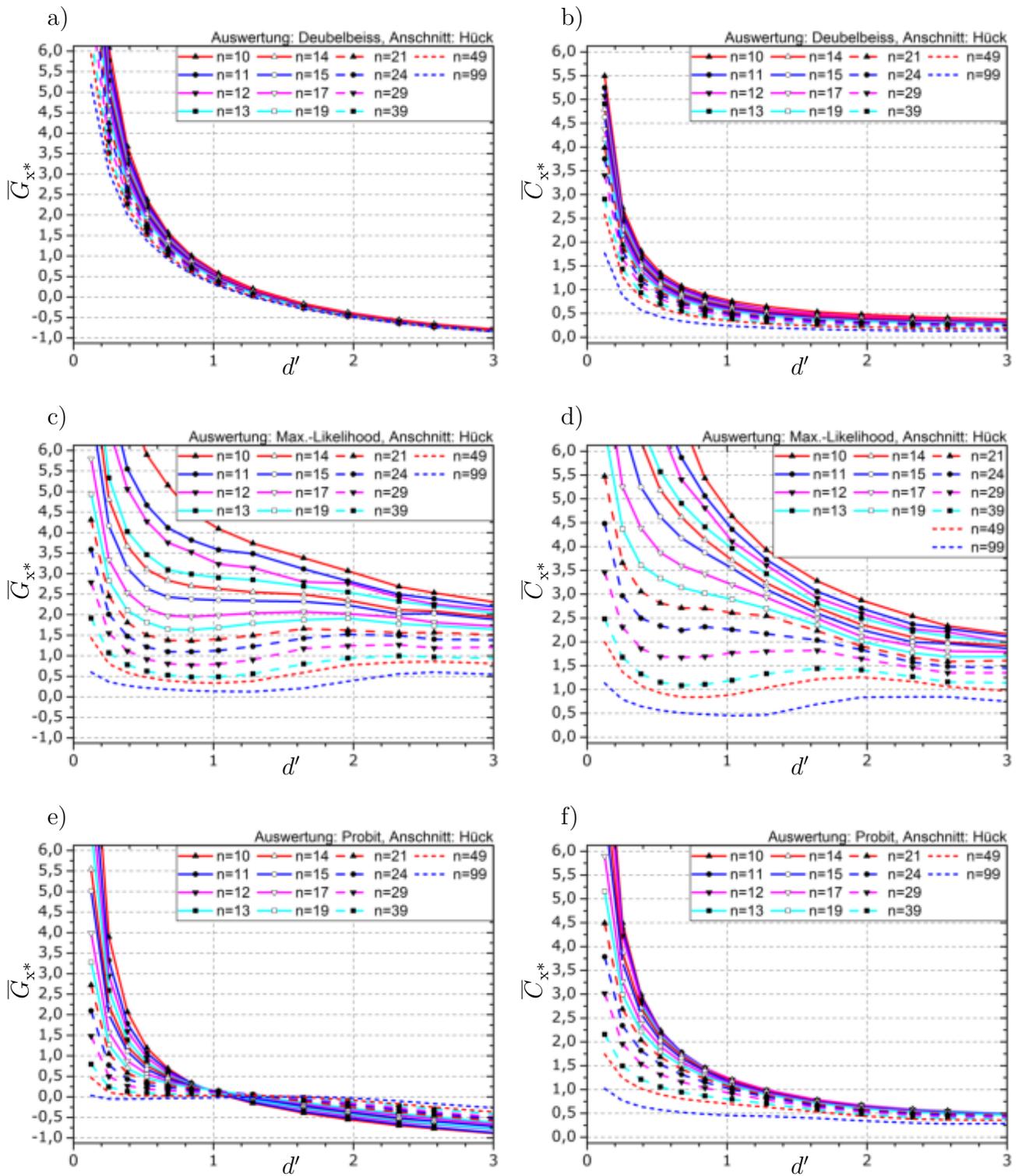


Abb. A.80: Mittlere Verzerrungen und Standardfehler des Schätzers x^* für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

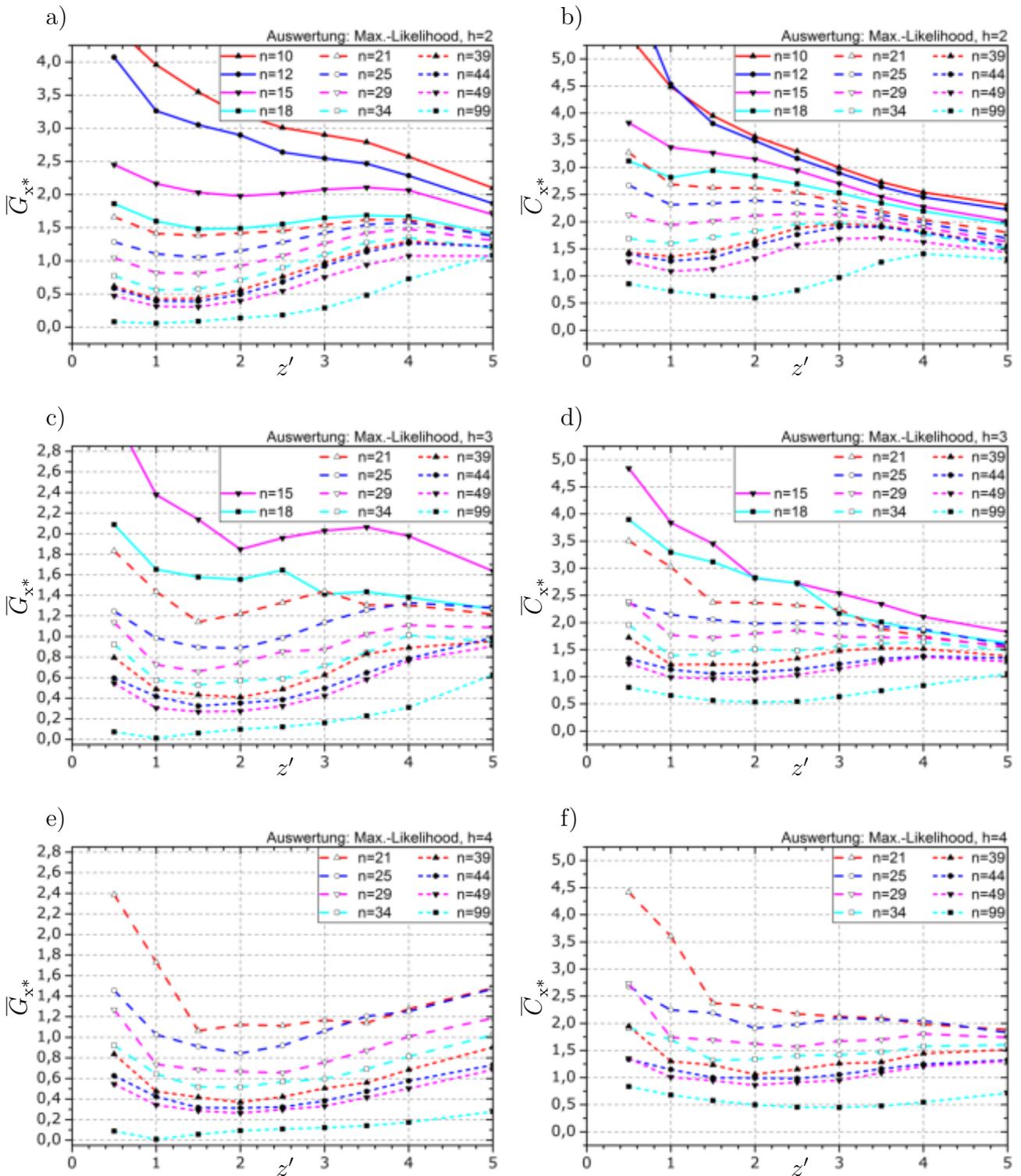


Abb. A.81: Mittlere Verzerrungen und Standardfehler des Schätzers x^* für Horizontversuche mit Auswertung nach MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

A.6 Schätzung des 2,5%-Quantils bei unbekannter Standardabweichung

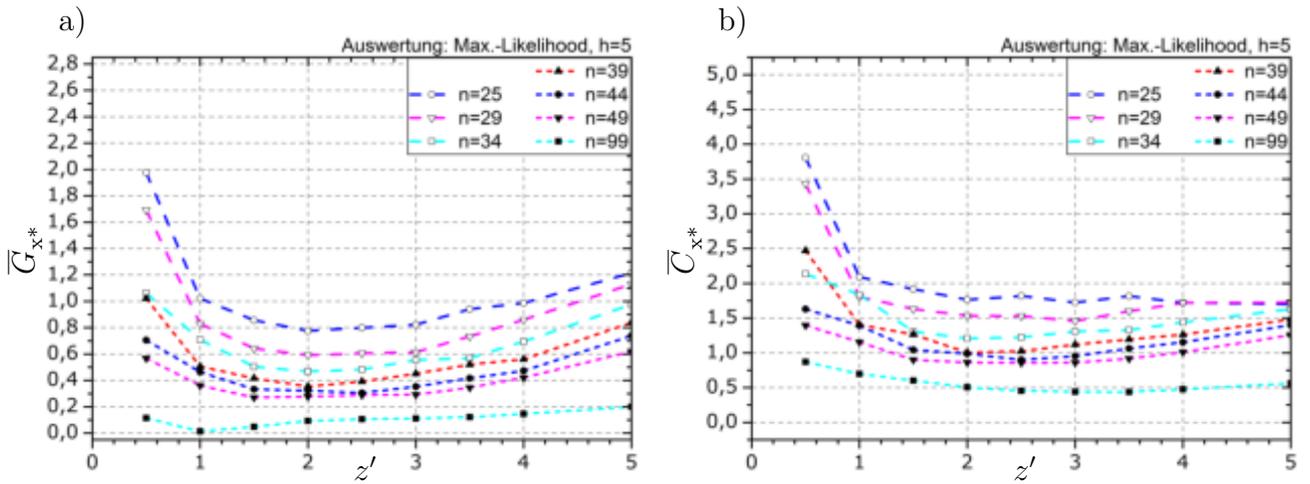


Abb. A.82: Mittlere Verzerrungen und Standardfehler des Schätzers x^* für Horizontversuche mit Auswertung nach MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

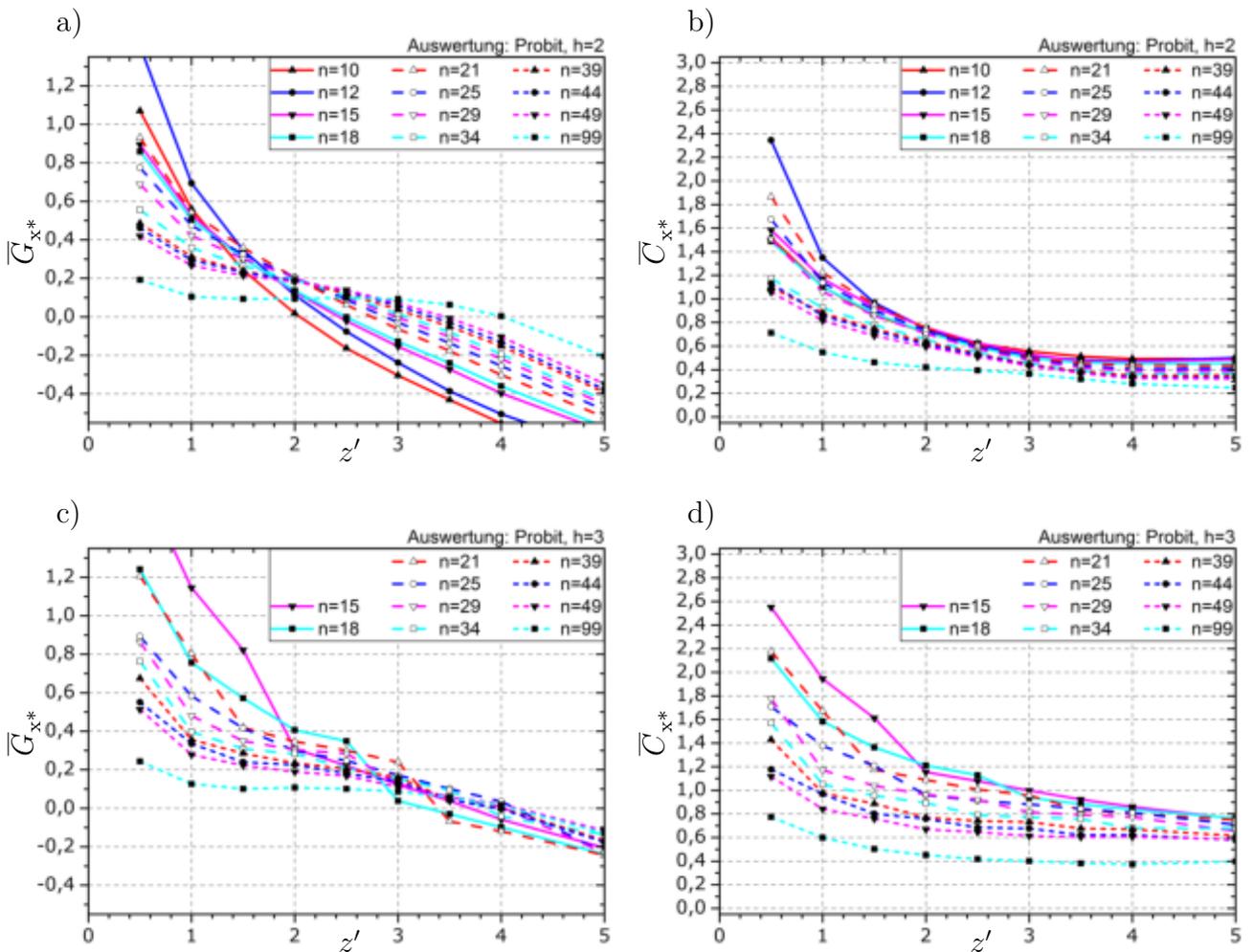


Abb. A.83: Mittlere Verzerrungen und Standardfehler des Schätzers x^* für Horizontversuche mit Auswertung nach PROBIT-METHODE, für Versuche auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

A Grafische Darstellung der Simulationsergebnisse

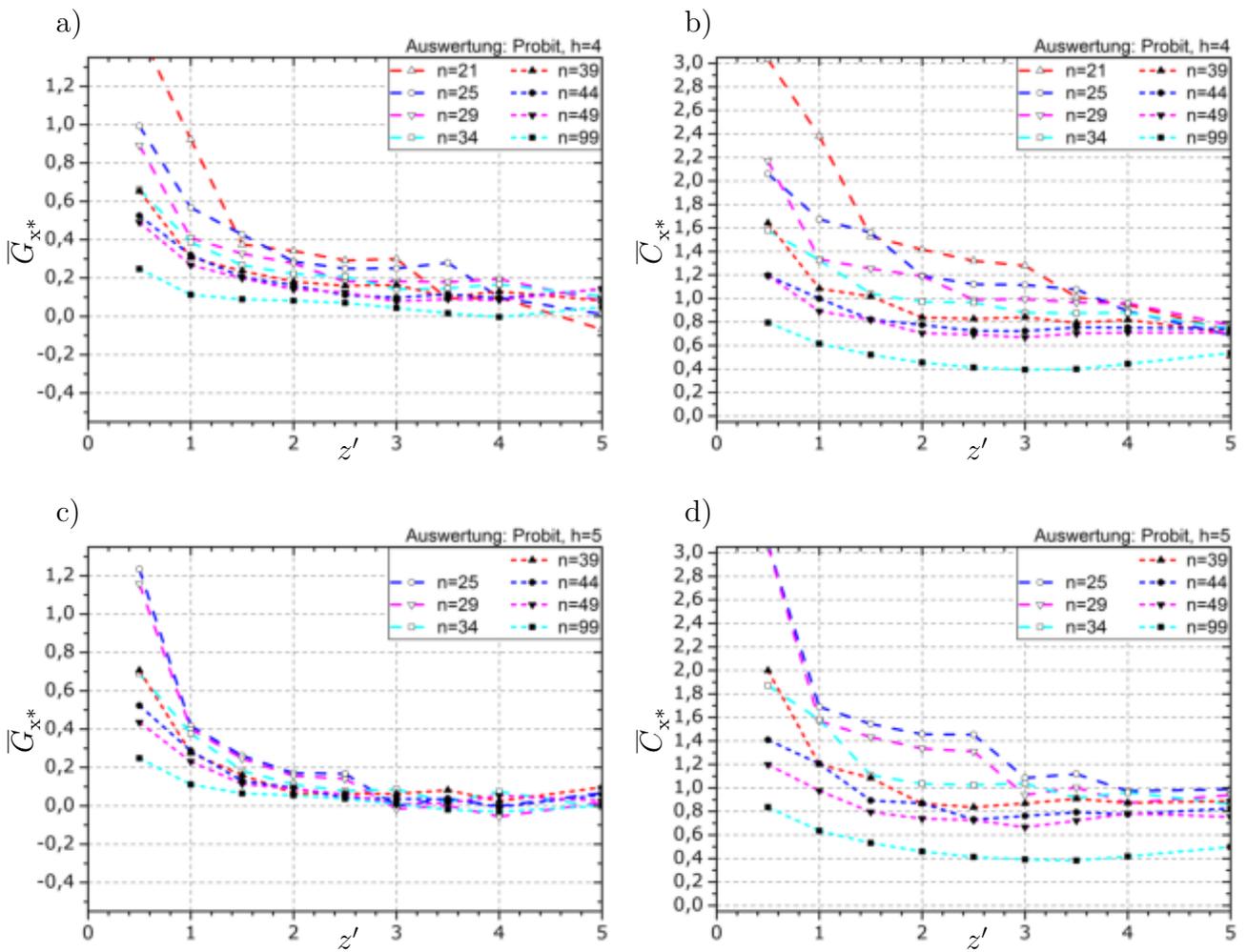


Abb. A.84: Mittlere Verzerrungen und Standardfehler des Schätzers x^* für Horizontversuche mit Auswertung nach PROBIT-METHODE, für Versuche auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.6.2 RMS-Fehler des Schätzers x^*

Simulationsergebnisse der Treppenstufenversuche

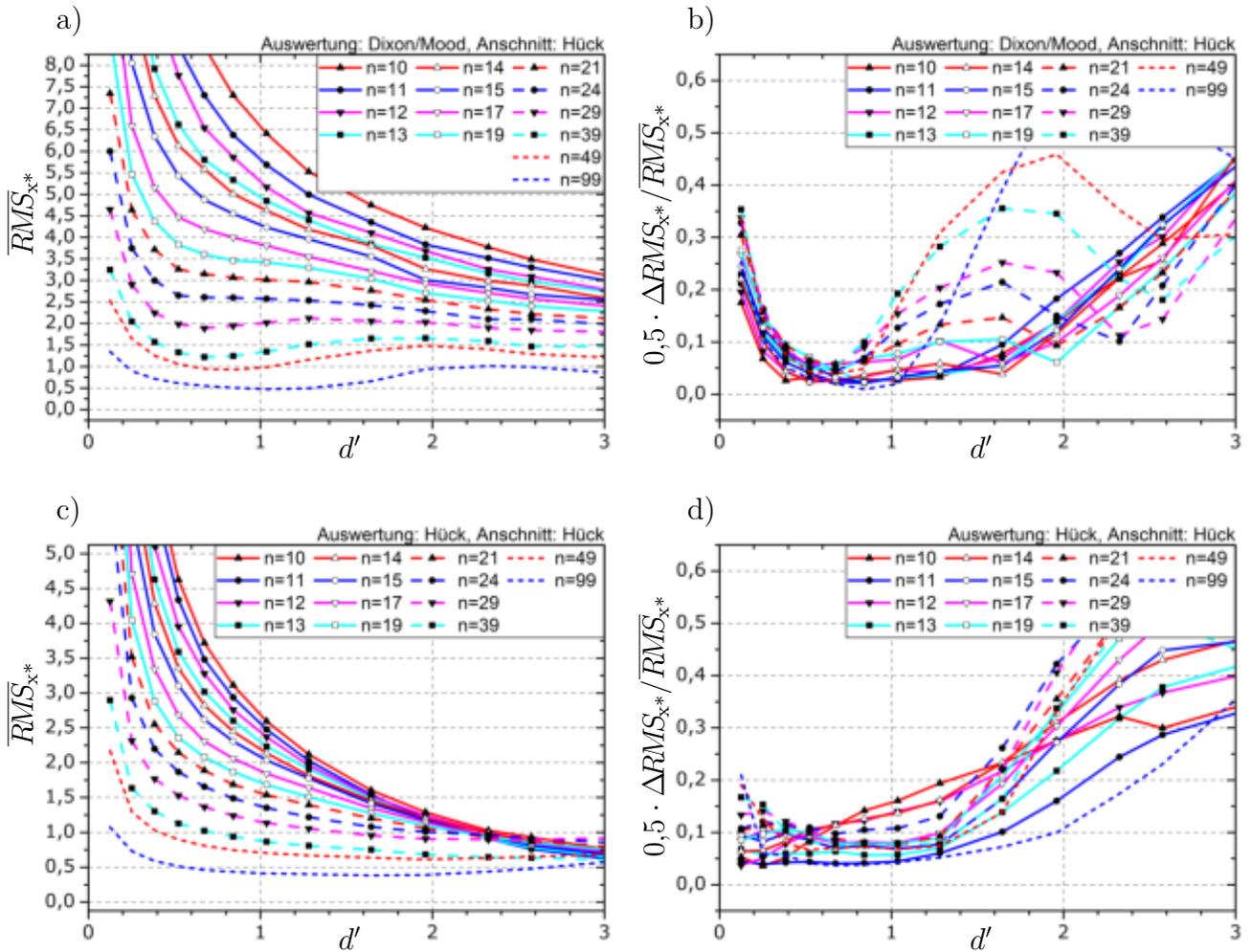


Abb. A.85: Mittlere RMS-Fehler \overline{RMS}_{x^*} und relative Abweichungen $0,5 \cdot \Delta RMS_{x^*} / \overline{RMS}_{x^*}$ des Schätzers x^* für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für die Auswertemethode nach DIXON/MOOD in a) und b) sowie HÜCK in c) und d).

A Grafische Darstellung der Simulationsergebnisse

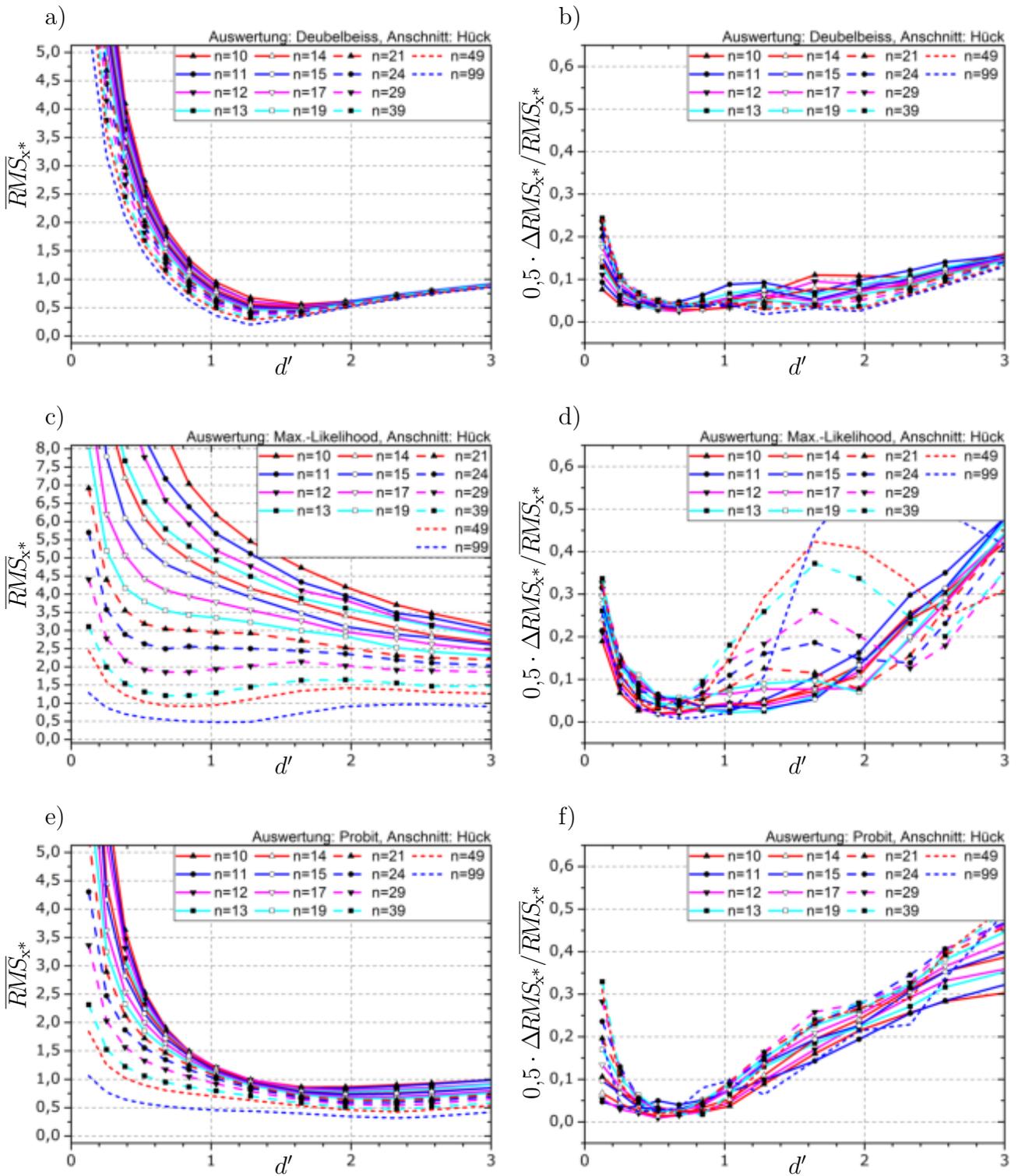


Abb. A.86: Mittlere RMS-Fehler \overline{RMS}_{x^*} und relative Abweichungen $0,5 \cdot \Delta RMS_{x^*} / \overline{RMS}_{x^*}$ des Schätzers x^* für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK, für Auswertemethoden nach DEUBELBEISS in a) und b), MAX.-LIKELIHOOD-METHODE in c) und d) sowie PROBIT-METHODE in e) und f).

Simulationsergebnisse der Horizontversuche

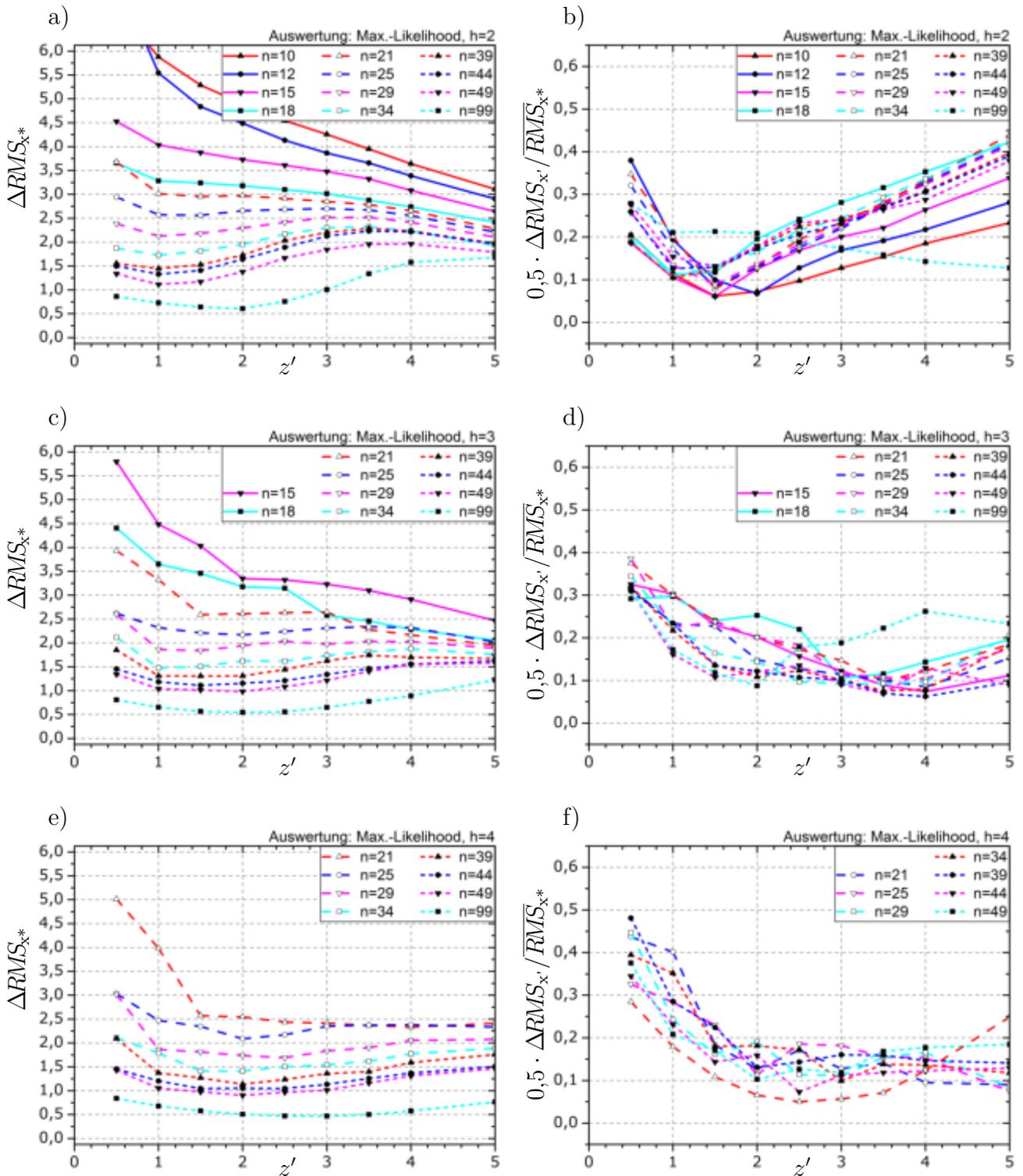


Abb. A.87: Mittlerer RMS-Fehler \overline{RMS}_{x^*} und relative Abweichung $0,5 \cdot \Delta RMS_{x^*} / \overline{RMS}_{x^*}$ des Schätzers x^* für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b), 3 Horizonten in c) und d) sowie 4 Horizonten in e) und f).

A Grafische Darstellung der Simulationsergebnisse

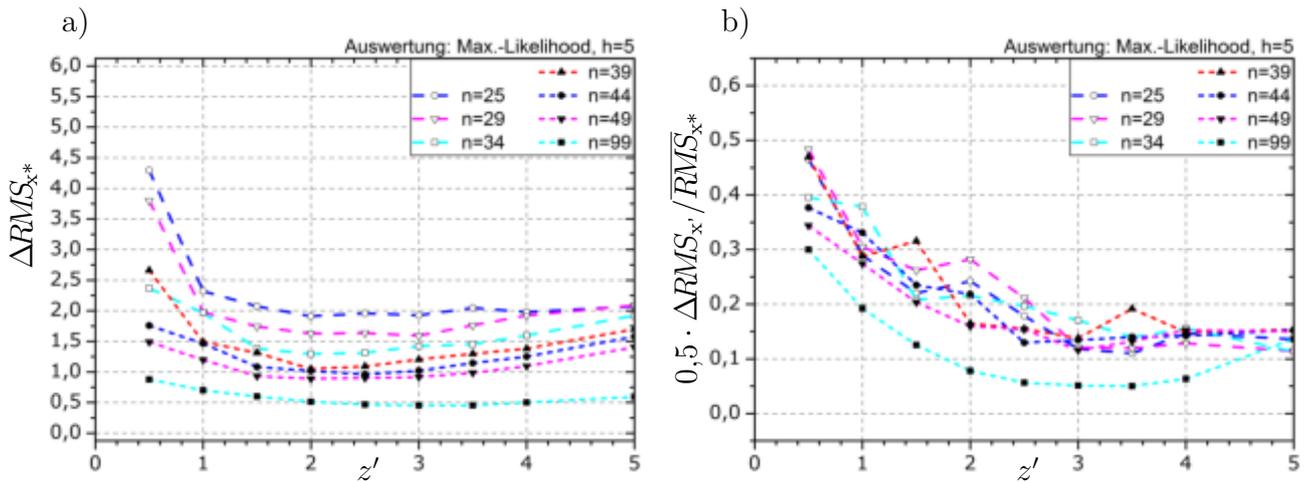


Abb. A.88: Mittlere RMS-Fehler \overline{RMS}_{x^*} und relative Abweichungen $0,5 \cdot \Delta RMS_{x^*} / \overline{RMS}_{x^*}$ des Schätzers x^* für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 5 Horizonten.

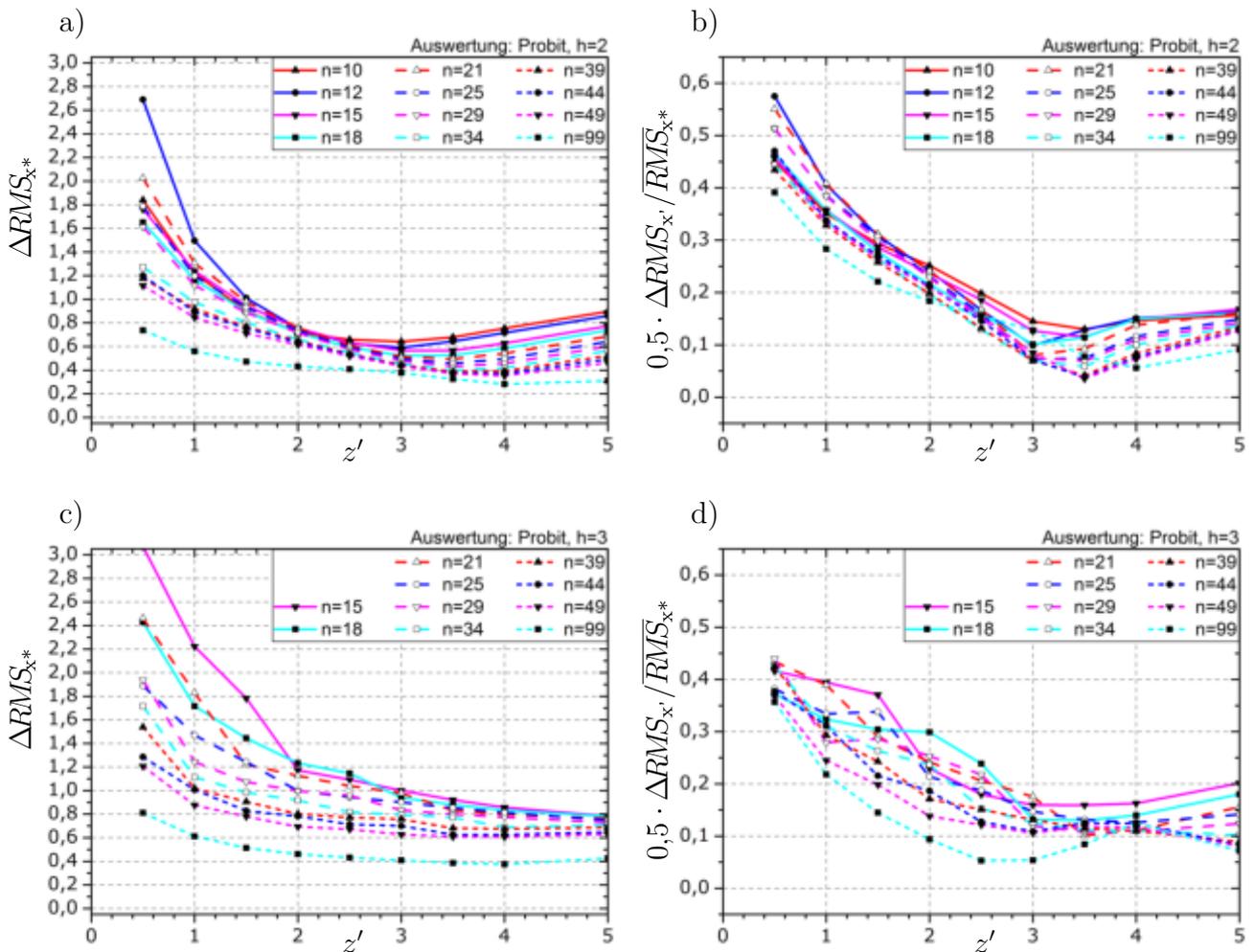


Abb. A.89: Mittlere RMS-Fehler \overline{RMS}_{x^*} und relative Abweichungen $0,5 \cdot \Delta RMS_{x^*} / \overline{RMS}_{x^*}$ des Schätzers x^* für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 2 Horizonten in a) und b) sowie 3 Horizonten in c) und d).

A.6 Schätzung des 2,5%-Quantils bei unbekannter Standardabweichung

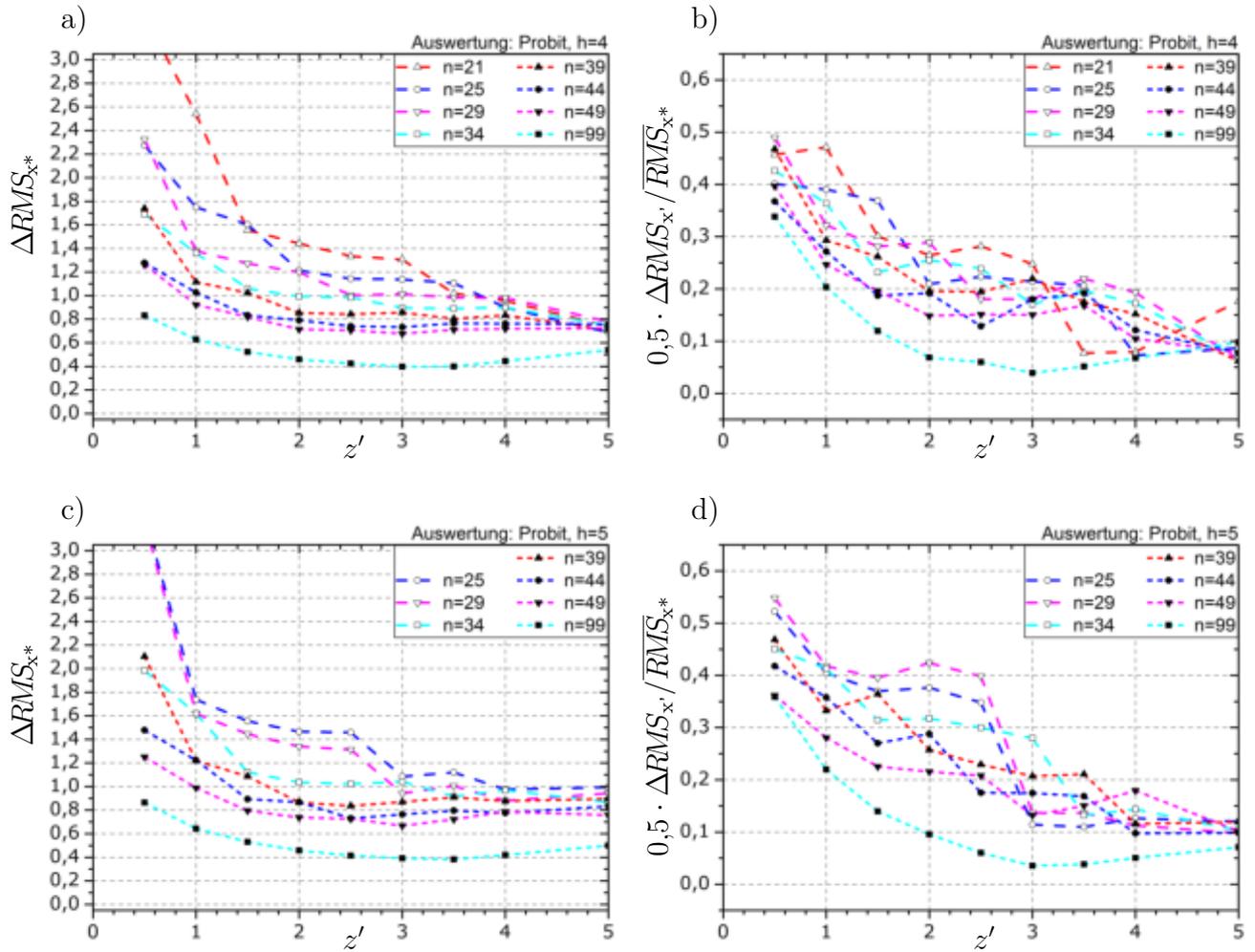


Abb. A.90: Mittlerer RMS-Fehler \overline{RMS}_{x^*} und relative Abweichung $0,5 \cdot \Delta RMS_{x^*} / \overline{RMS}_{x^*}$ des Schätzers x^* für Horizontversuche. Auswertung mittels MAX.-LIKELIHOOD-METHODE, für Versuche auf 4 Horizonten in a) und b) sowie 5 Horizonten in c) und d).

A.6.3 Mittlere 95%-Quantile des Schätzers x^*

Simulationsergebnisse der Treppenstufenversuche

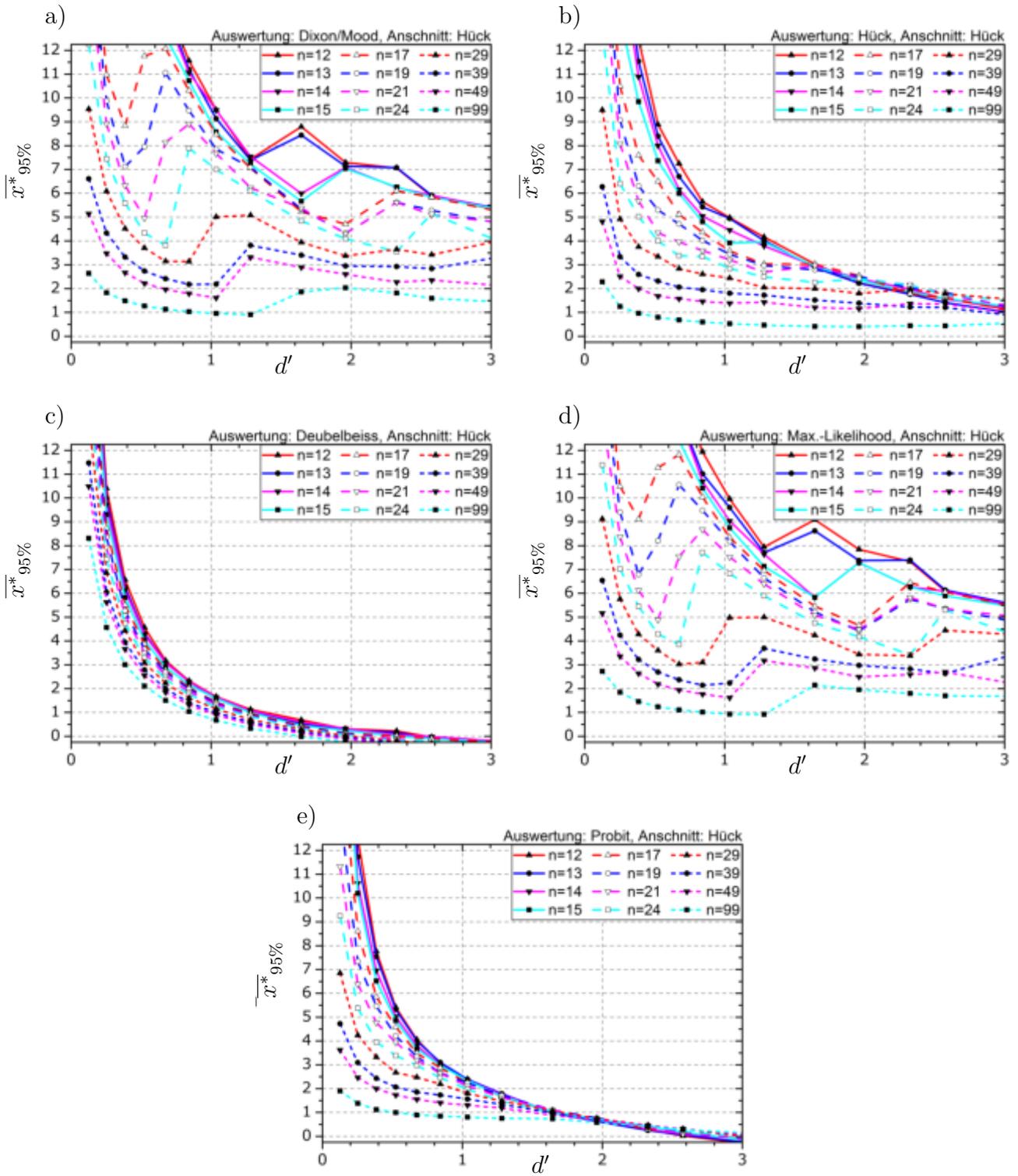


Abb. A.91: Mittlere 95%-Quantile $\overline{x^*}_{95\%}$ des Schätzers x^* für Treppenstufenversuche mit Anschnittsdefinition nach HÜCK. Für die Auswertemethode nach DIXON/MOOD in a), nach HÜCK in b), nach DEUBELBEISS in c), MAX.-LIKELIHOOD-METHODE in d) sowie PROBIT-METHODE in e).

Simulationsergebnisse der Horizontversuche

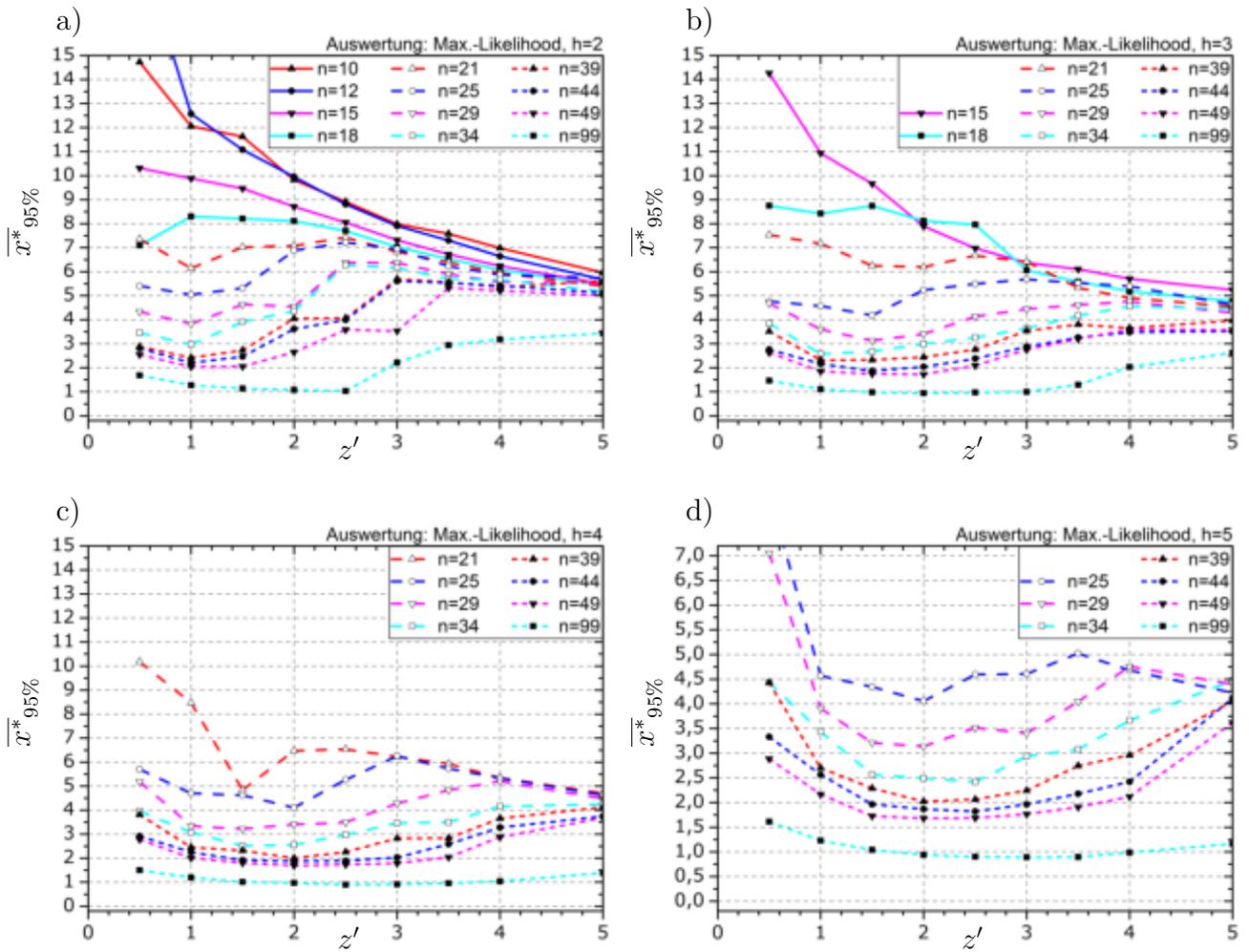


Abb. A.92: Mittlere 95%-Quantile $\bar{x}^*_{95\%}$ des Schätzers x^* aus Horizontversuchen für die Auswertung mittels MAX.-LIKELIHOOD-METHODE und Versuchen auf 2 Horzonten in a), 3 Horzonten in b), 4 Horzonten in c) sowie 5 Horzonten in d).

A Grafische Darstellung der Simulationsergebnisse

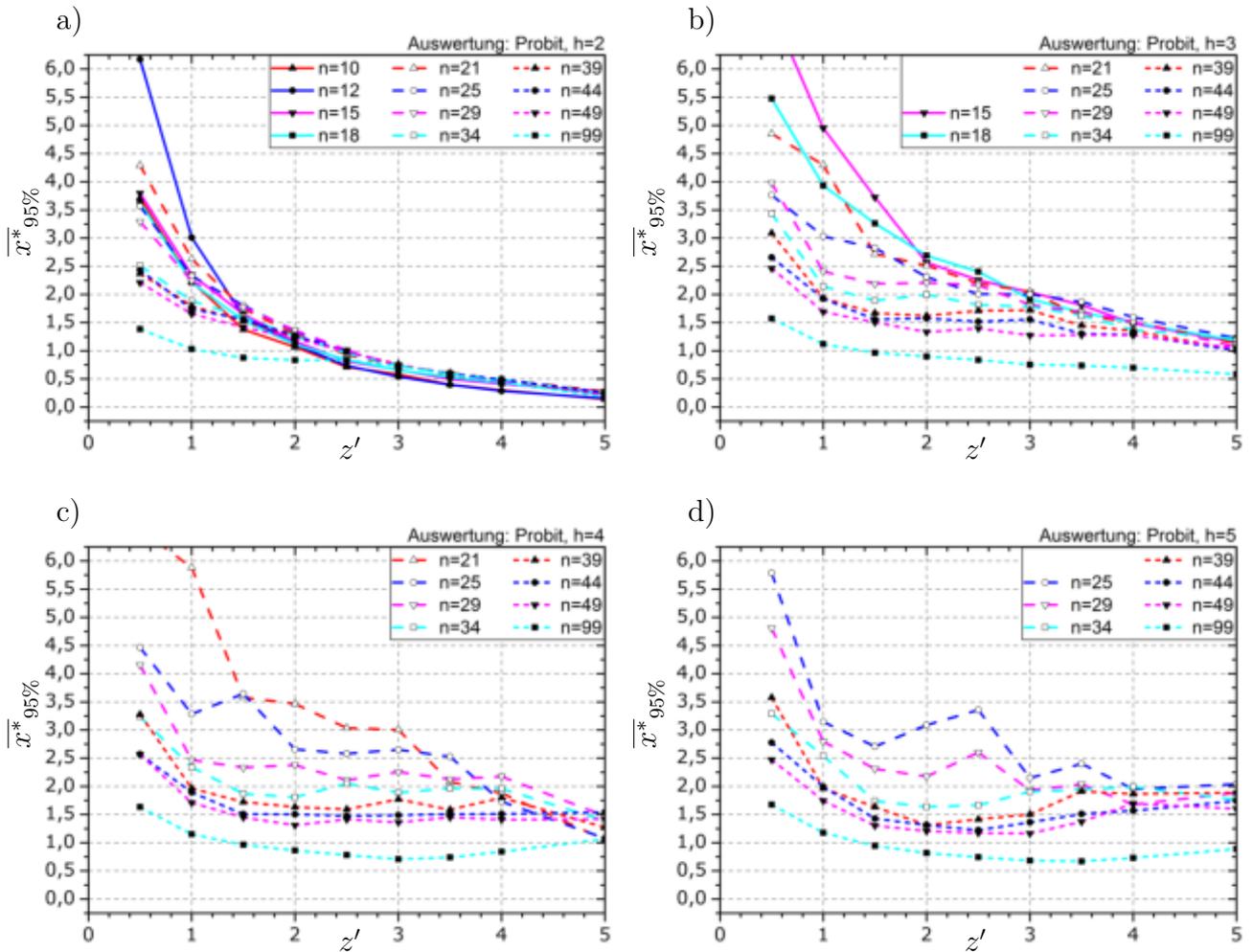


Abb. A.93: Mittlere 95%-Quantile $\overline{x^*}_{95\%}$ des Schätzers x^* aus Horizontversuchen für die Auswertung mittels PROBIT-METHODE und Versuchen auf 2 Horzonten in a), 3 Horzonten in b), 4 Horzonten in c) sowie 5 Horzonten in d).

B Sonstiges

B.1 Treppenstufenauswertung

B.1.1 Treppenstufenauswertung nach Hück

B.1.1.1 Schätzung der Standardabweichung

Für die Beschreibung der von HÜCK in [21] veröffentlichten Korrekturfunktionen aus Abbildung 2.9 a) existieren verschiedene Ansätze. Von HÜCK selbst [23] stammt für $k_H > 0,5$ und $n \geq 17$ der Formelsatz aus Gl. B.1 bis B.4.

$$\frac{s}{d} = A_1 \cdot (A_2 \cdot k_H)^{A_3} + 1,6 \cdot (A_2 \cdot k_H) - 0,35 \quad (\text{B.1})$$

$$A_1 = 1767 \cdot (n+1)^{-2,734} + 0,16 \quad (\text{B.2})$$

$$A_2 = -0,00018 \cdot (n+1)^{-2,07} + 0,014 \cdot (n+1) + 0,827 \quad (\text{B.3})$$

$$A_3 = 2732 \cdot (n+1)^{-2,765} + 2,458 \quad (\text{B.4})$$

Von MÜLLER wird in [31] das gleiche Diagramm durch die Gleichungen B.5 bis B.7 beschrieben.

$$\frac{s}{d} = \begin{cases} 10^{B_1} \cdot k_H^{B_2} & \text{für } k_H \geq 0,5 \\ 10^{B_1} \cdot 0,5^{B_2} & \text{für } k_H < 0,5 \end{cases} \quad (\text{B.5})$$

$$B_1 = 4,579494 \cdot (n+1)^{-0,889521} \quad (\text{B.6})$$

$$B_2 = 7,235548 \cdot (n+1)^{-0,405229} \quad (\text{B.7})$$

Innerhalb des Instituts für Festkörpermechanik der Technischen Universität Dresdens werden in den ProgrammROUTINEN zur Auswertung von Treppenstufenversuchen die von HANTSCHKE [15] ermittelten Gleichungen B.8 bis B.11 verwendet.

$$\frac{s}{d} = \max\{C_1 \cdot k_H^3 + C_2 \cdot k_H + C_3; 1,6 \cdot k_H - 0,34\} - \dots \quad (\text{B.8})$$
$$\dots \max\{C_1 \cdot k_0^3 + C_2 \cdot k_0 + C_3; 1,6 \cdot k_0 - 0,34\}$$

$$C_1 = 823,414 \cdot (n+1)^{-2,3352} \quad (\text{B.9})$$

$$C_2 = 1,913277 \cdot C_1^{0,01425556} \quad (\text{B.10})$$

$$C_3 = -0,77365 \cdot C_1^{0,186102} \quad (\text{B.11})$$

$$k_0 = 0,2125 \quad (\text{B.12})$$

Für sehr geringe und sehr große Stichprobenumfänge unterscheiden sich die Funktionswerte der Korrekturfunktionen erheblich. Während die letzte Beschreibung das Diagramm 2.9 a) auf S. 50 über für den größten Wertebereich am genauesten wieder gibt, sind die Beschreibung von HÜCK für $n + 1 < 15$ bzw. $n + 1 > 55$ und die Beschreibung von MÜLLER für $n + 1 > 60$ unbrauchbar. Der Beschreibung durch MÜLLER ist eine eigene Simulation von Treppenstufenversuchen vorangegangen. Ob dabei ein mit Abbildung 2.9 a) deckungsgleiches Korrekturdiagramm ermittelt wurde ist nicht dokumentiert. Daher lässt sich die Frage, welche Korrekturfunktion zu verwenden ist nicht aus einem relativen Vergleich heraus beantworten. Ohnehin wird der für die praktische Arbeit relevante Wertebereich von $15 < n < 50$ von allen drei Funktionen annähernd identisch beschrieben.

B.1.1.2 Schätzung der Standardfehler

Zur rechnerischen Bestimmung der Standardfehler aus den von HÜCK in [21] angegebenen und in Abbildung 2.10 auf S. 50 dargestellten Diagrammen werden von ihm selbst in [23] für $n > 17$ Gleichungen angegeben. Der Standardfehler des Schätzers des Erwartungswertes folgt demnach aus den Gleichungen B.13 bis B.17.

$$\frac{s_m}{s} = A_1 \cdot \left(\frac{s}{d}\right)^{A_2} + A_3 \cdot \left(\frac{s}{d}\right)^2 + A_4 \quad (\text{B.13})$$

$$A_1 = 0,625 \cdot (n + 1)^{-0,848} - 0,013 \quad (\text{B.14})$$

$$A_2 = -0,0021 \cdot (n + 1)^{1,74} - 0,5 \quad (\text{B.15})$$

$$A_3 = -0,1 \cdot (n + 1)^{-0,821} \quad (\text{B.16})$$

$$A_4 = 1,579 \cdot (n + 1)^{-0,0614} - 1,042 \quad (\text{B.17})$$

Der Standardfehler des Schätzers der Standardabweichung lässt sich aus den Gleichungen B.18 bis B.20 berechnen.

$$\frac{s_s}{d} = B_1 \cdot \left(\frac{s}{d}\right)^{B_2} \quad (\text{B.18})$$

$$B_1 = 405 \cdot (n + 1)^{-2,15} + 0,224 \quad (\text{B.19})$$

$$B_2 = \begin{cases} -35699 \cdot (n + 1)^{-4,18} - 0,0132 \cdot (n + 1) + 2,486 & \text{für } d \leq 22 \\ -12,7 \cdot (n + 1)^{-0,89} - 0,0305 \cdot (n + 1) + 3,59 & \text{für } d > 22 \end{cases} \quad (\text{B.20})$$

Von HANTSCHKE wurden für $n < 150$ die Gleichungen B.21 bis B.25 zur Beschreibung des Standardfehlers der Erwartungswertschätzung

$$\frac{s_m}{s} = D_1 \cdot \left(\frac{s}{d}\right)^{D_2} + D_3 \cdot \left(\frac{s}{d}\right) + D_4 \quad (\text{B.21})$$

$$D_1 = 1,205926 \cdot (n + 1)^{-0,673011} - 0,01196 \quad (\text{B.22})$$

$$D_2 = -0,000284458 \cdot (n + 1)^{1,349893} - 0,32165 \quad (\text{B.23})$$

$$D_3 = -0,072286444 \cdot (n + 1)^{-0,720352} + 0,00522 \quad (\text{B.24})$$

$$D_4 = 1,315079 \cdot (n + 1)^{-0,027113} - 1,051 \quad (\text{B.25})$$

und B.26 bis B.28 für verschiedene Standardfehler der Standardabweichungsschätzung aus den Diagrammen in Abbildung 2.10 abgeleitet.

$$\frac{s_s}{d} = E_1 \cdot \left(\frac{s}{d}\right)^{E_2} \quad (\text{B.26})$$

$$E_1 = 0,089708 \cdot e^{3,754456 \cdot (\lg(n+1))^{-1,926531}} \quad (\text{B.27})$$

$$E_2 = 2,325673 \cdot e^{-0,006177 \cdot (n+1)} \quad (\text{B.28})$$

B.2 Zur Umrechnungen zwischen Verteilungsparametern

Sind die Verteilungsparameter eines Verteilungsmodells bestimmt, kann eine näherungsweise Umrechnung auf ein anderes Verteilungsmodell erfolgen, indem die Forderung aufgestellt wird, dass die Verteilungsfunktionen in bestimmten Punkten übereinstimmen. Um die Streuung der Versuchsdaten zu erfassen werden dazu möglichst kleine bzw. möglichst große Quantile der Zufallsvariable verwendet, wobei darauf zu achten ist, dass diese nicht außerhalb des in der Messung beobachteten Bereichs liegen. Von HAIBACH werden in [14] für diese Näherung z. B. die Punkte, die sich im Abstand der Standardabweichung vom Mittelwert befinden, verwendet.

B.2.1 Näherung für die Parameter der Weibull-Verteilung

Wird die zweiparametrische WEIBULL-Verteilung

$$P(X \leq x) = F(x) = 1 - e^{-\left(\frac{x}{\alpha}\right)^k} \quad (\text{B.29})$$

doppelt logarithmiert, kann diese Form als eine Geradengleichung

$$\underbrace{\ln \left[\ln \left(\frac{1}{1 - F(x)} \right) \right]}_y = \underbrace{k \ln \left(\frac{x}{\alpha} \right)}_{m \cdot x} \quad (\text{B.30})$$

interpretiert werden, deren Anstieg durch den Exponenten bzw. Formparameter k dargestellt wird. Werden nun zwei Punkte auf dieser Geraden gewählt, dann berechnet sich der Anstieg der Geraden nach Gleichung B.32.

$$k = \frac{\ln \left[\frac{\ln \left(\frac{1}{1 - F(x_2)} \right)}{\ln \left(\frac{1}{1 - F(x_1)} \right)} \right]}{\ln \left[\frac{x_2}{x_1} \right]} \quad (\text{B.31})$$

$$k = \frac{\ln \left[\ln [1 - F(x_2)] \right] - \ln \left[\ln [1 - F(x_1)] \right]}{\ln [x_2] - \ln [x_1]} \quad (\text{B.32})$$

Der Maßstabsparameter α der WEIBULL-Verteilung kann bestimmt werden, indem für einen der gewählten Punkte die Gleichung B.30 erfüllt wird.

B.2.2 Näherung aus logarithmischer Normalverteilung

Wurde die Auswertung unter Annahme einer logarithmisch normalverteilten Zufallsgröße, also $X \sim \mathcal{N}_L(\mu; \sigma^2)$ bzw. $Y = \ln(X)$, $Y \sim \mathcal{N}(\mu_L; \sigma_L^2)$, durchgeführt und die Streuspanne T aus dem Verhältnis der abgeschätzten Quantile $X_{90\%}$ zu $X_{10\%}$ bestimmt, dann gilt, wenn wie in [14] dazu die Punkte $Y_1 = \mu_L - \sigma_L$ und $Y_2 = \mu_L + \sigma_L$ verwendet werden:

$$k = \frac{1,315}{\lg T} \quad \text{bzw.} \quad k = \frac{3,029}{\ln T} \quad (\text{B.33})$$

Werden, wie in [24], die Quantile $Y_1 = Y_{10\%} = \mu_L - 1,2816 \sigma_L$ und $Y_2 = Y_{90\%} = \mu_L + 1,2816 \sigma_L$ ausgewählt, folgt:

$$k = \frac{1,34}{\lg T} \quad \text{bzw.} \quad k = \frac{3,084}{\ln T} \quad (\text{B.34})$$

Gleichung B.33 B.34 und gelten dabei unabhängig von der verwendeten Transformationsvorschrift $Y = \ln(X)$ oder $Y = \lg X$, da für beide die Gleiche Streuspanne bestimmt wird. Die dabei berechneten Standardabweichung σ_L unterschieden sich allerdings nach folgendem Zusammenhang:

$$\sigma_{\lg} = \frac{\sigma_{\ln}}{\ln 10} \quad (\text{B.35})$$

B.2.3 Näherung aus Normalverteilung

Wurde die Auswertung unter Annahme einer normalverteilten Zufallsgröße durchgeführt, ist also $X \sim \mathcal{N}(\mu; \sigma^2)$ dann ist die Streuspanne kein eindeutiges Maß für die Streuung, da sie zusätzlich vom Erwartungswert μ abhängig ist. Werden für x_1 und x_2 Werte symmetrisch zum Mittelwert verwendet, d.h. $x_{1,2} = \mu \mp u\sigma$, dann kann Gleichung B.32 mit den Verteilungsfunktion Φ der Normalverteilung beschrieben werden.

$$k = \frac{\ln [\ln [1 - \Phi(u)]] - \ln [\ln [\Phi(u)]]}{\ln \left[\frac{x_2}{x_1} \right]} = \frac{C_{\ln}}{\ln \left[\frac{x_2}{x_1} \right]} = \frac{C_{\lg}}{\lg \left[\frac{x_2}{x_1} \right]} \quad (\text{B.36})$$

Der Zähler in Gleichung B.36 ist aus Tabelle B.1 zu entnehmen.

Tabelle B.1: Umrechnungsfaktoren C nach Gleichung B.36 für den Exponenten k der WEIBULL-Verteilung

u	$\Phi(u)$	C_{\ln}	C_{\lg}
0,500	0,691	0,503	1,159
1,000	0,841	1,028	2,366
1,282	0,900	1,340	3,086
1,645	0,950	1,767	4,068
1,960	0,975	2,164	4,982

B.3 Zu erwartende Stützzahlen für ein ideal treffsicheres Stützzahlkonzept

Für den Bewertung der Treffsicherheit verschiedener Stützzahlkonzepte werden die nach dem jeweiligen Konzept abgeschätzten Stützzahlen n_{th} mit den experimentell bestimmten Stützzahlen n_{exp} verglichen. Hierfür wird an einem möglichst großen Datensatz die Verhältnisse aus experimentell ermittelten zu abgeschätzten Stützzahlen gebildet. Diese Verhältnisse sind identisch zu den Verhältnissen zwischen den experimentell ermittelten Festigkeiten Σ_{exp} und den nach dem jeweiligen Konzept abgeschätzten Festigkeiten Σ_{th} .¹

$$n_{rel} = \frac{n_{exp}}{n_{th}} = \frac{\frac{\Sigma_{exp}}{\Sigma_{ref}}}{\frac{\Sigma_{th}}{\Sigma_{ref}}} = \frac{\Sigma_{exp}}{\Sigma_{th}} \quad (B.37)$$

Die gebildeten relativen Stützzahlen werden statistisch ausgewertet. Dabei wird n_{rel} als logarithmischen normalverteilte Zufallsgröße aufgefasst, deren Logarithmus $\lg(n_{rel}) = n_{rel,lg}$ mit dem Erwartungswert $\mu_{lg,n_{rel}}$ und der Varianz $\sigma_{lg,n_{rel}}^2$ normalverteilt ist. Die Auswertung erfolgt meist im Wahrscheinlichkeitsnetz, kann aber genauso nach anderen Schätzmethoden, z. B. der Momentenmethode erfolgen. Es werden die Schätzer $m_{lg,n_{rel}}$ und $s_{lg,n_{rel}}$ bzw. $s_{lg,n_{rel}}^2$ für die wahren Parameter $\mu_{lg,n_{rel}}$ und $\sigma_{lg,n_{rel}}$ bzw. $\sigma_{lg,n_{rel}}^2$ ermittelt, aus denen anschließend Median $n_{rel,50\%}$ und Streuspanne $T_{n_{rel}}$

$$n_{rel,50\%} = 10^{(m_{lg,n_{rel}})} \quad T_{n_{rel}} = 10^{(2 \cdot 1,2816 s_{lg,n_{rel}})} \quad (B.38)$$

der Stützzahlen eines jeden Konzeptes bestimmt werden. Der Vergleich der Treffsicherheit der verschiedenen Stützzahlkonzepte kann hierbei immer nur relativ, also im Bezug auf andere am selben Datensatz überprüften Konzepte erfolgen, da „gute“ oder „schlechte“ Werte unbekannt sind.

Eine Möglichkeit die ermittelten Streuspannen besser einzuschätzen kann durch die Annahme eines ideal treffsicheren Stützzahlkonzeptes erreicht werden. Trifft jede theoretisch geschätzt Schwingfestigkeit $\Sigma_{th,i}$ die tatsächliche mittlere Schwingfestigkeit der logarithmisch normalverteilten Schwingfestigkeit, ist also $\lg(\Sigma_{th,i})$ immer genau der wahre Wert des Verteilungsparameters $\mu_{lg,\Sigma,i}$ und ist außerdem der Logarithmus der experimentell bestimmten mittleren Schwingfestigkeit $\lg(\Sigma_{exp,i})$ eine erwartungstreue Schätzung für $\mu_{lg,\Sigma,i}$, dann gilt:

$$\mathbf{E}[n_{rel,lg,i}] = \mathbf{E}[\lg \Sigma_{exp,i} - \lg \Sigma_{th,i}] \quad (B.39)$$

$$\mu_{lg,n_{rel,i}} = \mu_{lg,\Sigma,i} - \mu_{lg,\Sigma,i} = 0 \quad (B.40)$$

Die Streuung jeder einzelnen Stützzahl entsteht nur aus der Streuung, der jeweilige experimentell bestimmte mittlere Schwingfestigkeit.

$$\mathbf{V}[n_{rel,lg,i}] = \mathbf{V}[\lg \Sigma_{exp,i} - \lg \Sigma_{th,i}] \quad (B.41)$$

$$\sigma_{lg,n_{rel,i}}^2 = \sigma_{lg,\Sigma_{exp,i}}^2 \quad (B.42)$$

¹Das Zeichen Σ wird verwendet, da das Symbol σ bzw. σ^2 für die Standardabweichung bzw. Varianz der Normalverteilung verwendet wird. Σ_{exp} bezeichnet die experimentell bestimmte und Σ_{th} die theoretisch abgeschätzte Festigkeit. Diese werden für die Bildung der Stützzahl auf eine als bekannt vorausgesetzte Referenz Σ_{ref} bezogen

Diese ist, wie im Abschnitt 4.2 gezeigt, von der Art und Weise der Versuchsführung abhängig und proportional zur tatsächlichen Streuung der Schwingfestigkeit $\sigma_{\lg,\Sigma,i}$ der untersuchten Proben bzw. Bauteile.

$$\sigma_{\lg,\Sigma_{\text{exp}},i} = C_{m,i} \cdot \sigma_{\lg,\Sigma,i} \quad (\text{B.43})$$

Werden bei der Untersuchung alle Schwingfestigkeiten in gleicher Art und Weise ermittelt, kann ein konstantes C_m angenommen werden. Wird weiterhin unterstellt, dass die Standardabweichung der tatsächliche Schwingfestigkeit aller untersuchten Proben in einem festen Verhältnis zum tatsächlichen Erwartungswert der Schwingfestigkeit steht, also alle untersuchten Proben eine gemeinsame Streuspanne bzw. logarithmische Standardabweichung aufweisen, dann gilt:

$$\sigma_{\lg,\Sigma_{\text{exp}},i} = \sigma_{\lg,\Sigma_{\text{exp}}} = C_m \cdot \sigma_{\lg,\Sigma} \quad (\text{B.44})$$

Damit kann jede Stützzahl $\lg(n_{\text{rel},i})$ als eine identisch normalverteilte Zufallsgröße mit dem Erwartungswert

$$\mathbf{E}[n_{\text{rel},\lg,i}] = \mu_{\lg,n_{\text{rel}}} = 0 \quad (\text{B.45})$$

und der Varianz

$$\mathbf{V}[n_{\text{rel},\lg,i}] = \sigma_{\lg,n_{\text{rel}}}^2 = C_m^2 \cdot \sigma_{\lg,\Sigma}^2 \quad (\text{B.46})$$

betrachtet werden. Beim Vergleich verschiedener Stützzahlkonzepte wird davon eine Stichprobe aus k Elementen gezogen, aus der mit den Größen $m_{\lg,n_{\text{rel}}}$ und $s_{\lg,n_{\text{rel}}}$ die Verteilungsparameter $\mu_{\lg,n_{\text{rel}}}$ und $\sigma_{\lg,n_{\text{rel}}}$ geschätzt werden. Wie treffsicher diese Schätzungen sind, hängt von der Schätzmethode und der Größe der Stichprobe ab. Der Schätzer $m_{\lg,n_{\text{rel}}}$ liefert sowohl bei der Auswertung nach Momentenmethode als auch bei der Auswertung im Wahrscheinlichkeitsnetz erwartungstreue Werte, d. h.:

$$\overline{m_{\lg,n_{\text{rel}}}} = \mathbf{E}[m_{\lg,n_{\text{rel}}}] = \mu_{\lg,n_{\text{rel}}} = 0 \quad (\text{B.47})$$

Der Standardfehler des Schätzers $m_{\lg,n_{\text{rel}}}$, also die Schwankung des Wertes von $m_{\lg,n_{\text{rel}}}$ von Stichprobe zu Stichprobe, lässt sich für beide Auswertemethoden mit

$$s_{m_{\lg,n_{\text{rel}}}} = \sqrt{\mathbf{V}[m_{\lg,n_{\text{rel}}}] = \frac{\sigma_{\lg,n_{\text{rel}}}}{\sqrt{k}} = \sqrt{\frac{C_m^2 \cdot \sigma_{\lg,\Sigma}^2}{k}} \quad (\text{B.48})$$

angeben.

Für den Schätzer $s_{\lg,n_{\text{rel}}}$ unterscheiden sich die Ergebnisse je nach Auswertemethode. Wird die Standardabweichung von $n_{\text{rel},\lg}$ mit der Momentenmethode direkt geschätzt², ist der Erwartungswert des Schätzers³

$$\overline{s_{\lg,n_{\text{rel}}}} = \mathbf{E}[s_{\lg,n_{\text{rel}}}] = \sqrt{\frac{2}{k-1} \frac{\Gamma(\frac{k}{2})}{\Gamma(\frac{k-1}{2})}} \sigma_{\lg,n_{\text{rel}}} \quad (\text{B.49})$$

und der Standardfehler wir mit

$$s_{s_{\lg,n_{\text{rel}}}} = \sqrt{\mathbf{V}[s_{\lg,n_{\text{rel}}}] = \sqrt{1 - \frac{2}{k-1} \frac{\Gamma^2(\frac{k}{2})}{\Gamma^2(\frac{k-1}{2})}} \sigma_{\lg,n_{\text{rel}}} \quad (\text{B.50})$$

²Also die nicht erwartungstreue Schätzung der Standardabweichung aus der erwartungstreuen (korrigierten) Stichprobenvarianz

³[42], S. 66. Das in der Formel verwendete Symbol $\Gamma(\cdot)$ bezeichnet die Gammafunktion und $\Gamma^2(\cdot)$ deren Quadrat.

angegeben. Bei einer Auswertung im Wahrscheinlichkeitsnetz, wie in Abschnitt 2.1.3.3 beschrieben, ist die Schätzung der Standardabweichung nahezu erwartungstreu, d.h.

$$\overline{s_{\lg, n_{\text{rel}}}} = \mathbf{E}[s_{\lg, n_{\text{rel}}}] = \sigma_{\lg, n_{\text{rel}}} \quad (\text{B.51})$$

Der Standardfehler fällt allerdings größer als bei der Schätzung mit der Momentenmethode aus. Er kann näherungsweise mit

$$s_{s_{\lg, n_{\text{rel}}}} = \sqrt{\mathbf{V}[s_{\lg, n_{\text{rel}}}]} = \sqrt{\left(\frac{k}{k-1}\right)^{0,7} \left(1 - \frac{2}{k-1} \frac{\Gamma^2\left(\frac{k}{2}\right)}{\Gamma^2\left(\frac{k-1}{2}\right)}\right) \sigma_{\lg, n_{\text{rel}}}} \quad (\text{B.52})$$

angegeben werden.

Für eine beliebige Stichprobe mit k verschiedenen relativen Stützzahlen würde man bei einer Auswertung im Wahrscheinlichkeitsnetz eine mittlere Stützzahl \tilde{n}_{rel} von

$$\tilde{n}_{\text{rel}} = 10^{\mu_{\lg, n_{\text{rel}}} \pm (C_m \cdot \sigma_{\lg, \Sigma}) / \sqrt{k}} = 1 \cdot 10^{\pm (C_m \cdot \sigma_{\lg, \Sigma}) / \sqrt{k}} \quad (\text{B.53})$$

erwarten. Die Streuung der ermittelten Stützzahlen kann durch die zu erwartende Streuspanne $T_{n_{\text{rel}}} = 10^{2 \cdot 1,2816 \cdot s_{\lg, n_{\text{rel}}}}$ abgebildet werden, wie sie in Tabelle B.2 für verschiedene C_m und $\sigma_{\lg, \Sigma}$ angegeben ist. Die Unsicherheit $u_{T_{n_{\text{rel}}}}^2$ dieser ermittelten Streuspanne kann durch Linearisierung aus der Unsicherheit bzw. dem Standardfehler der Streuung bestimmt werden:

$$u_{T_{n_{\text{rel}}}}^2 = \left(\frac{\partial T_{n_{\text{rel}}}}{\partial s_{\lg, n_{\text{rel}}}}\right)^2 s_{s_{\lg, n_{\text{rel}}}}^2 \quad (\text{B.54})$$

In Tabelle B.3 sind die zu erwartenden Unsicherheiten für verschieden ermittelte Streuspannen bei verschiedenen Stichprobenumfängen angegeben. Die Frage wie sicher die Schätzung der Streuspanne $T_{n_{\text{rel}}}$ ist, kann zumindest, wenn wie oben beschrieben alle ermittelten Stützzahlen identische Verteilungsparameter besitzen, recht einfach beantwortet werden.

B.3 Zu erwartende Stützzahlen für ein ideal treffsicheres Stützzahlkonzept

Tabelle B.2: Zu erwartende Streuspannen $T_{n_{\text{rel}}}$ eines ideal treffsicheren Stützzahlkonzeptes.

		Streuspanne der Schwingfestigkeit T_{Σ}								
		1,005	1,010	1,025	1,050	1,100	1,125	1,250	1,500	2,000
		log. Standardabweichung der Schwingfestigkeit $\sigma_{\lg, \Sigma}$								
		0,00085	0,00169	0,00418	0,00827	0,01615	0,01996	0,03781	0,06870	0,11744
C_m		$T_{n_{\text{rel}}}$								
0,01		1,0000	1,0001	1,0002	1,0005	1,0010	1,0012	1,0022	1,0041	1,0070
0,1		1,0005	1,0010	1,0025	1,0049	1,0096	1,0118	1,0226	1,0414	1,0718
0,2		1,0010	1,0020	1,0050	1,0098	1,0192	1,0238	1,0456	1,0845	1,1487
0,3		1,0015	1,0030	1,0074	1,0147	1,0290	1,0360	1,0692	1,1293	1,2311
0,4		1,0020	1,0040	1,0099	1,0197	1,0389	1,0482	1,0934	1,1761	1,3195
0,5		1,0025	1,0050	1,0124	1,0247	1,0488	1,0607	1,1180	1,2247	1,4142
0,6		1,0030	1,0060	1,0149	1,0297	1,0589	1,0732	1,1433	1,2754	1,5157
0,7		1,0035	1,0070	1,0174	1,0347	1,0690	1,0859	1,1691	1,3282	1,6245
0,8		1,0040	1,0080	1,0200	1,0398	1,0792	1,0988	1,1954	1,3832	1,7411
0,9		1,0045	1,0090	1,0225	1,0449	1,0896	1,1118	1,2224	1,4404	1,8661
1,0		1,0050	1,0100	1,0250	1,0500	1,1000	1,1250	1,2500	1,5000	2,0000

Tabelle B.3: Zu erwartende Unsicherheiten $u_{T_{n_{\text{rel}}}}$ der Streuspannen $T_{n_{\text{rel}}}$ eines ideal treffsicheren Stützzahlkonzeptes, für verschiedene Stichprobengrößen.

		Streuspanne der Stützzahlen $T_{n_{\text{rel}}}$								
k		1,005	1,010	1,025	1,050	1,100	1,125	1,250	1,500	2,000
3		0,0027	0,0054	0,0135	0,0274	0,0560	0,0707	0,1489	0,3247	0,7401
5		0,0018	0,0037	0,0093	0,0189	0,0387	0,0489	0,1029	0,2244	0,5114
10		0,0012	0,0024	0,0061	0,0123	0,0253	0,0319	0,0672	0,1466	0,3340
15		0,0010	0,0019	0,0049	0,0098	0,0201	0,0254	0,0535	0,1167	0,2659
20		0,0008	0,0016	0,0042	0,0084	0,0172	0,0217	0,0458	0,0998	0,2274
30		0,0007	0,0013	0,0033	0,0068	0,0139	0,0175	0,0369	0,0805	0,1834
40		0,0006	0,0011	0,0029	0,0058	0,0119	0,0151	0,0318	0,0693	0,1578
50		0,0005	0,0010	0,0026	0,0052	0,0106	0,0134	0,0283	0,0617	0,1407
70		0,0004	0,0009	0,0022	0,0044	0,0090	0,0113	0,0238	0,0519	0,1184
100		0,0004	0,0007	0,0018	0,0036	0,0075	0,0094	0,0199	0,0433	0,0987
150		0,0003	0,0006	0,0015	0,0030	0,0061	0,0077	0,0162	0,0353	0,0804