Ein nichtlineares, hierarchisches und gemischtes Modell für das Baum-Höhenwachstum der Fichte (*Picea abies (L.) Karst.*) in Baden-Württemberg

Dissertation

zur Erlangung des Doktorgrades

des Forstwissenschaftlichen Fachbereichs

der Georg-August-Universität Göttingen

vorgelegt von

Arne Nothdurft

aus Nürnberg

Göttingen 2007

1. Berichterstatter: Prof. Dr. J. Saborowski

2. Berichterstatter: Prof. Dr. J. Nagel

3. Berichterstatter: Prof. Dr. H. Pretzsch

Tag der mündlichen Prüfung: 9. Februar 2007

Danksagung

Mein Dank gilt Herrn Prof. Dr. Joachim Saborowski für die Unterstützung dieser Arbeit und die wertvollen Diskussionen. Herrn Prof. Dr. Jürgen Nagel möchte ich für die Übernahme des Korreferats danken. Herrn Dr. Edgar Kublin verdanke ich eine intensive Betreuung vor Ort und die gemeinsame Entwicklung der Fragestellung. Von Herrn Ph.D. Juha Lappi erhielt ich während seines Forschungsaufenthalts an der FVA die entscheidenden Impulse. Herrn Dr. Gerald Kändler danke ich für die fruchtbaren Diskussionen und das Korrekturlesen. Herrn Prof. Dr. W. Schöpfer und Prof. Dr. J. Hradetzky gelang es, die historischen Daten aus Prof. Mitscherlichs Studie aufzuspüren, die von Herr N. Martinez gewissenhaft digitalisiert wurden.

Die Landesforstverwaltung Baden-Württemberg ermöglichte diese Dissertation durch meine Beschäftigung in der Abteilung für Biometrie und Informatik an der FVA Freiburg. Die gesamte Abteilung Biometrie und Informatik trug mit hervorragendem Arbeitsklima und technischer Unterstützung zum Gelingen dieser Arbeit bei.

Meiner Freundin Karolina Milewicz danke ich für die gemeinsame schöne Zeit und ihre zuverlässige Hilfe in allen Lebenslagen. Meiner Mutter Karin und meinem Vater Jürgen danke ich für ihre Unterstützung und für die Aufrechterhaltung meines Heimatgefühls in Buntenbock (Oberharz) und meinem Bruder Jörg für alle gemeinsamen, sportlichen und spaßigen Aktivitäten.

Zum nötigen Ausgleich in der Freizeit trug das Rennrad-Training mit Jens Kerll, Johannes Mühlenbruch, sowie den Oberrieder Bike-Bären bei. Bernd Schmidt, Thorsten Münch und Günther Moritz möchte ich für das jahrelange Training beim SC Buntenbock danken. Der Ski-Marathon-Mannschaft des SV Kirchzarten verdanke ich eindrucksvolle Ski-Ausflüge.

Inhaltsverzeichnis

1 Einleitung **1**

1.1 Zielsetzung 1

1.2 Gegenwärtiger Kenntnisstand 4

1.3 Hypothesen 9

2 Datengrundlage **13**

2.1 Daten für die Modellbildung 13

2.1.1 Immissionsökologische Waldzustandserhebung (IWE) 13

2.1.2 Mitscherlichs Erhebung 14

2.1.3 Modellierungsdatensatz 14

2.2 Daten für die Modellevaluierung 15

3 Methode **17**

3.1 Das nichtlineare Regressionsmodell 17

3.1.1 Begründung der Nichtlinearität 17

3.1.2 Die nichtlineare Grundfunktion 18

3.2 Theorie hierarchischer und gemischter Modelle 20

3.2.1 Lineare gemischte Modelle zur Einführung 20

3.2.2 Das nichtlineare, hierarchische und gemischte Modell 21

3.2.3 Parameterschätzung 26

3.2.4 Vorhersage der Zufallsparameter 34

4 Ergebnisse **39**

4.1 Modellkonstruktion 39

4.1.1 Einstufiges Zufallsmodell 39

4.1.2 Zweistufiges Zufallsmodell 48

4.1.3 Zweistufiges gemischtes Modell 54

4.1.3.1 Überlegungen zur Beschreibung räumlicher Variation . . . 54

4.1.3.2 Überlegungen zur Beschreibung zeitlicher Variation 62

4.1.3.3 Beschreibung zeitlicher Variation mit dem Keimjahr . . . 72

4.2 Modellevaluierung 82

5 Diskussion **115**

5.1 Methodischer Erkenntnisgewinn 115

5.1.1 Hierarchische, gemischte Modelle vs. Traditionelle Site-Index Modelle 115

5.1.2 Würdigung des methodischen Erkenntnisgewinns 126

5.1.3 Einsatzgebiet des hierarchischen, nichtlinearen und gemischten Modells 129

5.2 Waldwachstums- und ertragskundlicher Erkenntnisgewinn 132

5.3 Würdigung der zugrunde liegenden Daten 137

6 Schlussfolgerung **139**

7 Zusammenfassung **143**

8 Anhang **i**

Abbildungsverzeichnis

2.1 Lage der Probeflächen 16

4.1 Messwerte der Baumhöhen 44
4.2 Vorhersagen der Baumhöhen aus dem einstufigen Zufallsmodell 45
4.3 Residuen des einstufigen Zufallsmodells 46
4.4 Unplausible Kurvenverläufe des einstufigen Zufallsmodells 47
4.5 Gruppierungsmerkmal für die Bäume der IWE 49
4.6 Stichproben-Mittelwertkurven 51
4.7 Validität des zweistufigen Zufallsmodell 52
4.8 Variation zwischen Bäumen 55
4.9 Verteilungen für die Vorhersagewerte der Zufallsparameter $z_i^{(2)}$ und $z_i^{(4)}$ nach Regionen - Zufallsmodell 58
4.10 Verteilungen für die Vorhersagewerte der Zufallsparameter $z_i^{(2)}$ und $z_i^{(4)}$ nach Regionen - gemischtes Modell 61
4.11 Mittelwertkurven für Regionen 62
4.12 Verteilung der Keimjahreswerte 65
4.13 Schätzungen für den Höhenerwartungswert-Parameter 66
4.14 Überprüfung der Annahmen für die Zufallseffekte 68
4.15 Mittelwertkurven nach Regionen und Datenherkunft 69
4.16 Mittelwertkurven und Beobachtungen für die Merkmalsgruppen 73
4.17 Stichproben-Mittelwertkurven nach Merkmalsgruppen 74
4.18 Zeitliche Variation 76

4.19 Zufallseffekt auf Baumebene 77
4.20 Zufallsparameter auf Baumebene bei fixem Effekt 80
4.21 Baumhöhenmessungen an einem Stichprobenpunkt 95
4.22 Mittelwertkurven und erste Beobachtungen zur Kalibrierung 95
4.23 Individuelle Kurven mit vorhergesagten Zufallsparametern 95
4.24 Individuelle Kurven und alle Beobachtungen 95
4.25 Kalibrierung anhand von zwei Vorbeobachtungen 96
4.26 Beobachtungen und Vorhersagen der Kalibrierungsdaten 96
4.27 Betrachtung einzelner Stichprobenpunkte hinsichtlich plausibler Vorhersagen 97
4.28 Ausgleich empirischer Richtungsfelder 101
4.29 Vorhersagen durch Bonitätsfächer mit der ersten Beobachtung als Startinformation 102
4.30 Vorhersagen durch Bonitätsfächer mit der elften Beobachtung als Startinformation 103
4.31 Residuen der Baumhöhen-Prognosewerte mit der ersten Beobachtung als Startinformation 104
4.32 Residuen der Baumhöhen-Prognosewerte mit der elften Beobachtung als Startinformation 105
4.33 Vergleiche der Verteilungsmaße für die Residuen 109
4.34 Genauigkeit von Baumhöhenvorhersagen 110

Tabellenverzeichnis

4.1 Likelihood-Ratio-Test 46
4.2 Parameterwerte aus OLS-Schätzung für den fixen Effekt auf Baum-Ebene. 79
4.3 Höhenlage für die Stichproben nach Regionen 80
4.4 Parameterschätzwerte aus der Anpassung des empirischen Richtungsfelds . 101
4.5 Geschätzte Vorhersageungenauigkeit 111

8.1 Parameter-Schätzwerte aus den Modellanpassungen ii
8.2 Geschätzte Kovarianzen der fixen Parameter des endgültigen Modells . . . iii
8.3 Beschreibung der Datengrundlage iv
8.4 Beschreibung der Datengrundlage nach Altersklassen v
8.5 Verteilungen der Residuen über die Prognosezeiträume bei Kenntnis der ersten Beobachtung - Teil 1 vi
8.6 Verteilungen der Residuen über die Prognosezeiträume bei Kenntnis der ersten Beobachtung - Teil 2 vii
8.7 Verteilungen der Residuen über die Prognosezeiträume bei Kenntnis der elften Beobachtung für die Fortschreibung mittels der Bonitätsfächer und der ersten und elften Beobachtungen für BLUPs der Zufallsparameter . . . viii

Kapitel 1

Einleitung

1.1 Zielsetzung

Die Volumen des Holzvorrats und dessen Zuwachs haben große Bedeutung für forstbetriebliche und wissenschaftliche Fragestellungen. Die Kenntnis des Holzzuwachses der Waldbäume dient der Planung einer nachhaltigen Waldbewirtschaftung (Ministerkonferenz zum Schutz der Wälder in Europa, 1993). Zur Anschauung und für das Verständnis muss die Wirklichkeit mit möglichst genauen und plausiblen Modellen abstrahiert werden. Die Disziplin der forstlichen Biometrie nimmt sich unter anderen der Aufgabe an, Modelle zur Schätzung des Zuwachses zu formulieren. Diese forstlichen Wuchsmodelle sollten einer ständigen Evaluierung unterliegen (Pretzsch, 2001). Bewährte Modelle müssen erneut auf Validität geprüft werden, wenn Anzeichen für Veränderungen der Wirklichkeit gegeben sind oder eine verbesserte Informationsbasis aufgrund reichhaltigerer Messdaten entsteht. Technische Neuerungen für die Erfassung von Daten oder aktuelle und genauere statistische Regressionsverfahren können zugleich Anlass und Beitrag zur Verbesserung von Systemmodellen sein.

Weil das Volumen der stehenden Bäume nur schwer messbar ist, wird es meist mittels Regressionen über einfach zu messende Hilfsgrößen hergeleitet. Der Durchmesser an einer beliebigen Stelle des Baumschaftes wird in Abhängigkeit von der Baumhöhe H, dem Durchmesser in Brusthöhe Bhd und ggf. noch von einem weiteren Durchmesser außerhalb

des Wurzelanlaufes beschrieben, die der Baum zum Zeitpunkt t hat. In der von Kublin (2003) entwickelten Programmbibliothek *BDat* wird das Holzvolumen durch Integration der Rotationskörper von abschnittsweise definierten Funktionen mit $V_t = f(t\,|Bhd, H)$ geschätzt. Der Zuwachs eines Baumes ergibt sich aus der Veränderung seines Holzvolumens zwischen zwei Zeitpunkten $t2$ und $t1$ mit $Z_{t2,t1} = V_{t2} - V_{t1}$. Unter der Annahme einer über die Zeit konstanten Schaftform kann der Volumenzuwachs eines Baumes über den Zuwachs an Höhe und Durchmesser hergeleitet werden. Wenn für eine Waldinventur nach einer Zeitperiode eine Folgeaufnahme vorgenommen wird, liegen für die wiederholt aufgenommenen Bäume grundsätzlich jeweils zwei Beobachtungen für die Merkmale Durchmesser und Höhe vor. Der geleistete Volumenzuwachs der vergangenen Zeitperiode kann für diese Bäume über deren Vorratsveränderung bilanziert werden. Es gibt jedoch auch Gründe, weshalb der Zuwachs von einzelnen Bäumen über eine Periode nicht über direkte Messungen geschätzt werden kann, sondern modelliert werden muss. Zum einen können für eine rückwärts gerichtete Bilanzierung Messungen fehlen. Zum anderen ist oft ein Blick in die Zukunft für Prognosen von Interesse, der durch Fortschreibungen unterstützt werden kann.

Fehlende Messungen müssen geschätzt werden, wenn für die Bäume keine beobachteten Hilfsgrößen zum Periodenanfang oder -ende vorliegen. Bei Waldinventuren mit konzentrischen oder variablen Probekreisdurchmessern ist die Auswahlwahrscheinlichkeit eines Baumes meist vom Durchmesser in Brusthöhe und seiner Entfernung zum Probekreismittelpunkt abhängig. Deswegen können trotz Wiederholungsaufnahmen für einzelne Bäume, nur zum späteren Zeitpunkt Messungen vorhanden sein. Es handelt sich dabei um Bäume des sogenannten Einwuchses. Deren Brusthöhendurchmesser hatten bei der Aufnahme zum ersten Zeitpunkt noch nicht die Kluppschwelle überschritten und mussten nicht erfasst werden. Wenn der Baum im Duchmesser wächst, kann der Bhd bei der Folgeaufnahme über der Kluppschwelle liegen; eine Aufnahme wird dann notwendig. Außerdem kann es Bäume geben, die zwar bei der früheren Aufnahme gemessen wurden, aber zwischenzeitlich vom Probekreis entfernt wurden; für diese können keine Messungen zum Stichtag der Folgeaufnahme vorliegen. Der Holzzuwachs, den die Probekreisfläche

repräsentiert setzt sich zusammen aus dem Zuwachs der überlebenden Bäume, dem des Einwuchses und dem der entnommenen oder abgestorbenen Bäume (Hradetzky, 2002). Der Zuwachs des Einwuchses und der abgegangenen Bäume kann nicht bilanziert werden, sondern muss modelliert werden. Außerdem muss der Zuwachs für eine vergangene Periode für alle gemessenen Bäume auf einmalig erhobenen Stichprobenflächen modelliert werden, sofern er von Interesse ist. Auf zwischenzeitlich angekauften Holzbodenflächen entstehen oft einmalig erhobene Stichproben, auch wenn der übrige Teil des Forstbetriebes zum wiederholten mal inventarisiert wurde. Sofern für den gesamten Betrieb keine Folgeaufnahmen vorliegen, die betreffenden Flächen also bislang noch nicht inventarisiert wurden oder bei Inventurverfahren, bei denen die Lage des Probekreismittelpunktes und der einzelnen Bäume nicht erfasst werden, müssen die Zuwächse aller Bäume modelliert werden.

Für die Planung im Zuge der Forsteinrichtung dient nach gültiger Dienstanweisung für die Forsteinrichtung im öffentlichen Wald Baden-Württembergs (Landesforstverwaltung Baden-Württemberg, 2002) bei Betrieben mit wiederholter Aufnahme eines permanenten Stichprobennetzes der Zuwachs der abgelaufenen Periode als Gesamtweiser für die Festsetzung des Hiebssatzes. Eine Modellschätzung des Zuwachses für die zukünftige Periode könnte die Planungssicherheit erhöhen. Deshalb haben sich in der näheren Vergangenheit die Waldwachstumssimulatoren BWINPro (Nagel et al., 2002) in Niedersachsen und SILVA (Pretzsch et al., 2002) in Bayern zur Entscheidungsunterstützung bei forstlichen Planungen etabliert; sie ermöglichen Fortschreibungen von Inventurdaten aus Rasterstichproben.

Das Ziel dieser Studie ist die Erstellung eines stochastischen Modells für das Baumhöhenwachstum in Baden-Württemberg am Beispiel der Baumart Fichte. In der Zukunft soll auf diese Arbeit weiter aufbauend ein Ansatz zur simultanen Modellierung von Höhen- und Durchmesserwachstum entwickelt werden. Dann wären Fortschreibungsmodelle für die beiden Attribute Bhd und Höhe zur Bestimmung des Volumens vorhanden. Über Bilanzierung der Volumina zu zwei verschiedenen Zeitpunkten ergäbe sich der Zuwachs an Holzvolumen.

Diese Arbeit entspringt wichtigen Beweggründen. Aus Waldinventuren sind aktuelle Daten aus Zeitreihenbeobachtungen einzelner Bäume verfügbar, die bisher nicht zur Wachstumsmodellierung verwendet wurden. Bei diesen Daten handelt es sich um Wiederholungsmessungen an gefällten Bäumen.

Bisher sind im deutschsprachigen Raum keine Wachstumsmodelle erstellt worden, bei denen die für longitudinale Daten charakteristische Abhängigkeit der Beobachtungen eines Individuums Berücksichtigung gefunden hätte. Erst die Entwicklung neuer statistischer Regressionsverfahren zur Analyse von Daten aus Wiederholungsmessungen und die Verfügbarkeit der Methoden in statistischer Standardsoftware ermöglichen eine Modellbildung, wie sie in dieser Studie aufgezeigt wird.

1.2 Gegenwärtiger Kenntnisstand

Ein mathematisches Modell als deterministischer Ansatz schätzt die Ziel-Variable (Response) in Abhängigkeit von Regressoren mit absoluter Sicherheit ohne Reststreuung. Nach Schabenberger und Pierce (2002, S. 2) ist ein derartiges Modell nur geeignet, wenn die abhängige Variable ohne Fehler gemessen werden kann und die Beziehungen zwischen dem Baumwachstum und allen Einflussgrößen konstant und bekannt sind. Da alle für das Baumwachstum verantwortlichen Einflussgrößen weder bekannt sind, noch gemessen werden können und die Beobachtungen der Response mit Fehlern behaftet sind, bietet sich deshalb ein statistisches Modell an.

Bei forstwissenschaftlichen Experimenten werden oft Stichproben gezogen, weil die Waldbestände als Untersuchungseinheiten wegen ihrer großen Anzahl von Bäumen nur schwer in ihrer Gesamtheit zu messen sind. In den Zufallsstichproben von Waldbeständen wird innerhalb der Waldbestände häufig eine bestimmte Zahl von Bäumen für Messungen ausgewählt. Die Waldbestände werden bei forstlichen Inventuren für Betriebe oder größere Räume meistens systematisch beprobt. Über das Untersuchungsgebiet werden Raster gelegt, an deren Knotenpunkten die zu erhebenden Stichproben installiert werden. An den Stichprobenpunkten werden die Merkmale entweder von einer bestimmten Anzahl von

nächststehenden Bäumen erhoben oder von den Bäumen, die eine bestimmte Entfernung zum Stichprobenpunkt nicht überschreiten.

Der Waldbestand oder der Stichprobenpunkt stellt im Sinne der Methoden, die in dieser Studie angewendet werden ein Merkmal zur Gruppierung dar. Die Bäume an einem Stichprobenpunkt sind die Elemente dieser Gruppe. Für den Begriff Gruppe wird der englischsprachige Begriff Cluster synonym verwendet. Die Bäume an einem Stichprobenpunkt stehen in den meisten Fällen im gleichen Waldbestand. Die Vermutung liegt nahe, dass die Bäume innerhalb eines Bestandes einen ähnlichen Verlauf ihres Wachstums zeigen und sich der Verlauf des Wachstums von Bäumen aus verschiedenen Beständen stärker unterscheidet. In der vorliegenden Arbeit soll diese Vermutung überprüft werden. Sollte sie sich bewahrheiten, kann jenes Phänomen der „Gruppenähnlichkeit“ bei der Modellkonstruktion berücksichtigt werden.

Werden am selben Individuum zu unterschiedlichen Zeitpunkten Messungen durchgeführt, dann sprechen wir von Wiederholungsmessungen. Nur wenn die verstrichene Zeit der Anlass für die wiederholte Beobachtung ist, spricht man streng genommen von longitudinalen Daten. Abschwächungen der strengen Terminologie finden sich bei Davidian (2005, S. 2). Beim Studium von Wachstumsprozessen, werden oft an den selben Individuen zu unterschiedlichen Zeitpunkten Messungen vorgenommen, die zu solchen longitudinalen Daten führen. Dagegen kann z.B. in pharmakokinetischen Studien die verabreichte Dosis eines Pharmazeutikums zu unterschiedlichen Zeitpunkten der Anlass zur Beobachtung sein; da dort nicht die verstrichene Zeit, sondern die unterschiedliche Dosis den Anlass zur erneuten Beobachtung gibt, spricht man streng genommen nur von Wiederholungsmessungen über die Dosis und nicht von longitudinalen Daten. Die in dieser Studie verwendeten Daten von Baumhöhenbeobachtungen stammen aus Trieblängenmessungen an gefällten Bäumen. Nach Ansicht von Davidian (2005) ist dabei nicht die Zeit, sondern die Position am Baumstamm der Anlass für die Messung. Da die Positionen der Messstellen am Ende von Höhentrieben liegen, die ein Baum innerhalb des Zeitraumes eines Jahres produziert hat, soll jedoch auch bei den vorliegenden Messdaten aus Höhentriebanalysen von longitudinalen Daten gesprochen werden.

Weil bei longitudinalen Daten mehrere Beobachtungen an ein und dem selben Individuum unternommen werden, ergibt sich mit dem Individuum selbst eine weitere Gruppierung der Daten. Die verschiedenen Gruppierungsmerkmale können hierarchisch strukturiert sein. Mehrere Messungen gehören zu einem Baum und mehrere Bäume gehören zu einem Stichprobenpunkt.

Ein Ziel der longitudinalen Datenanalyse, ist die Entwicklung eines Regressionsmodells, das die Variation der Response-Variable innerhalb der Gruppen und zwischen den Gruppen berücksichtigt, indem eine Mittelwertkurve für die gesamte Population und Mittelwertkurven für die einzelnen Cluster geschätzt werden (Schabenberger und Pierce, 2002, S. 408).

Die Variation der Parameter und damit auch die der Response-Variablen kann durch sogenannte fixe, zufällige und gemischte Effekte beschrieben werden. Bei einem Modell mit Zufallsparametern wird angenommen, dass die Werte für die Response um einen mittleren Wert streuen. Wenn die genaue Ursache der Abweichung vom Mittelwert nicht aufgeklärt werden kann, dann wird angenommen, dass sie von unbekannten zufälligen Effekten und dem Restfehler herrührt. Wenn es gelingt eine Einflussgröße als Ursache auszumachen, dann kann sie in das Modell als zusätzlicher Regressor eingefügt werden. Den Einfluss, den diese Kovariablen auf die Response ausüben nennt man fixe Effekte. Wird die Variation in einem Modell durch die Kombination von fixen und zufälligen Effekten modelliert, spricht man von gemischten Effekten in gemischten Modellen. Es wird dann unterstellt, dass nur ein Teil der Variation in Abhängigkeit von bekannten Regressoren beschrieben werden kann und die restliche Variation von unbekannten Einflüssen und einem stochastischen Restfehlerterm verursacht wird. Wir nehmen an, dass es sowohl für die gesamte Population, als auch für jede Gruppe über ihre Individuen hinweg einen Mittelwert bzw. bei den vorliegenden longitudinalen Baumhöhenbeobachtungen eine mittlere Wachstumsfunktion gibt. Wenn die Daten eine hierarchische Gruppierung aufweisen, kann es Zufallseffekte auf mehreren Ebenen geben. Die zufälligen Effekte auf einer Gruppierungsebene verursachen die nicht erklärbare Variation *zwischen* den Gruppen-Mittelwertkurven. Die

zufälligen Effekte auf der nächsten darunter liegenden Gruppierungsebene bewirken die nicht erklärbare Variation *um* die jeweiligen Gruppen-Mittelwertkurven *herum*.

In ähnlicher Weise gibt es fixe Effekte auf Ebene der Gruppen die den bekannten Teil der Variation zwischen den Gruppen beschreiben können und fixe Effekte auf der Gruppierungsebene darunter, die den bekannten Teil der Variation innerhalb einer Gruppe zwischen ihren Einheiten abbilden. Wenn gemischte Modelle die Variation wie beschrieben durch Effekte auf unterschiedlichen Ebenen erklären, nennen wir sie hierarchische und gemischte Modelle.

Die Wachstumsverläufe der Durchmesser und Höhen von Waldbäumen folgen meist Kurven sigmoider Form mit einem Wendepunkt und einer Asymptote. Die physiologische Begründung dazu findet sich bei Pretzsch (2001). Solche Kurvenverläufe können biologisch plausibel mit Hilfe nichtlinearer Funktionen beschrieben werden.

Von Lappi und Bailey (1988); Lappi (1991, 1997); Lappi und Malinen (1994); Mehtatalo (2004, 2005) wurden lineare gemischte Modelle für die Baumhöhe und das Volumen entwickelt, die durch Vorhersagen von Zufallsparametern kalibriert werden. Wenn die Grundfunktion eine nichtlineare Exponentialfunktion war, dann wurde sie durch Logarithmieren linearisiert. Seit Anfang der 90er Jahre gibt es forstwissenschaftliche Studien zu nichtlinearen, gemischten Modellen. Die Untersuchungsgegenstände und die Modellkonstruktionen der existierenden Arbeiten sollen nun dargelegt werden.

Lindstrom und Bates (1990) stellten in ihrer Pionierarbeit verschiedene Methoden zur Parameterschätzung vor. In einer Anwendung modellierten sie den Stammumfang von einzelnen Orangen-Bäumen über das Alter in Tagen mit einer logistischen Funktion. Gregoire und Schabenberger (1996a,b) beschrieben das verkaufsfähige Holzvolumen von Amberbäumen *(Liquidambar styraciflua L.)* in Abhängigkeit vom Zopfdurchmesser für einzelne Bäume. Dabei diskutierten sie ebenfalls verschiedene Parameterschätzmethoden.

Hall und Bailey (2001) bildeten ein hierarchisches gemischtes Modell für das Höhenwachstum von *Pinus taeda L.* über dem Alter. Die bei forstwissenschaftlichen Fragestellungen häufig verwendete Richards-Funktion diente als nichtlineares Grundmodell. Die Datenbasis waren longitudinale Beobachtungen der Baumhöhe für Alter von 3 bis 14 Jah-

ren an 365 Bäumen auf 24 Stichprobenflächen. Bei dem Modell wurden Zufallseffekte auf 2 Hierarchie-Ebenen implementiert.

Fang und Bailey (2001) entwickelten ein nichtlineares Modell für die Bestandesoberhöhe von *Pinus elliottii Egelm.* über den Altersrahmen von 5 bis 17 Jahren basierend auf der Richards-Funktion. Sie diskutierten verschiedene Modelle für den Restfehlerterm bei autokorrelierten Residuen. Einen Schwerpunkt ihrer Arbeit bildeten Anwendungen zur Fortschreibung bei unterschiedlicher Informationsbasis. Fang et al. (2001) erstellten anhand der selben Datenbasis ein multivariates simultanes System, das aus einem nichtlinearen gemischten Modell für das Oberhöhenwachstum und aus log-linearen Modellen für die Grundfläche und das Volumen über dem Alter bestand.

Calama und Montero (2004) formulierten ein nichtlineares, hierarchisches und gemischtes Modell für die individuelle Baumhöhe über den Durchmesser von *Pinus pinea L.* in Spanien. Ihre Studie stützte sich auf 8614 Bäume von 455 Plots in vier Regionen. Die Daten von 45 Plots wurden zur Validierung von der Modellanpassung ausgespart. Die nicht erklärbare Variation auf Ebene der Plots (zwischen den Plots) wurde mit zufälligen Effekten abgebildet. Die Herkunft des einzelnen Plots bezüglich der Region wurde als fixer Effekt begriffen und in den Daten mit Indikatorvariablen kodiert. Die nichterklärbare Variation auf Baumebene wurde durch einen Restfehlerterm modelliert.

Hall und Clutter (2004) verfassten ein trivariates, hierarchisches Modell für die Oberhöhe, die Grundfläche und die Stammzahl von *Pinus elliottii Egelm.*-Beständen. Das verkaufsfähige Holzvolumen konnte in Abhängigkeit einer multiplikativen Verkettung der drei Responsevariablen mit jeweils bestimmter Potenz geschätzt werden. Für die Beschreibung des Höhen- und des Grundflächenwachstums über dem Alter diente die Richards-Funktion als Basis und für die Stammzahl pro Hektar über dem Alter eine modifizierte negative Exponentialfunktion. Zufallseffekte wurden auf der Ebene der Bestände und der Ebene der Plots zugelassen. Die individuelle Variation auf der Ebene der einzelnen Bäume wurde durch das Restfehlermodell ausgedrückt. Als Vorteil dieses simultanen Ansatzes wurde herausgestellt, dass man sowohl Korrelationen zwischen den Restfehlern der drei

Modellkompartimente, als auch zwischen den Zufallseffekten für die Parameter auf der Bestandes- und Plotebene nutzen konnte.

Calegario et al. (2005) beschrieben mit einem nichtlinearen, gemischten Modell auf Basis einer logistischen Grundfunktion das Wachstum der Oberhöhe über dem Alter von Eukalyptus-Hybriden. Die klonale Herkunft auf 115 Plots wurde als fixer Effekt implementiert.

Zhao et al. (2005) stellten die Gesamtwuchsleistung von *Pinus taeda L.* über dem Alter mit Hilfe eines hierarchischen, nichtlinearen Modells auf Basis der Richards-Funktion dar. Fixe Effekte waren drei unterschiedliche Kombinationen der Behandlung aus Düngung und Einflußnahme auf die Begleitvegetation, sowie eine Kontrollvariante ohne Behandlung. Zufällige Effekte wurden auf den drei Hierarchieebenen Standort, Block und Plot untergebracht. Außerdem wurde die Autokorrelation durch verschiedene Fehlermodelle berücksichtigt.

Ziel dieser Arbeit soll die Erstellung eines nichtlinearen, hierarchischen und gemischten Modells (abgekürzt durch HNLMM für **h**ierarchical **n**on-**l**inear **m**ixed **m**odel) für das Baumhöhenwachstum der Fichte in Baden-Württemberg sein. Als Datengrundlage dienen longitudinale Baumhöhenmessungen aus aktuellen Inventuren der Immissionsökologischen Waldschadenserhebung (IWE) und historische Messungen von Mitscherlich (1957). In dieser Arbeit wird die Theorie der nichtlinearen, hierarchischen und gemischten Modelle erörtert. Nach der Modellbildung und der Parameterschätzung folgt eine Anwendung zur Evaluierung an einem unabhängigen Datenbestand. Die Ergebnisse der Anwendung werden kritisch gewürdigt und diskutiert. Es folgt ein Vergleich der nichtlinearen, gemischten Modelle mit den bewährten traditionellen Site-Index-Modellen. Abschließend werden der Erkenntnisgewinn und der Nutzen für die Praxis durch das erstellte Modell diskutiert.

1.3 Hypothesen

Für die vorliegende Arbeit werden folgende Hypothesen aufgestellt:

1. Das Baumhöhenwachstum der Fichte in Baden-Württemberg lässt sich durch ein nichtlineares, hierarchisches und gemischtes Modell beschreiben.

2. Mit einem nichtlinearen, hierarchischen und gemischten Modell lassen sich Baumhöhen treffgenauer und effizienter fortschreiben als mit herkömmlichen Bonitätsfächern.

3. Mit einem nichtlinearen, hierarchischen und gemischten Modell lassen sich zeitliche und räumliche Variation des Baumhöhenwachstums der Fichte in Baden-Württemberg abbilden.

Im deutschsprachigen Raum sind bislang keine Erfahrungen mit nichtlinearen, hierarchischen und gemischten Modellen für forstliche Wachstumsprozesse vorhanden. Die dargestellten Studien aus dem Abschnitt 1.2 behandeln zudem fast ausnahmslos das Wachstum von schnellwüchsigen Plantagenbeständen. Die Reichweite der Regressorvariable, meist das Baumalter, ist in diesen Studien vergleichsweise gering im Gegensatz zu dem in Baden-Württemberg beobachtbaren Rahmen von Alterswerten der Fichte. Die longitudinalen Beobachtungen jener Studien beginnen meistens in sehr frühen Altern.

Im Datenbestand der Baden-Württembergischen IWE finden sich über 100-jährige Bäume. Davon stellen Höhentriebsmessungen für die letzten 20 Lebensjahre nur einen kurzen Zeitabschnitt der gesamten Lebensdauer dar. Dieser Sachverhalt kann Probleme bei den numerischen Verfahren zur Parameterschätzung auslösen. Deswegen ist die Hypothese 1 zugleich als technische Herausforderung anzusehen. Die Brauchbarkeit verfügbarer Statistik-Software wird zugleich geprüft. Außerdem können bei Verifizierung der Hypothese die zugrunde liegenden Daten als geeignet für die Modellierung angesehen werden.

Wie später noch genauer dargestellt wird, besitzen die auf Zufallsparametern basierenden gemischten Modelle eine höhere Flexibilität im Vergleich zu starren Bonitätsfächern. Diese manifestiert sich in der Fähigkeit, anhand von Beobachtungen effiziente Kalibrierungen an die Verhältnisse vor Ort vorzunehmen. Gelänge es in dieser Studie, Hypothese 2 zu verifizieren, wäre ein wissenschaftlicher Fortschritt mit der Verfahrensverbesserung von Modellen zur Fortschreibung des Baumwachstums erzielt.

Ein weiteres Ziel dieser Studie ist die Untersuchung zeitlicher und räumlicher Variation des Fichtenhöhenwachstums. Mit Hypothese 3 nehmen wir an, dass sich der Höhenwachstumsverlauf von Fichten in Baden-Württemberg im Laufe der Zeit verändert hat. Anlässe für diese Annahme sind die Ergebnisse von Untheim (1996) und Moosmayer et al. (2003). Wenn es gelänge, mittels fixer Effekte aus dem gemischten Ansatz dieser Studie, die zeitlichen und regionalen Trends zu modellieren, wären wichtige wissenschaftliche Erkenntnisse gewonnen. Die Trends könnten dann nicht nur regressionsanalytisch beschrieben werden, sondern sie würden auch als Modellgrößen dienen. Es wäre dann ein Beitrag geleistet, die zu beobachtenden Phänomene modellhaft abzubilden. Die Integration der Zusammenhänge in ein realistisches Höhenwachstumsmodell erlaubt statistisch abgesicherte Schlussfolgerungen über Bonitätsveränderungen im Laufe der Zeit.

Kapitel 2

Datengrundlage

2.1 Daten für die Modellbildung

2.1.1 Immissionsökologische Waldzustandserhebung (IWE)

Ergänzend zur jährlichen terrestrischen Waldschadensinventur werden in Baden-Württemberg periodische Immissionsökologische Waldzustandserhebungen (IWE) vorgenommen (Schöpfer und Hradetzky, 1983; Schöpfer, 1994). Die IWE wurde in den Jahren 1983, 1988, 1994 und 2001 auf einem permanenten Probeflächennetz mit systematischer Rasterweite von 4x4 km durchgeführt. Bei den Feldaufnahmen wurden in den Jahren 1983, 1988 und 1994 pro Stichprobenpunkt zwei Bäume gefällt und im Jahr 2001 jeweils nur ein Baum. An den Probebäumen wurde je nach Inventur die Länge einer unterschiedlichen Anzahl an Jahreshöhentrieben von der Baumspitze abwärts gemessen (1983: 20 Stück, 1988: 25, 1994: 10, 2001: 10). Außerdem wurden an Stammscheiben in 1,3 m Höhe die Jahrringe analysiert. Die Daten der IWE, die in diese Studie eingeflossen sind, werden in den Tabellen 8.3 und 8.4 des Anhangs beschrieben.

2.1.2 Mitscherlichs Erhebung

In den Jahren 1954 und 1955 wurden unter der Leitung von Prof. Mitscherlich im Badischen Landesteil 306 einmalige Versuchsflächen aufgenommen [1]. Die Flächen wurden in ihrer Größe so gewählt, dass sie in etwa 100 Bäume umfassten. An allen Stämmen wurde der Brusthöhendurchmesser gemessen. An der Hälfte der Bäume wurden zwei Bohrspäne von gegenüberliegenden Seiten entnommen. In der Nähe der Probeflächen wurden jeweils 10 vorherrschende Stämme gefällt und ihre Höhenentwicklung in fünfjährigen Abständen gemessen. Dazu wurden die Probestämme auf der Oberseite entastet, die Äste am Ort belassen, hinter dem 15. und 30. Quirl ein Querschnitt geführt und die Länge von jeweils 5 Jahrestrieben beginnend von der Baumspitze gemessen. Zur Kontrolle wurde die Anzahl der Jahrestriebe mit der Anzahl der Jahrringe abgeglichen. Diese Kontrolle wurde als unbedingt notwendig erachtet. Besonders bei älteren Fichten war das Messen von weit zurückliegenden Höhentrieben schwierig, da die Knospenschuppenspuren an der Quirlbasis nicht mehr deutlich zu erkennen gewesen sind. Die Höhentriebe wurden so weit wie möglich zurückverfolgt, im allgemeinen 40 bis 50 Jahre. Zusätzlich wurden zur Bestimmung der Schaftform die Durchmesser in Brusthöhe und in 5 Sektionen gleicher Länge gemessen. Außerdem wurde versucht, ein oder zwei Bohrspäne bis auf den Kern zu entnehmen. Das Alter der Bäume wurde durch Zählung der Jahrringe am Stock festgestellt.

2.1.3 Modellierungsdatensatz

Für die Aufklärung räumlicher und zeitlicher Variation (Hypothese 3) wurden die aktuellen longitudinalen Daten aus der IWE und die historischen Messreihen von Mitscherlich zusammengeführt. Die Probeflächen Mitscherlichs waren geklumpt in folgenden drei Großräumen Baden-Württembergs angelegt:

- im Odenwald

[1] Dem unermüdlichen Einsatz von Prof. Schöpfer und Prof. Hradetzky ist es zu verdanken, dass die gesamten Erfassungsdaten aus Mitscherlichs Studie digital gesichert werden konnten.

- auf den westlichen Teilen der Jung- und Altmoräne des südwestdeutschen Alpenvorlandes
- im Südschwarzwald und Baar-Wutach

Da die Erhebungen Mitscherlichs nicht landesweit systematisch durchgeführt wurden, mussten die im landesweit systematischen Raster erfassten Daten der IWE vor der Zusammenführung auf die betreffenden Regionen reduziert werden. Für die weitere Analyse werden die Probekreisdaten in dieser Studie wie Zufallsauswahlen aus den einzelnen Regionen behandelt. Einzelne wenige Daten von Probekreisen wurden wegen nicht reparabler Erfassungsfehler von der weiteren Untersuchung als Ausreißer ausgeschlossen. In Abbildung 2.1 ist die Lage der Probeflächen grafisch dargestellt, deren longitudinale Daten in dieser Studie verwendet wurden.

Insgesamt wurden in der vorliegenden Arbeit Baumhöhenmesswerte von 433 Probeflächen verwertet. Mitscherlichs Daten umfassen 255 einmalige Probeflächenaufnahmen aus den Jahren 1954 und 1955. Aus der IWE stammen 178 Probeflächen. 106 Probeflächen der IWE wurden zu vier Inventurterminen aufgenommen, 25 Probeflächen wurden dreimal erfasst, 39 zweimal und 8 einmal. Insgesamt wurden die Höhentriebsmessungen von 3397 Bäumen in diese Studie einbezogen; davon stammen 2374 Bäume aus Mitscherlichs Erhebung und 1023 aus den Inventuren der IWE. Zusammenfassende Statistiken können den Tabellen 8.3 und 8.4 des Anhangs entnommen werden.

2.2 Daten für die Modellevaluierung

Nach der Erstellung des Modells soll dessen Güte an unabhängigen Daten geprüft werden. In den Jahren 1999 bis 2002 wurden von der Forstlichen Versuchs- und Forschungsanstalt Freiburg (FVA) longitudinale Messdaten zu Baumhöhen der Fichte im Südschwarzwald gewonnen (Kändler und Riemer, 2005). Diese Erhebungen waren Teil der Untersuchungen im Rahmen des Projektes „Zukunftsorientierte Forstwirtschaft“ und wurden vom Bundesministerium für Ernährung, Landwirtschaft und Forsten (BMBF) finanziell unterstützt.

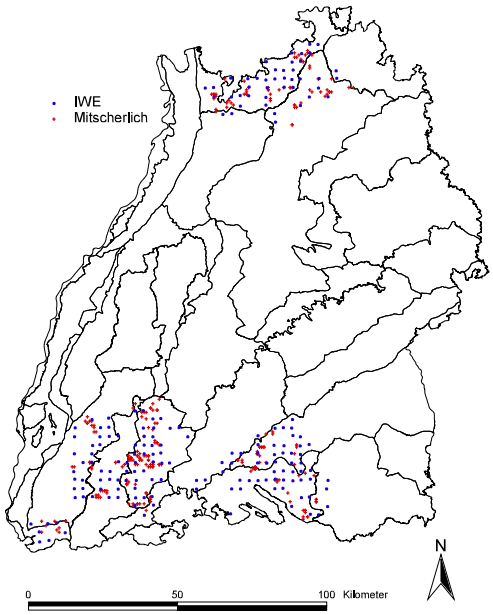


Abbildung 2.1: *Lage der Probeflächen, von denen Baumhöhenbeobachtungen für die Modellanpassung verwendet wurden.*

Das Vorhaben hatte zum Ziel, die Wuchsleistung der Hauptbaumarten Fichte, Tanne und Buche in Abhängigkeit von Standortsgrößen zu modellieren. Dafür wurden auf 51 Probeflächen Höhentriebsmessungen und Stammscheibenanalysen an 368 gefällten Fichten durchgeführt. In Tabelle 8.3 des Anhangs sind zusammenfassende Statistiken der Evaluierungsdaten aufgeführt. Die Daten aus dieser Studie werden in Anlehnung an das Projekt „Zukunftsorientierte Forstwirtschaft“ fortan als ZF-Daten bezeichnet.

Kapitel 3

Methode

3.1 Das nichtlineare Regressionsmodell

3.1.1 Begründung der Nichtlinearität

Unter einem nichtlinearen Modell wird nach Ratkowsky (1989) ein Modell verstanden, in dem mindestens ein Parameter in nichtlinearer Form erscheint. Schabenberger und Pierce (2002, S. 27) spezifizieren, dass es sich um ein lineares Modell handelt, wenn die partiellen Ableitungen des Modells nach allen Parametern von keinem einzigen Parameter abhängen, andernfalls handelt es sich um ein nichtlineares. Für den gemischten Modell-Ansatz müssen wir diese Definition nach Pinheiro und Bates (2002, S. 274) erweitern und fassen nichtlineare gemischte Modelle als Erweiterung linearer gemischter Modelle auf, bei denen die Regressionsfunktion einen nichtlinearen Zusammenhang zwischen Response auf der einen Seite und fixen und zufälligen Effekten auf der anderen Seite darstellt. Mit Hilfe nichtlinearer Modelle werden oft Sachverhalte deterministischer Natur beschrieben. Für unsere Studie wird ein Modell für einen Mechanismus benötigt, der die Responsevariable Baumhöhe hervorbringt. Die Parameter in einem nichtlinearen Modell haben folglich eine natürliche und physikalische Bedeutung; ihre Werte können deshalb oft biologisch interpretiert werden. Ein nichtlineares Modell benötigt in der Regel nur wenige Parameter und genügt damit dem Prinzip der Sparsamkeit. Außerdem liefern nichtlineare Modelle

überwiegend glaubwürdige und biologisch plausible Schätzungen für die Responsevariable außerhalb des Beobachtungsrahmens des Datenbereiches; dieses leisten lineare Modelle und speziell hochgradige Polynome nicht.

Mit der Nichtlinearität werden jedoch auch Probleme erkauft. Die Parameterschätzungen werden vornehmlich mittels Maximum-Likelihood-Techniken vorgenommen. Die Likelihood ist nach Sorensen (2002, S. 121) eine Funktion von Parametern bei gegebenen Realisationen der Zielvariable, die proportional zur Dichtefunktion der Beobachtungen ist. Das Ziel der Optimierung ist, diejenigen Parameterwerte zu finden, die die Likelihood maximieren. Für den Fall eines linearen Modells kann die Likelihood direkt berechnet werden. Bei einer nichtlinearen Grundfunktion ist die Likelihood nichtlinear in den Zufallsparametern, deshalb kann sie nicht explizit mit einer Gleichung ausgedrückt werden und muss über eine Näherungsfunktion numerisch geschätzt werden. Die aufwändigen Optimierungsalgorithmen stellen hohe Anforderungen an die Rechnerleistung und können leicht unplausible Schätzergebnisse erzeugen.

Wegen dieser numerischen Hürden finden sich noch zahlreiche aktuelle Studien, in denen Wachstumsprozesse mittels linearer gemischter Modelle beschrieben werden.

3.1.2 Die nichtlineare Grundfunktion

In ersten Versuchen wurde die Richards-Funktion als nichtlineare Grundfunktion verwendet. Dabei wurde wie schon in der Arbeit von Fang und Bailey (2001) der asymptotische Parameter durch einen Erwartungswert-Parameter ersetzt. Bei den vorliegenden Daten traten jedoch Konvergenzschwierigkeiten und biologisch unplausible Kurvenformen auf. Die Substitutionen der anderen Parameter minderten diese Probleme nicht. Zur Richards-Funktion schreibt Ratkowsky (1989, S. 47, 109, 140-141), dass sie in den Forstwissenschaften weite Anwendung fände, aber allgemein sehr schlechte statistische Eigenschaften habe, die auch durch Reparametrisierung nicht zu verbessern seien. Die gleichen numerischen Probleme traten auch bei anderen Funktionen mit zwei und drei Parametern aus der

Monographie von Ratkowsky (1989) auf. Schließlich erwies sich die Trendfunktion nach Sloboda (1971) als stabil.

$$y\left(t\right) = 65^{\psi_1} \cdot \left(\frac{y\left(t_0\right)}{65^{\psi_1}}\right)^{\exp\left[\frac{\psi_2}{(\psi_3-1)\cdot t^{(\psi_3-1)}} - \frac{\psi_2}{(\psi_3-1)\cdot t_0^{(\psi_3-1)}}\right]} \tag{3.1}$$

Die Parameter erwiesen durchweg gute Eigenschaften und nach Vorgabe geeigneter Startwerte trat stets Konvergenz ein. Die Funktion aus Gleichung 3.1 kann im Ursprung durch $y(0) = 0$ stetig ergänzt werden. Sie weist einen einzigen Wendepunkt und eine zur Abszisse parallele Asymptote auf. Mit dieser Funktion ergibt sich die Höhe $y\left(t\right)$ eines Baumes im Alter t aus der Vorkenntnis über die Höhe $y\left(t_0\right)$ im Alter t_0. Mit variierenden $y\left(t_0\right)$ über konstantem t_0 spannt sich eine Kurvenschar als Höhen-Bonitätsfächer auf, bei dem durch einen beliebigen Punkt nur eine einzige Kurve geht. Normalerweise erfolgt die Parameterschätzung durch den Ausgleich empirischer Richtungsfelder mit der Differentialform nach Technik von Wolf (1957). Dabei werden für Alters- und Höhenklassen die durchschnittlichen Steigungswerte über numerische Minimierung der Residuenquadratsumme ausgeglichen. Aus numerischen Gründen formulierte Sloboda (1971) den Parameter ψ_1 als Exponent zur Basis 65. Dadurch liegen die Werte aller Parameter in annähernd gleicher Höhe; die Optimierungsalgorithmen sollten dadurch leichter Konvergenz erzielen.

Für diese Arbeit wurde für die Anpassung die Trendfunktion in der Integralform verwendet. Dazu wurde $y\left(t_0\right)$ durch einen Höhenerwartungswert-Parameter $\psi_{ij}^{(4)}$ für konstantes Alter ersetzt.

Der Höhenwachstumsverlauf für den Baum j am Plot i wird schließlich beschrieben durch:

$$y_{ij}\left(t\right) = 65^{\psi_{ij}^{(1)}} \cdot \left(\frac{\psi_{ij}^{(4)}}{65^{\psi_{ij}^{(1)}}}\right)^{\exp\left[\frac{\psi_{ij}^{(2)}}{\left(\psi_{ij}^{(3)}-1\right)\cdot t^{\left(\psi_{ij}^{(3)}-1\right)}} - \frac{\psi_{ij}^{(2)}}{\left(\psi_{ij}^{(3)}-1\right)\cdot t_0^{\left(\psi_{ij}^{(3)}-1\right)}}\right]} + \epsilon_{ij}\left(t\right) \tag{3.2}$$

$\epsilon_{ij}\left(t\right)$ sei der Einfachheit wegen ein unabhängiger, normalverteilter Restfehler, der im Mittel 0 ist.

3.2 Theorie hierarchischer und gemischter Modelle

3.2.1 Lineare gemischte Modelle zur Einführung

Wegen der zum Teil recht schwer zu überschaubaren Notation wird die Theorie der gemischten Modelle zunächst für den linearen Ansatz beschrieben. Die Darstellung in den folgenden Abschnitten folgt der Notation von Schabenberger und Pierce (2002). Die Gruppierung der Daten über mehrere Ebenen wird der Einfachheit erst einmal nicht berücksichtigt. Wir nehmen daher an, es läge eine Population mit m Gruppen vor. Für die i-te Gruppe dieser Population liegen n_i Beobachtungen der Responsevariable vor. Der Vektor y_i der Responsevariable für die Gruppe i ergibt sich aus:

$$\begin{aligned} &y_i = X_i\beta + Z_i b_i + \epsilon_i \qquad i = 1, \ldots, m \\ &b_i \sim N(0, D) \qquad \epsilon_i \sim N(0, R_i) \\ &R_i = \sigma^2 I \qquad Cov[b_i, \epsilon_i] = 0 \end{aligned} \tag{3.3}$$

Dabei ist β ein Vektor mit Parameterwerten für fixe Effekte in q Zeilen. Die Parameterwerte sind konstant für die Population. Der Vektor b_i enthält die zufälligen Effekte in r Zeilen. q und r sind konstant für die Population. Die Design-Matrix X_i mit der Dimension $n_i \times q$ enthält Regressoren für die korrespondierenden Parameterwerte der fixen Effekte und die Design-Matrix Z_i mit der Dimension $n_i \times r$ enthält Regressoren für die zufälligen Effekte. Die Matrix Z_i enthält nur Einträge mit Werten von 0 und 1. Der Vektor ϵ_i enthält in n_i Reihen den stochastischen Restfehler für die einzelnen Beobachtungen. Die $(n_i \times n_i)$ Restfehlermatrix R_i trägt lediglich auf der Hauptdiagonalen mit σ^2 von 0 verschiedene Elemente. Nach einer grundsätzlicher Annahme sind diese Restfehler also voneinander unabhängig und mit der Varianz σ^2 normalverteilt um den Mittelwert von 0. Die Zufallseffekte zwischen den Gruppen werden mit $Cov[b_i, b_{i'}] = 0$ für $i \neq i'$ als unabhängig angesehen. Auch zwischen den Zufallseffekten und den Restfehlern einer Gruppe besteht kein Zusammenhang. Die Abhängigkeit der verschiedenen Zufallseffekte einer Gruppe untereinander beschreibt die $(r \times r)$ Varianz-Kovarianz-Matrix D. Auf deren

Hauptdiagonale stehen die Varianzen der Zufallsparameter, die das Ausmaß der Variation beschreiben, die durch uns unbekannte Effekte hervorgerufen wird.

Die Anzahl der Beobachtungen kann von Gruppe zu Gruppe $n_i \neq n_{i'}$ mit $i \neq i'$ unterschiedlich sein und damit auch die Anzahl der Zeilen jener Design-Matrizen. Die Daten heißen in diesem Fall unbalanciert. Die gesamte Variation innerhalb der Population wird durch die gruppenspezifischen fixen und zufälligen Effekte sowie den Restfehler hervorgerufen. Die Variation zwischen den Gruppen heißt *Inter-Clustervariation*, sie wird verursacht durch die Zufallsparameter in b_i. Die Variation innerhalb einer Gruppe heißt *Intra-Clustervariation* und wird bei dem einstufigen Modell lediglich durch den Restfehler verursacht. Wenn die gruppenspezifische Response betrachtet werden soll, dann muss b_i als Bedingung formuliert werden. Die Intra-Cluster-Variation ist dann die einzige Zufallskomponente auf der Gruppenebene:

$$\begin{aligned} E\left[y_i|b_i\right] &= X_i\beta + Z_i b_i \\ Var\left[y_i|b_i\right] &= R_i \end{aligned} \tag{3.4}$$

Die mittlere für die Population zu erwartende Response (Populations-Mittelwert-Modell) ergibt sich aus:

$$E\left[y_i\right] = E\left[E\left[y_i|b_i\right]\right] = E\left[X_i\beta + Z_i b_i\right] = X_i\beta \tag{3.5}$$

mit der Varianz

$$Var\left[y_i\right] = R_i + Z_i D Z_i^T \tag{3.6}$$

3.2.2 Das nichtlineare, hierarchische und gemischte Modell

Die Matrix-Notationen der bedeutendsten Autoren unterscheiden sich nicht wesentlich (Lindstrom und Bates, 1990), (Davidian, 2001, Kapitel 15, S. 414 ff.), (Davidian und Giltinan, 2003, S. 392), (Davidian und Giltinan, 1995, Kapitel 4, S. 97 ff.), (Pinheiro und Bates, 2002, S. 306 ff.), (Schabenberger und Pierce, 2002, S. 528 ff.). Für eine bessere Anschaulichkeit wird im folgenden Bezug zur vorliegenden Fragestellung genommen. In

dieser Studie wurden 3397 Bäume an 433 Probeflächen aufgenommen. Man sollte annehmen, dass die Höhenwachstumsverläufe der Bäume an einer bestimmten Probefläche ähnlich sind. Denn die Bäume standen räumlich nah beisammen und gehörten alle zum selben Waldbestand. Dadurch genossen sie nahezu einheitliche waldbauliche Eingriffe, wie z.B. gleiche Zeitpunkte und Intensitäten von Durchforstungen oder ähnlich intensive historische Waldnutzungsformen, wie z.B. Streunutzungen. Die Bäume einer Probefläche standen fast am gleichen Ort von geographischer Länge, Breite und Höhe. Sie waren ungefähr dem gleichem Angebot an Wasser, Wärme, Nähr- und Schadstoffen aus Luft und Boden ausgesetzt. Die forstgenetische Herkunft von Bäumen auf ein und dem selben Probekreis dürfte stärker übereinstimmen als die von Bäumen auf verschiedenen Probekreisen. Wegen dieser Gemeinsamkeiten für Bäume am selben Probekreis wird im folgenden angenommen, dass die Kennung eines Probekreises das Gruppenmerkmal der ersten und damit obersten Ebene ist.

Die Höhen-Messwerte eines Baumes zu bestimmten Altern beschreiben seinen Höhenwachstumsverlauf über die Zeit. Sie können biologisch zu einer Wachstumskurve verbunden werden. Deshalb ist der Baum das Gruppierungsmerkmal der zweiten Ebene. Die Gruppierungs-Ebene der Bäume gliedert sich hierarchisch unter die Ebene der Probekreise ein. Zur Beschreibung des Wachstums unter besonderer Berücksichtigung der zugrunde liegenden Datenstruktur ist ein hierarchisches, nichtlineares und gemischtes Modell mit zwei Gruppierungs-Ebenen zweckmäßig.

Wir definieren die k-te Beobachtung des j-ten Baum am i-ten Plot durch:

$$\begin{aligned} &y_{ijk} = f\left(\phi_{ij}, t_{ijk}\right) + \epsilon_{ijk} \\ &i = 1, \ldots, m \qquad j = 1, \ldots, n_i \qquad k = 1, \ldots, o_{ij} \end{aligned} \tag{3.7}$$

Dabei ist f eine nichtlineare Funktion vom Prädiktor t_{ijk} und dem Parametervektor ϕ_{ij} mit der Länge p als die Anzahl der Funktions-Parameter. In diesem Fall ist der Prädiktor ein Skalar mit dem Wert des Baumalters zum korrespondierenden einzelnen Höhenmesswert. Die Funktionsparameter sind für alle Responsewerte desselben Baumes konstant. In dieser Arbeit wird das zweistufige nichtlineare Modell in Matrix-Notation auf der ers-

ten und damit höchsten Ebene beschrieben; auf dieser Ebene geschieht auch die spätere Kalibrierung mit der Vorhersage der Zufallsparameter. Die untereinander geschachtelten Gruppen unterschiedlicher Ebenen können sehr gut durch geblockte Matrizen dargestellt werden. Das hat zudem für die Notation den Vorteil, dass leicht auftretende Inkonsistenzen bei den Indices vermieden werden können. Wir definieren deshalb den Responsevektor für alle n_i Bäume am Probekreis i durch:

$$y_i = f_i\left(\phi_i, \theta_i\right) + \epsilon_i \tag{3.8}$$

$$y_i = \begin{bmatrix} y_{i11} \\ \vdots \\ y_{i1o_{i1}} \\ \vdots \\ \vdots \\ y_{in_i1} \\ \vdots \\ y_{in_io_{in_i}} \end{bmatrix}, \quad f_i\left(\phi_i, \theta_i\right) = \begin{bmatrix} f\left(\phi_{i1}, t_{i11}\right) \\ \vdots \\ f\left(\phi_{i1}, t_{i1o_{i1}}\right) \\ \vdots \\ \vdots \\ f\left(\phi_{in_i}, t_{in_i1}\right) \\ \vdots \\ f\left(\phi_{in_i}, t_{in_io_{in_i}}\right) \end{bmatrix}, \quad \epsilon_i = \begin{bmatrix} \epsilon_{i11} \\ \vdots \\ \epsilon_{i1o_{i1}} \\ \vdots \\ \vdots \\ \epsilon_{in_i1} \\ \vdots \\ \epsilon_{in_io_{in_i}} \end{bmatrix}$$

$$\phi_i = \begin{bmatrix} \phi_{i1} \\ \vdots \\ \phi_{in_i} \end{bmatrix} = \begin{bmatrix} \psi_{i1}^{(1)} \\ \vdots \\ \psi_{i1}^{(p)} \\ \vdots \\ \vdots \\ \psi_{in_i}^{(1)} \\ \vdots \\ \psi_{in_i}^{(p)} \end{bmatrix}, \quad \theta_i = \begin{bmatrix} \theta_{i1} \\ \vdots \\ \theta_{in_i} \end{bmatrix} = \begin{bmatrix} t_{i11} \\ \vdots \\ t_{i1o_{i1}} \\ \vdots \\ \vdots \\ t_{in_i1} \\ \vdots \\ t_{in_io_{in_i}} \end{bmatrix}$$

$$\epsilon_i \sim N\left[0, R_i\right] \qquad R_i = \sigma^2 I \left[\sum_{j=1}^{n_i} o_{ij}\right]$$

Der Spaltenvektor y_i enthält die Response für alle n_i Ebene-2-Gruppen (Bäume) der i-ten Ebene-1-Gruppe (Plot) in übereinanderstehenden Blocks. Der Vektor θ_i enthält die Werte für die Prädiktorvariable Baumalter und ϵ_i enthält die Restfehler. Für den j-ten Baum am i-ten Plot liegen o_{ij} Messungen vor.

Der Parametervektor ergibt sich aus den gemischten Effekten durch:

$$\begin{aligned}
&\phi_i = A_i\beta + B_i b_i \qquad (3.9)\\
&\beta = \left(\iota^T, \kappa^T, \lambda^T\right)^T\\
&b_i = \left(\zeta_i^T, \zeta_{i1}^T, \ldots, \zeta_{in_i}^T\right)^T\\
&\zeta_i = \left(z_i^{(1)}, \ldots, z_i^{(p)}\right)^T \quad \zeta_{ij} = \left(z_{ij}^{(1)}, \ldots, z_{ij}^{(p)}\right)^T\\
&b_i \sim N\left(0, D_i\right)\\
&D_i = \begin{bmatrix} \Delta_1 & 0 & \cdots & 0\\ 0 & \Delta_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \Delta_2 \end{bmatrix}\\
&\Delta_1 = \begin{bmatrix} Var\left[z_i^{(1)}\right] & \cdots & Cov\left[z_i^{(1)}, z_i^{(p)}\right]\\ \vdots & \ddots & \vdots\\ Cov\left[z_i^{(1)}, z_i^{(p)}\right] & \cdots & Var\left[z_i^{(p)}\right] \end{bmatrix}\\
&\Delta_2 = \begin{bmatrix} Var\left[z_{ij}^{(1)}\right] & \cdots & Cov\left[z_{ij}^{(1)}, z_{ij}^{(p)}\right]\\ \vdots & \ddots & \vdots\\ Cov\left[z_{ij}^{(1)}, z_{ij}^{(p)}\right] & \cdots & Var\left[z_{ij}^{(p)}\right] \end{bmatrix}
\end{aligned}$$

Der Vektor β gehört zu den fixen Effekten und b_i enthält die Zufallsparameter. Die Matrizen A_i und B_i sind korrespondierende Design-Matrizen. Der Parametervektor ϕ_i enthält die p Parameter der Wachstumsfunktion für alle n_i Bäume am Plot i in $p \cdot n_i$ Zeilen untereinander geblockt.

Der Vektor β enthält die Parameterwerte für die von Kovariablen abhängigen fixen Effekte, sowie die Werte für Interzepte. Im oberen Block enthält er mit ι die Parameterwerte für die fixen Effekte auf Populationsebene und im nächsten Block mit κ, diejenigen für die Plotebene. Darunter folgt ein Block mit den fixen-Effekt-Parametern für die Baumebene

in Form von λ. Die Matrix A_i trägt in ihren Elementen für Interzept-fixe-Effekte lediglich eine 1 und für die von Kovariablen abhängigen fixen Effekte die zu beobachtenden Werte dieser Kovariablen. Anschaulich gesprochen steuert die Matrix A_i die Einträge aus dem Vektor für die fixen Effekte an die richtige Stelle. Während die Parameter-Vektoren ι, κ und λ konstant für alle i und j sind, variieren die Einträge für die Regressor-Kovariablen in A_i.

Wir nehmen an, dass die fixen Effekte auf der obersten Ebene nicht die gesamte Variation zwischen den Mittelwertkurven für die einzelnen Probekreise bzw. nicht die gesamte Variation der Probekreismittelwertkurven um die Mittelwertkurve für die gesamte Population erklären können. Es gibt unbekannte Effekte, die den verbleibenden Teil dieser *Inter-Cluster-Variation* hervorrufen; diese bewirken die Zufallsparameter ζ_i auf der obersten Ebene, der Plot-Ebene. Gleichermaßen unterstellen wir, dass die fixen Effekte auf der Baumebene nicht die gesamte Variation innerhalb der einzelnen Probekreise erklären können. Der restliche Teil dieser *Intra-Cluster-Variation* wird durch zufällige Effekte mit den Parametern ζ_{ij} auf der Hierarchieebene der Bäume bewirkt. Der Vektor b_i enthält die Werte der Zufallsparameter in $p\,(1 + n_i)$ Zeilen. Die obersten p Zeilen enthalten die Zufallseffekte $z_i^{(1)}$ bis $z_i^{(p)}$ auf Plotebene und darunter reihen sich n_i Blocks zu je p Zeilen mit den Zufallsparametern auf Baumebene mit jeweils $z_{ij}^{(1)}$ bis $z_{ij}^{(p)}$ für jeden j-ten Baum auf. Die Matrix B_i trägt gewöhnlich nur die Werte 0 und 1 und lenkt die Zufallsparameterwerte aus b_i ihrerseits an die richtige Stelle. Während die Einträge der Matrix B_i konstant sind, variieren die Werte der Zufallsparameter zwischen den Plots und zwischen den Bäumen. Die Rest-Variation, die nicht durch fixe oder zufällige Effekte beschrieben werden kann, entfällt auf die unabhängigen, normal verteilten Restfehler in ϵ_i.

Die fixen Effekte auf Populationsebene gehen als Interzepte in die lineare Verkettung von fixen und zufälligen Effekte für die Funktionsparameter ein. Fixe Effekte können von einzelnen Kovariablen verschiedener Skalierung abhängen. Denkbar wäre zum Beispiel ein fixer Effekt auf Ebene der Plots, der von der metrischen Kovariable Höhenlage ü. NN abhängt oder ein fixer Effekt, der die Abhängigkeit von einer kategorial kodierten

Standortseinheit beschreibt. Die Ausprägungen dieser zusätzlichen Regressoren finden sich in Matrix A_i und ihre Regressor-Parameterwerte sind in β enthalten.

Sowohl die zufälligen Effekte auf der Plotebene, als auch diejenigen auf Baumebene seien normal verteilt mit bestimmten Varianzen. Die Zufallsparameter in b_i werden mit $Cov\,[b_i, b_{i'}] = 0$ für $i \neq i'$ als unabhängig zwischen den Gruppen angesehen. Zwischen den Zufallseffekten und dem Restfehler wird mit $Cov\,[b_i, e_i] = 0$ ebenfalls Unabhängigkeit postuliert (Pinheiro und Bates, 2002, S. 58). Die Zufallseffekte zwischen den Ebenen seien mit $Cov\,[\zeta_i, \zeta_{ij}] = 0$ für alle j ebenfalls unabhängig (Pinheiro und Bates, 2002, S. 60); auf einer Ebene können sie hingegen miteinander korrelieren. Die Struktur und Stärke der Abhängigkeit wird in der Varianz-Kovarianz-Matrix D_i manifestiert. Die Begriffe Varianz-Kovarianz-Matrix und Kovarianz-Matrix werden synonym verwendet. Der linke obere Block der Kovarianz-Matrix enthält mit Δ_1 die Kovarianzen der Zufallseffekte auf der ersten Gruppierungsebene. Darunter sind die Kovarianzen der zweiten Ebene Δ_2 in der Diagonale n_i mal diagonal geblockt.

Weil die allgemeine Darstellung des Modells in diesem Abschnitt recht abstrakt erscheinen kann, wird später im Abschnitt 4.2 eine detaillierte Beschreibung anhand eines Beispiels vorgenommen.

3.2.3 Parameterschätzung

Definitionsgemäß werden unter den Parametern eines gemischten Modells die Regressor-Parameter für die fixen Effekte und die Varianzen und Kovarianzen der Zufallsparameter, sowie die Varianz des Restfehlers verstanden. Davidian (2001, Kap. 15, S. 419) bezeichnet die Schätzung dieser Parameter als hauptsächliche Aufgabe der Analyse mit gemischten Modellen, da diese Parameter einerseits das typische Verhalten von individuellen Funktionsparametern und andererseits deren Variation in systematischer (fixe Effekte) sowie unerklärlicher (Zufallseffekte) Weise innerhalb der Population beschreiben. Davidian und Giltinan (1995), Davidian (2001) und Schabenberger und Pierce (2002, S. 530) vergleichen und beschreiben verschiedene Methoden zur Parameterschätzung. Diese Autoren unterscheiden grundsätzlich drei Methoden zur Parameterschätzung:

- individuelle Schätzmethoden
- Linearisierungs-Methoden
- Methoden, die auf Integral-Approximation basieren

Die bedeutendsten Parameterschätzmethoden basieren auf der Maximum-Likelihood-Technik. Die Likelihood ist (Sorensen, 2002, S. 121) eine Funktion von Parametern bei festen Realisationen der Responsevariable, die proportional zur Dichtefunktion der Beobachtungen ist. Der grundsätzliche Unterschied zwischen hierarchischen, gemischten Modellen in der linearen und nichtlinearen Form besteht in der Fähigkeit, die marginale Verteilungsfunktion der Response y_i für die Likelihood-Schätzung explizit zu formulieren (Davidian und Giltinan, 1995, S. 118).

Wir nehmen an, dass wir für alle Bäume eines jeden Probekreises das Tripel (y_i, θ_i, A_i) beobachten können. Die Baumhöhen des Baumes sind in y_i und die zugehörigen Werte der Prädiktorvariable Baumalter in θ_i enthalten. A_i enthält die Werte von Kovariablen für die fixen Effekte, die außerdem am Baum und Probekreis zu beobachten sind.

Die marginale Dichte für die Response y_i gegeben θ_i ist $p\left(y_i|\theta_i, A_i, \beta, R_i, D_i\right)$. Weil die nicht beobachtbaren Zufallseffekte mit ihren Parametern b_i Bestandteil des Modells sind, muss die marginale Dichte für gegebene b_i integriert werden (vgl. Davidian, 2001, S. 420):

$$\begin{aligned} p\left(y_i|\theta_i, A_i, \beta, R_i, D_i\right) &= \int p\left(y_i, b_i|\theta_i, A_i, \beta, b_i, R_i, D_i\right) db_i \\ &= \int p\left(y_i|\theta_i, A_i, \beta, b_i, R_i\right) p\left(b_i|\theta_i, A_i, \beta, R_i, D_i\right) db_i \\ &= \int p\left(y_i|\theta_i, A_i, \beta, b_i, R_i\right) p\left(b_i|D_i, \right) db_i \end{aligned} \tag{3.10}$$

Die letzte Gleichung folgt, weil für die Zufallsparameter Unabhängigkeit von der Prädiktorvariable, von allen Kovariablen und von dem Restfehler angenommen wurde. Dabei ist $p\left(y_i|\theta_i, A_i, \beta, b_i, R_i\right)$ die von den Parametern bedingte Wahrscheinlichkeitsdichte für die Response y_i aller n_i Bäume am Plot i bei gegebenen b_i und $p\left(b_i|D_i\right)$ die Dichte von b_i.

Die Likelihood für das Modell aus den Gleichungsfolgen 3.8 und 3.9 als Dichtefunktion aller Realisationen in Abhängigkeit von den Parametern ist dann:

$$L = \prod_{i=1}^{m} \int p\left(y_i|\theta_i, A_i, \beta, b_i, R_i\right) p\left(b_i|D_i\right) db_i \tag{3.11}$$

Für lineare, hierarchische und gemischte Modelle sind sowohl die bedingte Verteilung von y_i gegeben b_i als auch b_i normal verteilt und deshalb kann das Integral explizit formuliert werden. Im Gegensatz dazu kann bei nichtlinearen, hierarchischen und gemischten Modellen wegen der nichtlinearen Responsefunktion f bzw. f_i (siehe Gleichung 3.8) das Integral grundsätzlich nicht geschätzt werden. Dieser Sachverhalt wird bei Davidian und Giltinan (1995, Kap. 3.3 und 4.4) ausführlich dargelegt. Die wichtigsten Lösungsansätze für das Integralproblem können in zwei Kategorien eingeteilt werden; die eine basiert auf *Linearisierung* und die andere auf *Approximation*. Bei der Linearisierungsmethode wird die nichtlineare Funktion durch ein lineares Modell näherungsweise ersetzt. Dieses Verfahren wird auch *Pseudo-Daten-Methode* genannt, weil nicht die Response selbst, sondern eine Funktion von ihr modelliert wird (Schabenberger und Pierce, 2002). Dabei sind einige Verfahren parametrisch und andere semiparametrisch. Bei den parametrischen Verfahren wird eine Verteilung für die Response angenommen, bei den semiparametrischen nicht. Die Methoden der Integralapproximation treffen Annahmen über die Verteilung der Zufallseffekte b_i und für die bedingte Verteilung der Response von y_i gegeben b_i und schätzen das Integral näherungsweise mit numerischen Verfahren.

Neben der *Linearisierung* und *Approximation* gibt es den *2-Stufen-Ansatz* oder auch *individuelle Schätzmethode* genannt. Weil mit der Schätzung der Likelihood numerische Probleme verbunden sein können, diskutieren Davidian und Giltinan (1995, S. 127) den *2-Stufen-Ansatz* als eine sinnvolle Alternative. In der ersten Stufe wird die nichtlineare Grundfunktion für jedes Individuum mittels GLS-Verfahren einzeln angepasst; damit erhält man vorläufige Funktionsparameter. In der zweiten Stufe werden in einem iterativen Prozess mit 2 Schritten die individuellen Parameterschätzwerte für die Population gemittelt. Im ersten Schritt der zweiten Stufe werden mit den Residuen der individuellen Anpassungen die Varianz und Kovarianz des Restfehlers geschätzt. Diese Schätzer

gehen dann in individuelle Gewichtungsmatrizen ein. Im zweiten Schritt werden die Funktionsparameter für jedes Individuum iterativ erneut geschätzt, indem die gewichtete Residuenquadratsumme individuell minimiert wird. Die nun erzeugten Funktionsparameter werden dann wieder als vorläufig angesehen und zur erneuten Formulierung einer Varianz-Kovarianz-Schätzung für den ersten Schritt der zweiten Stufe benutzt. Die Schätzung der Varianz-Kovarianz-Parameter für die Population im ersten Schritt der zweiten Stufe kann nach Davidian und Giltinan (1995) und Davidian (2001) über sogenannte gepoolte (zusammengefasste) Schätzer erfolgen. Dabei wird eine Pseudo-Likelihood-Größe aus den Residuen gebildet und über alle Individuen zusammengefasst geschätzt. Details sind bei Davidian (2001, Kap. 15.3, S. 426 ff.) zu finden.

Eine Voraussetzung für diesen zweistufigen Ansatz ist eine genügend große Anzahl von Beobachtungen für alle Individuen, damit glaubwürdige und biologisch plausible Kurvenverläufe angepasst werden können. In ersten Ansätzen dieser Arbeit wurden Versuche unternommen, dem *2-Stufen-Ansatz* folgend die Höhenbeobachtungen eines jeden Baumes über dem Alter mittels OLS und der `SAS 9.1`-Routine `Proc NLIN` funktional auszugleichen. Die Ergebnisse waren unbefriedigend. Bei rund einem Drittel aller Bäume kam es entweder zu Konvergenzproblemen bei der iterativen Minimierung der Residuenquadratsumme oder zu Parameterschätzungen, die zu unlogischen Kurvenformen geführt hätten. Die Ursache lag darin, dass die Beobachtungen eines Baumes eine zu geringe Anzahl und Reichweite für eine erfolgreiche individuelle Schätzung aufwiesen. Bei einem über hundert Jahre alten Baum stellen die Beobachtungen der letzten 10 Jahreshöhentriebe, wie bei der IWE 1994 vorgenommen, nur einen kurzen Abschnitt aus seinem gesamten Leben dar. Bei nicht wenigen Bäumen finden sich in diesem kurzen Abschnitt keine Anzeichen für biologisch plausible Positionen von Wendepunkt oder Asymptote. Die Ausgleichskurven, basierend auf individuellen Schätzern, können dann zwar nominell und optisch eine gute Anpassung darstellen, aber bei der Extrapolation würden manche Bäume „in den Himmel wachsen“. Für die individuellen Schätzungen schien der Datenumfang von Mitscherlich geeigneter zu sein. Die Höhentriebsmessungen wurden zwar nicht jährlich aber weit zurück bis in junges Alter vorgenommen. Sämtliche „Ausreißer-“Bäume mit Konvergenzschwierigkeiten

oder unlogischen Kurvenanpassungen nicht zu berücksichtigen, würde eine verzerrte Parameterschätzung zur Folge haben können, da die verbleibenden Bäume nicht länger eine Zufallsstichprobe aus dem Untersuchungsgebiet darstellen würden. Aufgrund der dargelegten Probleme wurde der *2-Stufen-Ansatz* zur Parameterschätzung in dieser Studie nicht weiter angewendet. Der Versuch mit diesem Ansatz kann dennoch als wertvoll betrachtet werden, da in dieser Phase der Modellbildung verschiedene nichtlineare Grundfunktionen hinsichtlich ihrer Flexibilität und Plausibilität erprobt wurden.

An dieser Stelle sei nochmals auf das eigentliche Problem hingewiesen, das mit der nichtlinearen Grundfunktion verbunden ist. Es gibt für das Ziel der Likelihoodschätzung keine explizite Darstellung für das Integral der marginalen Dichtefunktion. Die zweite Methode (Linearisierung) setzt direkt an der Wurzel des Übels an und linearisiert die Funktion mit einer Taylor-Entwicklung um einen Wert für b_i.

Mit den Annahmen, dass die Zufallseffekte in b_i und die Restfehler in ϵ_i normalverteilt sind, kann nach Pinheiro und Bates (2002) auch für die marginale Verteilung von y_i eine Normalverteilung unterstellt werden mit:

$$\begin{aligned} p\left(y_i|\theta_i, A_i, \beta, R_i, b_i\right) = {} & (2\pi)^{-h_i/2} \left|R_i\right|^{-1/2} \\ & \exp\left\{-(1/2)\left[y_i - f_i\left(\theta_i, A_i, \beta, b_i\right)\right]^T R_i \left[y_i - f_i\left(\theta_i, A_i, \beta, b_i\right)\right]\right\} \end{aligned} \tag{3.12}$$

Dabei ist f_i der Mittelwertvektor von y_i mit der Dimension $h_i = \sum_{j=1}^{n_i} o_{ij}$. Zunächst wird auf die *first-order-Methode* von Beal und Sheiner (1982, 1988) mit einer Taylor-Entwicklung um $b_i = E\left[b_i\right] = 0$ eingegangen. Dabei werden die Zufallsparameter auf ihren Erwartungswert von 0 gesetzt. Mit der Approximation von f_i

$$f_i\left(\theta_i, A_i, \beta, b_i\right) \approx f_i\left(\theta_i, A_i, \beta, 0\right) + Z_i\left(\theta_i, A_i, \beta\right) b_i \tag{3.13}$$

erhält man für Gleichung 3.12:

$$\begin{aligned} p\left(y_i|\theta_i, A_i, \beta, R_i, b_i\right) \approx {} & (2\pi)^{-h_i/2} \left|R_i\right|^{-1/2} \\ & \exp\Big\{-(1/2)\left[y_i - f_i\left(\theta_i, A_i, \beta, 0\right) - Z_i\left(\theta_i, A_i, \beta\right) b_i\right]^T \\ & R_i \left[y_i - f_i\left(\theta_i, A_i, \beta, 0\right) - Z_i\left(\theta_i, A_i, \beta\right) b_i\right]\Big\} \end{aligned} \tag{3.14}$$

Dabei ist $Z_i(\theta_i, A_i, \beta)$ die Jacobi-Matrix $\partial f_i(\theta_i, A_i, \beta, b_i)/\partial b_i$ geschätzt mit $b_i = 0$. Mit der Annahme, dass $p(b_i|D_i)$ normal verteilt mit Mittelwert 0 und Varianz-Kovarianz-Matrix D_i ist, ergibt sich hiermit das näherungsweise geschätzte Integral aus Gleichung 3.10 durch:

$$\int p(y_i|\theta_i, A_i, \beta, b_i)\, p_b(b_i|D_i,)\, db_i \approx (2\pi)^{-h_i/2} |V_i|^{-1/2} \exp\left\{-(1/2)\left[y_i - f_i(\theta_i, A_i, \beta, 0)\right]^T R_i \left[y_i - f_i(\theta_i, A_i, \beta, 0)\right]\right\} \tag{3.15}$$

Hierbei ist $V_i = Z_i D_i Z_i^T + R_i$ die Varianz der marginalen Dichteverteilung des Responsemittelwertvektors. Das Integral kann nun explizit ausgerechnet werden. Es ist numerisch einfacher, die *Log-Likelihood-Funktion* zu maximieren (Härdle und Simar, 2003, S. 174) bzw. die *negative Log-Likelihood-Funktion* zu minimieren. Der negative Logarithmus der Likelihood-Approximation wird über alle m Gruppen der ersten Hierarchieebene (Probeflächen) mittels iterativer Techniken minimiert:

$$-LL = -log\left(\prod_{i=1}^{m} \int p(y_i|\theta_i, A_i, \beta, b_i)\, p(b_i|D_i,)\, db_i\right) \tag{3.16}$$

Weil die Zufallseffekte auf ihren Erwartungswert von 0 gesetzt werden, bezeichnen Schabenberger und Pierce (2002) diese *first-order-Methode* auch als *Populations-Mittelwert-Expansion*. Ein besonderer Vorteil der Linearisierungs-Methoden ist nach Schabenberger und Pierce (2002), dass sie gerade nicht von der Notwendigkeit abhängigen, separat für jeden Cluster (Baum, Plot) das Modell anzupassen. Auch wenn Cluster nicht Informationen über den gesamten Beobachtungsrahmen beitragen, können sie trotzdem einen Beitrag bei der Linearisierungs-Methode leisten. Aus diesem Grund wurde in dieser Studie schließlich auch die Linearisierungs-Methode angewendet.

Anfangs wurde die Parameterschätzung mit der `SAS`-Prozedur `NLMIXED` vorgenommen. Da bei dieser Prozedur standardmäßig nur eine Gruppierungsebene unterstützt wird, wurde das Attribut mit den Kennungen für die Probeflächen als `subject` angewiesen und die einzelnen Bäume eines Plots mit Dummies kodiert. In der vorliegenden Studie wurde wie in der Programmhilfe von `SAS 9.1` angeraten, eine geeignete Methode zur numerischen Optimierung durch das Prinzip nach „trial and error“ ausgewählt. Die Quasi-Newton-

Methode, die auch die Voreinstellung ist, erwies die beste Balance zwischen Geschwindigkeit und Stabilität. Bei dieser Optimierung kann noch die Art der Aktualisierung der Hesse'schen Matrix vorgegeben werden. Hier tat ebenfalls die voreingestellte Methode mit der zweifachen Aktualisierung der Cholesky-faktorisierten Hesse'schen Matrix nach Broyden, Fletcher, Goldfarb und Shanno (SAS Institute Inc., 2003) gute Dienste. Desweiteren können Schrittweite und Suchrichtung für das Näherungsoptimum vorgegeben werden. Ein weiterer Schlüssel zur erfolgreichen Modellanpassung sind plausible Startwerte. Eine Eingrenzung der Parameterschätzwerte half in den meisten Fällen nicht weiter. Die häufigste Ursache für Konvergenzschwierigkeiten waren Modellmissspezifikationen wie z.B. Kodierungsfehler oder Überparametrisierungen. Deshalb sind die Bemühungen für die eigentliche Modellkonstruktion bezüglich des Zusammenspiels von fixen und zufälligen Effekten am wichtigsten. Die `SAS`-Prozedur `NLMIXED` zeigte jedoch mehrere Schwächen. Die Schätzungen für einzelne Fälle der IWE-Daten waren am Rande des Beobachtungsbereiches verzerrt. Die Ursachen dafür konnten nicht geklärt werden. Das Phänomen trat aber fast ausschließlich an IWE-Bäumen mit ihren relativ kurzen Beobachtungsreihen auf. Außerdem befolgten die Verteilungen der Zufallsparameter nicht die zuvor getroffenen Annahmen. So waren die vorhergesagten empirischen Zufallsparameter häufig um einen Mittelwert ungleich von 0 verteilt. Desweiteren konnte es passieren, dass für die von Kovariablen abhängigen Parameter zwar nominell hoch signifikante, jedoch unplausible Werte geschätzt wurden. Dadurch erhielten die Zufallsparameter ihrerseits wieder einen Trend. Der Theorie zu Folge sollten bei der Modellkonstruktion die Zufallsparameter aus Zufallsmodellen auf einen Trend über Kovariablen untersucht werden. Wenn sich Anzeichen für solche Abhängigkeiten ergäben, könnten die Kovariablen als fixe Effekte implementiert werden. Der Trend der Zufallsparameter im resultierenden gemischten Modell sollte dann bereinigt worden sein. Genau dieses stellte sich aber oft nicht ein. Die Parameterwerte der fixen Effekte wurden überschätzt und damit auch die Trends in den Mittelwertkurven. Dieses mussten die Zufallseffekte ihrerseits durch einen Trend wieder ausgleichen. Wegen dieser Schwierigkeiten mit der `SAS`-Prozedur `NLMIXED` wurde das Verhalten der `nlme`-Funktion aus dem `R`-Softwarepaket untersucht. Sämtliche Probleme der

SAS-Routine traten erfreulicherweise nicht auf. Die empirischen Vorhersagewerte der Zufallsparameter waren stets um den Mittelwert von 0 verteilt und die Parameter der fixen Effekte wurden so plausibel geschätzt, dass nach ihrer Implementierung die Zufallsparameter über die Kovariablen gleichmäßig um 0 streuten. Aus diesen Gründen wurde die Parameterschätzung schließlich mit der nlme-Funktion aus R vorgenommen.

Mit der nlme-Funktion wird eine erweiterte Methode der Linearisierung angeboten. Der Algorithmus wird beschrieben von Lindstrom und Bates (1990) und Pinheiro und Bates (2002, S. 313). Der Algorithmus alterniert zwischen zwei Schritten, einem „penalised nonlinear least squares"-Schritt (PNLS) und einem „linear mixed effects"-Schritt (LME). Diese Methode der Linearisierung wird von Schabenberger und Pierce (2002) als subjektspezifische Expansion bezeichnet, da die Linearisierung nicht um den Erwartungswert der Zufallsparameter von 0 vorgenommen wird, sondern um die individuellen Schätzwerte der Zufallsparameter. Die bedingten Schätzungen für die Parameterwerte der fixen Effekte β werden erreicht durch die Minimierung folgender PNLS-Funktion:

$$\sum_{i=1}^{m} \left[\|y_i - f_i(\beta, b_i, \theta_i)\|^2 + \|\Delta b_i\|^2 \right] \tag{3.17}$$

Im LME-Schritt wird der Präzisionsfaktor Δ basierend auf einer Taylor-Entwicklung der Modell-Funktion um momentane Schätzungen von β und b_i aktualisiert, ausgedrückt durch $\hat{\beta}^{(w)}$ und $\hat{b}_i^{(w)}$. Mit

$$\hat{X}_i^{(w)} = \left.\frac{\partial f_i}{\partial \beta^T}\right|_{\hat{\beta}^{(w)}, \hat{b}_i^{(w)}}, \qquad \hat{Z}_i^{(w)} = \left.\frac{\partial f_i}{\partial b_i^T}\right|_{\hat{\beta}^{(w)}, \hat{b}_i^{(w)}} \tag{3.18}$$

und

$$\hat{w}_i^{(w)} = y_i - f_i\left(\hat{\beta}^{(w)}, \hat{b}_i^{(w)}\right) + \hat{X}_i^{(w)}\hat{\beta}^{(w)} + \hat{Z}_i^{(w)}\hat{b}_i^{(w)} \tag{3.19}$$

ist

$$\begin{aligned} l_{LME}\left(\beta, \sigma^2, \Delta | y\right) = -\frac{N}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2}\sum_{i=1}^{m} \Big\{ &\log|\Sigma_i(\Delta)| \\ &+\sigma^{-2}\left[\hat{w}_i^{(w)} - \hat{X}_i^{(w)}\beta\right]^T \Sigma_i^{-1}(\Delta)\left[\hat{w}_i^{(w)} - \hat{X}_i^{(w)}\beta\right]\Big\} \end{aligned} \tag{3.20}$$

die Näherungsfunktion für die Log-Likelihood zur Schätzung von Δ. Dabei ist $\Sigma_i(\Delta) = I + \hat{Z}_i^{(w)} \Delta^{-1} \Delta^{-T} \hat{Z}_i^{(w)^T}$. Diese Log-Likelihood ist identisch zu der eines linearen gemischten Modells, nur hier ist die Response $\hat{w}^{(w)}$ (Pinheiro und Bates (2002), Kap. 2.2). Der Algorithmus wechselt zwischen den beiden Schritten, bis ein Konvergenzkriterium eingetroffen ist.

Zur vollständigen Darstellung der eigentlichen Parameterschätzmethoden wird die Methode der Integral-Approximation an dieser Stelle erwähnt. Diese Methode ist in `Proc NLMIXED` aus `SAS 9.1` mit der adaptiven Gauß'schen Quadratur implementiert. Dabei wird ein bestimmtes Integral über vorgegebene Abszissenwerte der Zufallseffekte durch gewichtete Summen approximiert. Die Gewichte sind tabelliert oder algorithmisch generiert. Die Zufallsparameterwerte stammen von einer Zufallsstichprobe aus ihrer Normalverteilungsdichte $N[0, D_i]$. Details zu dieser Methode sind bei Pinheiro und Bates (2002, S. 319) nachzulesen. Für diese Studie wurden Versuche mit der Methode der adaptiven Gauß'schen Quadratur unternommen. Da es nach mehrtägigen Iterationen kein Anzeichen für eine erfolgreiche Optimierung gab, wurde diese Methode nicht weiter verwendet. Schabenberger und Pierce (2002, S. 536) stellen die hohe Genauigkeit dieser Methode heraus, führen aber auch an, dass sie bezüglich ihres Ressourcenverbrauchs sehr teuer sein kann und dies besonders bei einer hohen Zahl von Zufallsparametern.

3.2.4 Vorhersage der Zufallsparameter

Obwohl sich die Zufallseffekte wie Parameter verhalten können, sind sie formal gesehen neben dem Restfehlerterm des Modells nur weitere Ebenen mit Zufallsvariation. Deshalb werden die Zufallsparameter formal gesehen nicht geschätzt, sondern vorhergesagt (Pinheiro und Bates, 2002, S. 8). Ein Ziel der Modellbildung in dieser Arbeit, ist die treffgenaue Vorhersage von Baumhöhen aus Fortschreibungen (Hypothese 2). Dabei soll das Modell nicht nur die Baumhöhen aus dem Modelldatensatz vorhersagen können, sondern auch die Höhenbeobachtungen von anderen Individuen. Die Schätzung der Parameter für die fixen Effekte und die Schätzung der Varianzen der Zufallsparameter ist die eigentliche Modellanpassung und wird auch als Fit bezeichnet. Die Anpassung des Modells an die Individuen

über die Vorhersage der Zufallsparameter wird im folgenden nach dem Verständnis von Lappi und Bailey (1988) als Kalibrierung bezeichnet. Wie bereits dargestellt wurde, ist die Modellanpassung kompliziert und stellt wegen der numerischen Maximierung der Likelihood hohe Anforderungen an die Rechnerleistungen. Die Kalibrierung mittels einfacher Matrixalgebra ist hingegen effizient und sparsam.

Die Daten, die der Modellanpassung zugrunde lagen, werden als Modelldaten bezeichnet und die neuen Daten, anhand derer eine Kalibrierung stattfindet, bezeichnen wir als Kalibrierungsdaten.

Wir nehmen an, dass die Parameterwerte für die fixen Effekte in β und die Kovarianz-Varianz-Matrizen der Zufallseffekte in D_i aus einer vorangegangenen Modellanpassung bekannt seien. Wenn an Probekreisen Höhen y_i^* für Bäume fortgeschrieben werden sollen, die nicht zur Modellanpassung verwendet wurden, dann nehmen wir die Gültigkeit der Schätzungen jener Parameter für die Responsevektoren von neuen Daten an. Für den neuen Plot i mit seinen n_i Bäumen ergibt sich somit das Tripel $(y_i^*, \theta_i^*, A_i^*)$. Für eine individuelle Baumhöhenfortschreibung müssen die Werte der Zufallsparameter auf allen Hierarchieebenen vorhergesagt werden. Denn wie später gezeigt wird, verursachen die Zufallseffekte trotz aller Anstrengungen, fixe Effekte zur Beschreibung der Inter- und Intra-Cluster-Variation zu finden, immer noch den größeren Teil der gesamten Variation.

Bei den Modelldaten ergibt sich durch eine Expansion von $f_i(\theta_i, A_i, \beta, b_i)$ um $b_i = 0$ die Mittelwertkurve entsprechend Gleichung 3.13 annähernd mit:

$$E[f_i] \approx f_i(\theta_i, A_i, \beta, 0) \tag{3.21}$$

Das Residuum $e_i = y_i - f_i(\theta_i, A_i, \beta, 0)$ stellt die gesamte restliche Abweichung der Response vom Mittelwertvektor dar, bei dem lediglich die fixen Effekte eingehen. Die Abweichung in Höhe von e_i wird also von den Zufallseffekten mit b_i und den Restfehlern ϵ_i zusammen verursacht.

Unter der Annahme, dass f_i linear in b_i ist, können mit Hilfe einer Taylor-Entwicklung erster Ordnung um $b_i = 0 = E[b_i]$ für die Matrix
$Z_i = F_i^T B_i = [\partial f_i(\theta_i, A_i, \beta, 0)/\partial\beta] B_i$ und für die Varianzen D_i von $(b_i|y_i, A_i, \beta)$ nach

Vonesh und Chinchilli (1996, S. 362) die Werte der Zufallsparameter erwartet werden mit:

$$E\left[b_i|y_i,\beta,\theta_i,A_i\right] = D_i Z_i^T \left(Z_i D_i Z_i^T + R_i\right)^{-1} e_i \tag{3.22}$$

Eigentlich sind die Parameter der fixen Effekte für die Vektoren y_i^* der Kalibrierungsdaten nicht bekannt. Deshalb verwenden wir für ι, κ und λ die Schätzungen $\hat{\iota}$, $\hat{\kappa}$ und $\hat{\lambda}$ in $\hat{\beta}$ aus der Modellanpassung.

Die Residuen um die Mittelwertkurven der Kalibrierungsdaten können approximiert werden durch:

$$\hat{e}_i = y_i^* - f_i\left(\theta_i^*, A_i^*, \hat{\beta}, 0\right) \tag{3.23}$$

Da auch die Varianz der Zufallseffekte bei den Kalibrierungsdaten nicht bekannt ist, verwenden wir $\hat{D}_i = diag\left(\hat{\Delta}_1, \hat{\Delta}_2, \ldots, \hat{\Delta}_2\right)$ als Schätzer. Die Matrix Z_i wird durch $\hat{Z}_i = \hat{F}_i^T B_i$ mit $\hat{F}_i = \partial f_i\left(\theta_i^*, A_i^*, \hat{\beta}\right) / \partial\left(\hat{\beta}\right)$ geschätzt.

Aus den Gleichungen 3.22 und 3.23 ergibt sich der Prädiktor für die Zufallsparameter b_i am i-ten Plot der *Kalibrierungsdaten* mit:

$$\hat{b}_i \simeq \hat{D}_i \hat{Z}_i^T \left(\hat{Z}_i \hat{D}_i \hat{Z}_i^T + \hat{R}_i\right)^{-1} \hat{e}_i \tag{3.24}$$

Nach Davidian und Giltinan (1995, S. 78) ist dieser Prädiktor ein BLUP (best linear unbiased predictor) für b_i. Dieser Prädiktor hat einen Bayes'schen Charakter als gewichtetes Mittel von $E\left[b_i\right] = 0$ und gewichteten Kleinstquadratschätzern, die aus der Behandlung von b_i als fixe Effekte resultieren. Anschaulich gesprochen, wird die approximierte Gesamtabweichung von der Mittelwertkurve auf die Zufallseffekte und den Restfehler aufgeteilt. Die von den Zufallseffekten verursachte Abweichung wird außerdem auf die einzelnen Hierarchieebenen verteilt.

Die Varianz des Vorhersagefehlers der Zufallsparameter kann für den j-ten Baum am i-ten Plot nach Vonesh und Chinchilli (1996, S. 363) approximiert werden durch:

$$\begin{aligned} V_{ij} = Var\left(b_{ij} - \hat{b}_{ij}\right) \simeq \\ \hat{D}_{ij} - \hat{D}_{ij}\hat{Z}_{ij}^T \Sigma_{ij}\left(\hat{\beta}_{ij}, \theta_{ij}^*, A_{ij}^*\right)^{-1} \hat{Z}_{ij}\hat{D}_{ij} + \\ \hat{D}_{ij}\hat{Z}_{ij}^T \Sigma_{ij}\left(\hat{\beta}_{ij}, \theta_{ij}^*, A_{ij}^*\right)^{-1} \hat{F}_{ij}^T \Omega_{ij}\left(\hat{\beta}_{ij}\right) \hat{F}_{ij} \Sigma_{ij}\left(\hat{\beta}_{ij}, \theta_{ij}^*, A_{ij}^*\right)^{-1} \hat{Z}_{ij}\hat{D}_{ij} \end{aligned} \tag{3.25}$$

Dabei sind

$$\begin{aligned}
&\hat{b}_{ij} = \left(\hat{\zeta}_i^T, \hat{\zeta}_{ij}^T\right)^T \quad , \quad \hat{\beta}_{ij} = \left(\hat{\iota}^T, \hat{\kappa}^T, \hat{\lambda}^T\right)^T \\
&\hat{F}_{ij} = \partial f\left(\theta_{ij}^*, A_{ij}^*, \hat{\beta}_{ij}\right) / \partial \hat{\beta}_{ij} \quad \hat{Z}_{ij} = \hat{F}_{ij}^T B_{ij} \\
&\hat{D}_{ij} = \begin{bmatrix} \hat{\Delta}_1 & 0 \\ 0 & \hat{\Delta}_2 \end{bmatrix} \\
&\Sigma_{ij}\left(\hat{\beta}_{ij}, \theta_{ij}^*, A_{ij}^*\right) = \hat{Z}_{ij}\hat{D}_{ij}\hat{Z}_{ij}^T + \hat{R}_{ij} \ , \quad \hat{R}_{ij} = \hat{\sigma}^2 I \\
&\Omega_{ij}\left(\hat{\beta}_{ij}\right) = A_{ij}^* \Omega A_{ij}^{*T}
\end{aligned} \tag{3.26}$$

A_{ij}^* und B_{ij} sind Design-Matrizen für die Baumebene. Ω ist die geschätzte Kovarianzmatrix für die Schätzer der Parameterwerte der fixen Effekte. Vonesh und Chinchilli (1996) ließen in ihrer Darstellung $\Sigma_{ij}\left(\hat{\beta}_{ij}, \theta_{ij}^*, A_{ij}^*\right)^{-1}$ im rechten Summanden von Gleichung 3.25 aus. Dieses wurde von Fang und Bailey (2001) übernommen.

Nach Vorhersage der Zufallsparameter durch Gleichung 3.24 ergibt sich der Vektor der Vorhersagewerte für die Responsevariable Baumhöhe zu beliebigen Baumalterswerten in θ_i schließlich mit:

$$\hat{y}_i = f_i\left(\theta_i, A_i^*\hat{\beta} + B_i\hat{b}_i\right) \tag{3.27}$$

Die Varianz des Vorhersagefehlers $\hat{y}_{ijk} - y_{ijk}$ für eine Baumhöhenschätzung zum Alter t_{ijk} kann approximiert werden durch:

$$Var\left(\hat{y}_{ijk} - y_{ijk}\right) \simeq \hat{F}_{ijk}^T \Omega_{ij}\left(\hat{\beta}_{ij}\right)\hat{F}_{ijk} + \hat{Z}_{ijk}V_{ij}\hat{Z}_{ijk}^T + \hat{F}_{ijk}^T C_{ij}\hat{Z}_{ijk}^T + \hat{Z}_{ijk}C_{ij}^T\hat{F}_{ijk} + \hat{\sigma}^2 \tag{3.28}$$

Hierbei ist
$C_{ij} = Cov\left[\left(\hat{\beta}_{ij} - \beta_{ij}\right), \left(\hat{b}_{ij} - b_{ij}\right)\right] \simeq -\Omega_{ij}\left(\hat{\beta}_{ij}\right)\hat{F}_{ij}\left(\hat{\beta}_{ij}\right)\Sigma_{ij}\left(\hat{\beta}_{ij}, \theta_{ij}^*, A_{ij}^*\right)^{-1}\hat{Z}_{ij}\hat{D}_{ij}$ und $\hat{F}_{ijk} = \partial f\left(\hat{\phi}_{ij}, t_{ijk}\right) / \partial \hat{\phi}_{ij}$ mit $\hat{\phi}_{ij} = A_{ij}^*\hat{\beta}_{ij} + B_{ij}\hat{b}_{ij}$. Mit der Varianz des Vorhersagefehlers aus Gleichung 3.28 können für die Baumhöhenschätzungen $100\,(1-\alpha)\,\%$- Vertrauensbereiche angegeben werden über:

$$\hat{y}_{ijk} \pm t\left(df, 1-\alpha/2\right)\sqrt{\hat{Var}\left(\hat{y}_{ijk} - y_{ijk}\right)} \tag{3.29}$$

Kapitel 4

Ergebnisse

4.1 Modellkonstruktion

4.1.1 Einstufiges Zufallsmodell

Die Erstellung des endgültigen Modells verlief in mehreren Schritten. Dabei bewährte es sich, vorab die Zufallsparameter zu gestalten und danach die von Kovariablen abhängigen fixen Effekte zu identifizieren. Deshalb wurden zunächst folgende Fragestellungen untersucht:

1. Welche Funktionsparameter sollen Zufallseffekte enthalten?

2. Für welche Gruppierungsebenen sollen Zufallseffekte zugelassen werden?

Die Parameter des Modells sollen gerade genügende und nicht mehr als nötige Zufallskomponenten enthalten, damit das Modell flexibel genug ist, um die gesamte Variation zu erklären. Dieses gebietet sich nicht nur aus Gründen der Sparsamkeit und Überschaubarkeit, sondern hilft auch ein unlogisches Fehlverhalten der Kurvenverläufe und Konvergenzprobleme der numerischen Optimierung zu vermeiden. Bei den vorliegenden Daten sind nicht auf jeder Probefläche die gleiche Anzahl von Bäumen vermessen worden. Wenn in einzelnen Clustern unterschiedliche Anzahlen von Elementen enthalten sind, dann heißen die Daten unbalanciert. Wie im Abschnitt 2.1 ausgeführt wurde, sind für die Studie

Mitscherlichs an grundsätzlich 10 Bäumen die Höhentrieblängen erfasst worden, bei den IWEs 1983, 1988 und 1994 dagegen immer nur an zwei Bäumen pro Inventurtermin und bei der IWE 2001 nur an einem. Aber selbst beim Datenbestandteil von Mitscherlich sind nicht von jedem Plot tatsächlich 10 Bäume in die Studie einbezogen worden. Für manche Bäume waren von den damaligen Bearbeitern unplausible Messwerte angemerkt worden. Teilweise konnten offensichtliche Messfehler aus den Feldaufnahmen vorliegen, die bei der Nachbereitung der Daten oder den Stammscheibenanalysen im Labor nicht mehr korrigiert werden konnten. Einige wenige Bäume wurden nach eingehenden visuellen Prüfungen eigenhändig aus dem Modelldatensatz entfernt. Messfehler bei Stammanalysen können hauptsächlich von Kronenbrüchen herrühren oder wenn für einen Höhentrieb keine Merkmale erkennbar sind. In beiden Fällen korrespondieren die Beobachtungswerte für das Alter aus der Stammscheibe und die der Baumhöhe aus der Differenz der Endhöhe zum Fällzeitpunkt und den kumulierten Trieblängen nicht miteinander. Wenn der Kronenbruch längere Zeit zurückliegt und dabei insbesondere nur wenige Triebe verloren gingen, können die äußeren Merkmale des Schadens mit der Zeit überwachsen werden. Der jeweilige Anfang eines neuen bzw. das jeweilige Ende eines vorherigen Höhentriebes wird durch die Spuren von Knospenschuppen an der Triebbasis gekennzeichnet. Bei lang zurückliegenden Höhentrieben besteht die Gefahr, dass diese nicht mehr sichtbar sind oder wegen ihrer Unscheinbarkeit übersehen werden können. Bei der Erhebung von Mitscherlich wurden im Zweifelsfall orientierende Querschnitte entnommen; anhand der Jahrringzahl konnte die Baumhöhe mit dem Baumalter synchronisiert werden. Dadurch konnten zwar die Fälle von Messfehlern vermindert, aber nicht gänzlich vermieden werden.

Ein Ziel der Modellbildung ist die Aufklärung der Variation innerhalb einer Gruppe von Bäumen an einem Probekreis. Bei der vorliegenden großen Menge an Daten kann ein mehrstufiges Modell viel Rechenzeit beanspruchen. Deshalb wurde zunächst ungeachtet der eigentlich mehrstufig gruppierten Datenstruktur ein einstufiges Modell formuliert. Zunächst wird zudem postuliert, dass sämtliche Variation von den unbekannten Zufallseffekten und dem Restfehler bewirkt wird. Da mit Gleichung 3.9 angenommen wird, dass die Erwartungswerte der Zufallsparameter gleich 0 sind, ist es unüblich die Funktions-

parameter nur aus Zufallsparametern zu formulieren (Pinheiro und Bates, 2002, S. 355). Zusätzlich zu den Zufallsparametern müssen die Funktionsparameter aus festen Interzepten bestehen, die für die Population konstant sind. Diese sind die Parameter aus dem Vektor ι in Gleichung 3.9. Obwohl diese Parameter streng genommen von fixen Effekten hervorgerufen werden, und zwar von solchen auf der Ebene der Population, soll das Modell aus diesem Konstruktionsschritt ein einstufiges Zufallsmodell genannt werden, da sämtliche Variation von Zufallseffekten verursacht wird.

Mit den Vereinfachungen, dass nur eine Hierarchieebene vorherrscht und jegliche Variation durch den Zufall bedingt ist, konnte leichter der Frage nachgegangen werden, welche Funktionsparameter Zufallseffekte beinhalten sollen. Dabei galt es, für das Modell einen Kompromiss zwischen Flexibilität auf der einen Seite sowie Stabilität und Plausibilität auf der anderen Seite zu finden. Das Modell sollte flexibel genug sein, die individuellen Kurvenformen jedes einzelnen Baumes zu beschreiben und zugleich keine unlogischen Anpassungsergebnisse auftreten lassen. Außerdem mussten in der Anfangsphase geeignete Startwerte für die Funktionsparameter gesucht werden. Zur Untersuchung dieser Fragestellungen reichte ein einstufiges Modell aus.

Das einstufige Modell ergibt sich aus Vereinfachung des zweistufigen Modells aus 3.8. Danach erhalten wir die n_i Höhenbeobachtungen am i-ten von m Bäumen mit:

$$y_i = f_i\left(\phi_i, \theta_i\right) + \epsilon_i \tag{4.1}$$

$$y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{in_i} \end{bmatrix}, \quad f_i\left(\phi_i, \theta_i\right) = \begin{bmatrix} f\left(\phi_i, t_{i1}\right) \\ \vdots \\ f\left(\phi_i, t_{in_i}\right) \end{bmatrix}, \quad \epsilon_i = \begin{bmatrix} \epsilon_{i1} \\ \vdots \\ \epsilon_{in_i} \end{bmatrix}$$

$$\phi_i = \begin{bmatrix} \psi_i^{(1)} \\ \vdots \\ \psi_i^{(p)} \end{bmatrix}, \quad \theta_i = \begin{bmatrix} t_{i1} \\ \vdots \\ t_{in_i} \end{bmatrix}, \quad \epsilon_i \sim N\left[0, R_i\right] \quad R_i = \sigma^2 I\left[n_i\right]$$

Der Parametervektor ergibt sich entsprechend aus:

$$\begin{aligned}
&\phi_i = A_i\beta + B_i b_i \qquad (4.2)\\
&\beta = \iota\\
&b_i = \zeta_i\\
&\zeta_i = \left(z_i^{(1)}, \ldots, z_i^{(p)}\right)^T\\
&b_i \sim N\left(0, D_i\right)\\
&D_i = \begin{bmatrix} Var\left[z_i^{(1)}\right] & \cdots & Cov\left[z_i^{(1)}, z_i^{(p)}\right] \\ \vdots & \ddots & \vdots \\ Cov\left[z_i^{(1)}, z_i^{(p)}\right] & \cdots & Var\left[z_i^{(p)}\right] \end{bmatrix}
\end{aligned}$$

In Gleichung 3.2 wurde der Parameter $\psi_{ij}^{(4)}$ als Erwartungswert für die Baumhöhe in einem festen Alter t_0 formuliert. In einem Site-Index-Modell würde man diesen Parameter als Bonitätsparameter für das feste Referenzalter t_0 bezeichnen. Da in dieser Studie nicht der Verlauf von Bestandesoberhöhen beschrieben werden soll, sondern das Wachstum von einzelnen Bäumen, stellt $\psi_{ij}^{(4)}$ gewissermaßen den *individuellen* Höhenerwartungswert im Alter t_0 dar. Das Referenzalter t_0 muss einen festen Wert annehmen. Wenn für t_0 sehr große oder sehr kleine Alterswerte gewählt wurden, kam es zu Konvergenzschwierigkeiten bei der Optimierung im Zuge der Parameterschätzung; deshalb wurde t_0 auf 50 Jahre festgesetzt.

Damit das Modell ein vielfältiges Höhenwachstum darstellen kann, sollte zumindest dieser Erwartungswertparameter einen Zufallsparameter und vielleicht später gemischte Effekte enthalten. Für große Flexibilität wäre es deshalb wünschenswert, wenn zusätzlich mindestens einer der beiden formbestimmenden Parameter $\psi_i^{(2)}$, $\psi_i^{(3)}$ eine Zufallskomponente enthalten würde. Zuerst wurden nur Zufallseffekte auf die zwei Funktionsparameter $\psi_i^{(2)}$ und $\psi_i^{(4)}$ installiert. Der Parameter $\psi_{ij}^{(1)}$ wurde nicht als solcher geschätzt, sondern heuristisch auf 0.98 festgesetzt; andernfalls traten Konvergenzschwierigkeiten auf. In Abbildung 4.1 sind die Höhenmesswerte nach dem Ursprung ihrer Erhebung (Mitscherlich, IWE) und der Lage der Probekreise bezüglich der Region aus der Einteilung im Abschnitt 2.1.3 dargestellt. In Abbildung 4.2 werden die Vorhersagen, die mit dem einstufi-

gen Zufallsmodell erzielt wurden, in entsprechender Weise gezeigt. Im augenscheinlichen Vergleich scheinen alle Modell-Kurven biologisch plausible Verläufe anzunehmen, wenn $\psi_i^{(2)}$ und $\psi_i^{(4)}$ Zufallskomponenten enthalten. Die Residuen zeigen zudem keinen Trend, weder über die Prädiktorvariable Baumalter noch über die Schätzwerte (Abbildung 4.3). Die Modellanpassung scheint gelungen zu sein. Die geschätzten Parameterwerte sind der Spalte „Modell 1“ in Tabelle 8.1 aus dem Anhang zu entnehmen.

An dieser Stelle stellt sich die Frage, ob nicht auch eine Zufallskomponente auf einen einzigen Funktionsparameter des Modells ausreicht. Dazu wurde zunächst ein Modell angepasst, das lediglich einen Zufallsparameter in $\psi_i^{(4)}$ enthält. Die Parameterwerte sind aus Gründen der Übersicht nicht in Tabelle 8.1 dargestellt. Ein Modell, das nur einem Zufallsparameter in $\psi_i^{(2)}$ besitzt ist unlogisch, da angenommen werden würde, dass alle Bäume im Referenzalter gleich hoch wären; dieses Modell wird deshalb nicht angepasst und nicht zum Vergleich herangezogen. Um die Güte der Anpassung zweier Modelle zu vergleichen, die sich in der Spezifikation ihrer Zufallsparameter unterscheiden, bietet sich der Likelihood-Ratio-Test (LRT) an ((Pinheiro und Bates, 2002, S. 83); (Schabenberger und Pierce, 2002, S. 13)). L_u sei die Likelihood des vollen Modells und L_r die eines beschränkten Modells, in dem mindestens ein Parameter aus dem vollen Modell auf 0 gesetzt ist. Im allgemeinen hat die LRT-Statistik

$$\Lambda = 2\log\left(L_u/L_r\right) = 2\left[\log(L_u) - \log(L_r)\right] \tag{4.3}$$

eine asymptotische χ^2-Verteilung mit q Freiheitsgraden. Dabei ergibt sich q aus der Anzahl der Restriktionen, d.h. der Anzahl Parameter, die im restringierten Modell im Vergleich zum vollen Modell auf 0 gesetzt sind.

Die `nlme`-Funktion aus `R` gibt den logarithmierten Likelihood-Wert aus. Das restringierte Modell mit lediglich einem Zufallsparameter in $\psi_i^{(4)}$ hat einen Log-Likelihood-Wert von $\log\left(L_r\right) = -59510.86$ und das volle Modell (Modell 1) mit Zufallsparametern in $\psi_i^{(2)}$ und $\psi_i^{(4)}$ einen von $\log\left(L_u\right) = -35169.35$. Die Prüfgröße ist $\Lambda = 2\log\left(L_u\right) - 2\log\left(L_r\right) = 2 \cdot (59510,86 - 35169,35) = 48683,02$ bei zwei Freiheitsgraden. Das restringierte Modell mit nur einem Zufallsparameter hat fünf Parameter; drei davon sind Interzepte für die

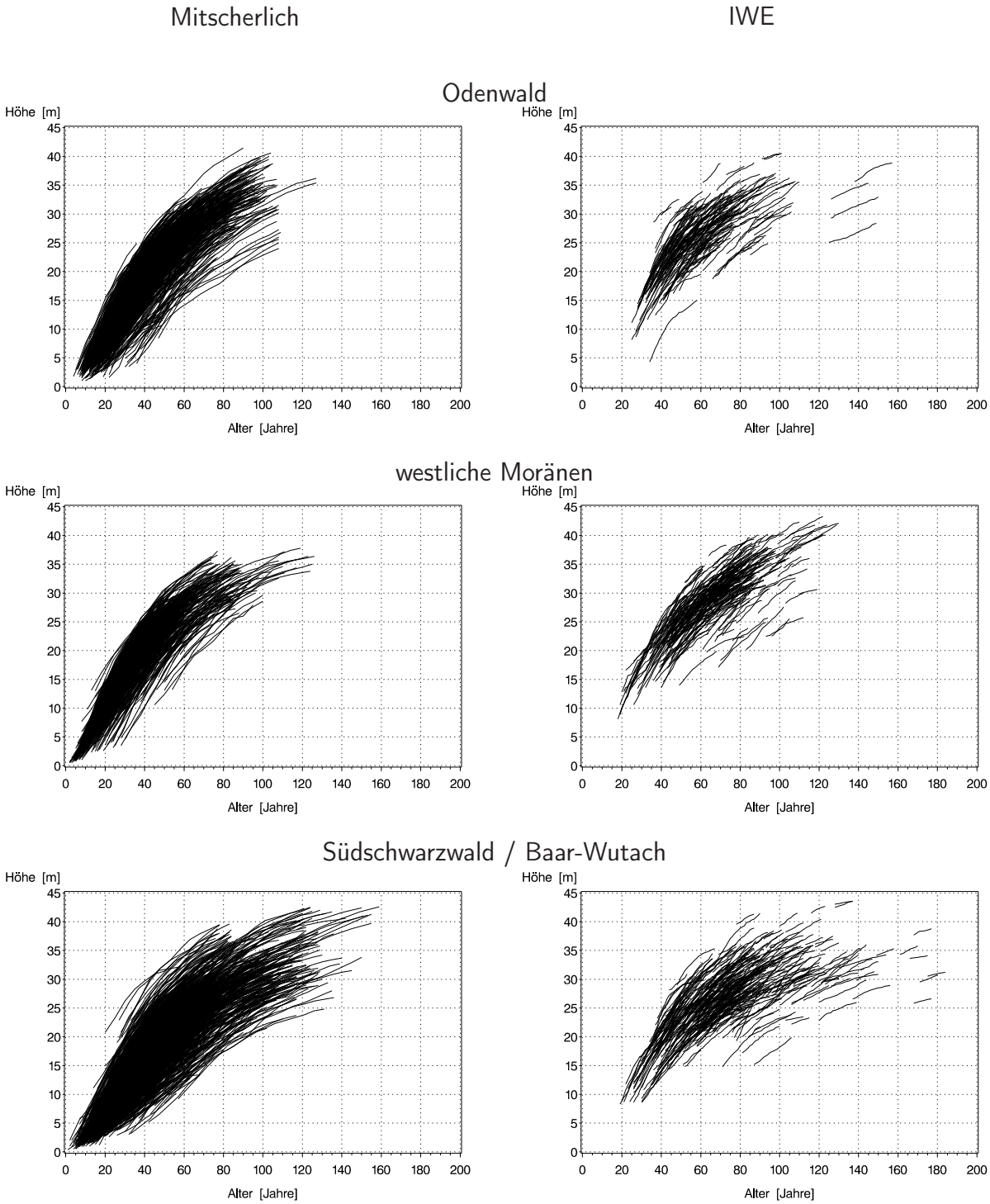


Abbildung 4.1: *Messwerte der Baumhöhen.*

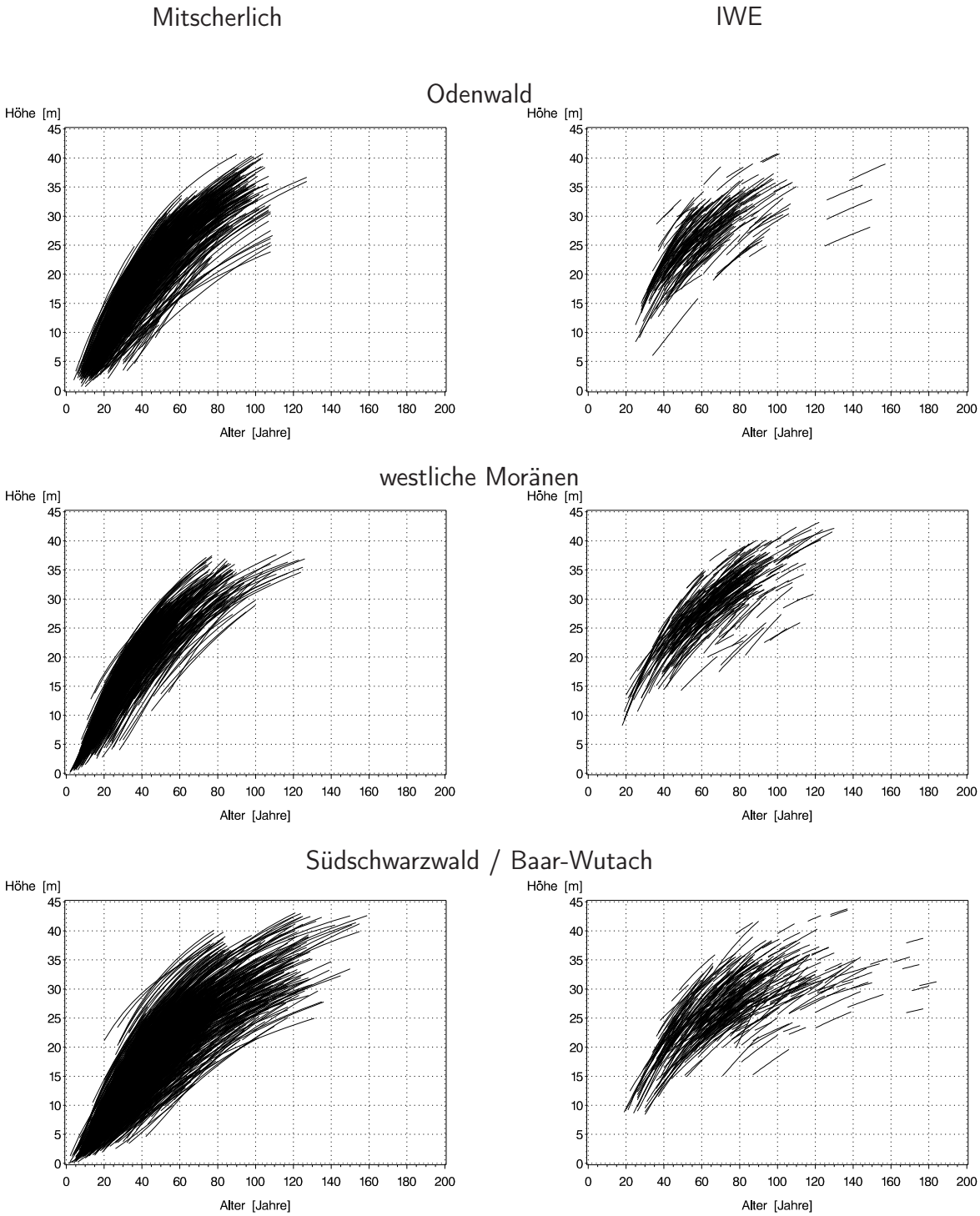


Abbildung 4.2: *Vorhersagen der Baumhöhen aus dem einstufigen Zufallsmodell.*

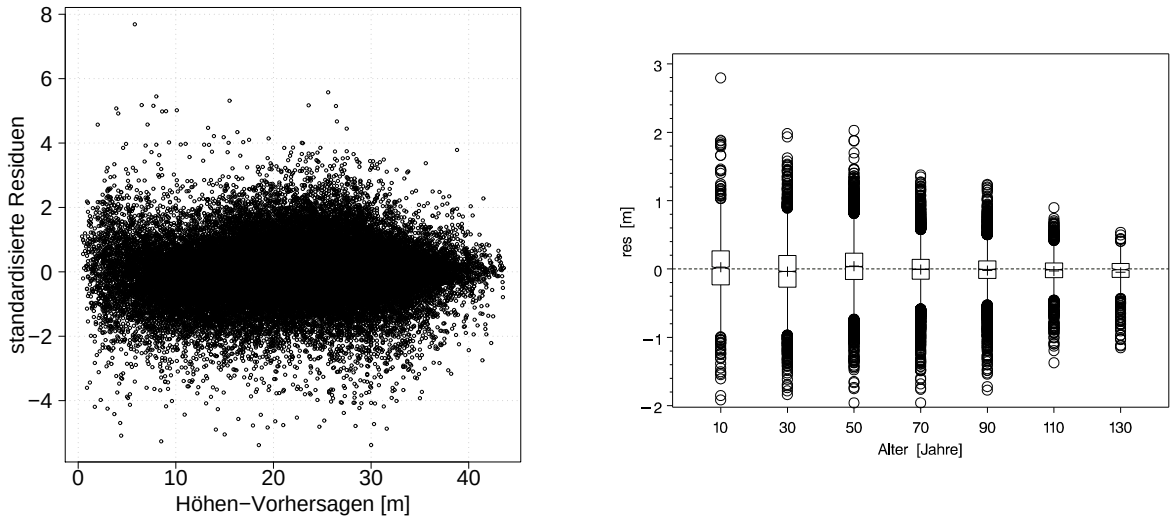


Abbildung 4.3: *Residuen des einstufigen Zufallsmodells. Links: Standardisierte Residuen über den Vorhersagewerten. Rechts: Boxplots der Residuen über der Prädiktorvariablen Baumalter in Altersklassen.*

Populations-fixen-Effekte, einer der Zufallsparameter und einer der Restfehler. Das unrestringierte Modell hat zwei Parameter mehr; mit $z_i^{(2)}$ einen zusätzlichen Zufallsparameter und die Kovarianz $Cov\left[z_i^{(2)}, z_i^{(4)}\right]$ zwischen den beiden Zufallsparametern. Dadurch ergeben sich die zwei Freiheitsgrade für den LRT. Die Quantile der χ^2-Verteilung können aus Tabellen entnommen werden. Die `anova`-Funktion aus `R` liefert jedoch eine komfortable Möglichkeit den LRT durchzuführen.

Modell	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
restringiert	1	5	119031.7	119074.78	-59510.86			
unrestringiert	2	7	70352.7	70412.97	-35169.35	1 vs 2	48683.03	<.0001

Tabelle 4.1: *Likelihood-Ratio-Test mit Hilfe der R-Funktion anova für Modelle mit unterschiedlicher Spezifikation der Zufallsparameter.*

Das reduzierte Modell kann bei einer Irrtumswahrscheinlichkeit von 0,01% verworfen werden. Durch die Einführung des zusätzlichen Zufallsparameters in $\psi_i^{(2)}$ wurde die Modellanpassung folglich signifikant verbessert.

Wenn neben $\psi_i^{(2)}$ und $\psi_i^{(4)}$ außerdem noch Zufallseffekte auf $\psi_i^{(1)}$ oder $\psi_i^{(3)}$ zugelassen wurden, traten bei der `nlme`-Funktion von `R` Konvergenzschwierigkeiten auf und bei

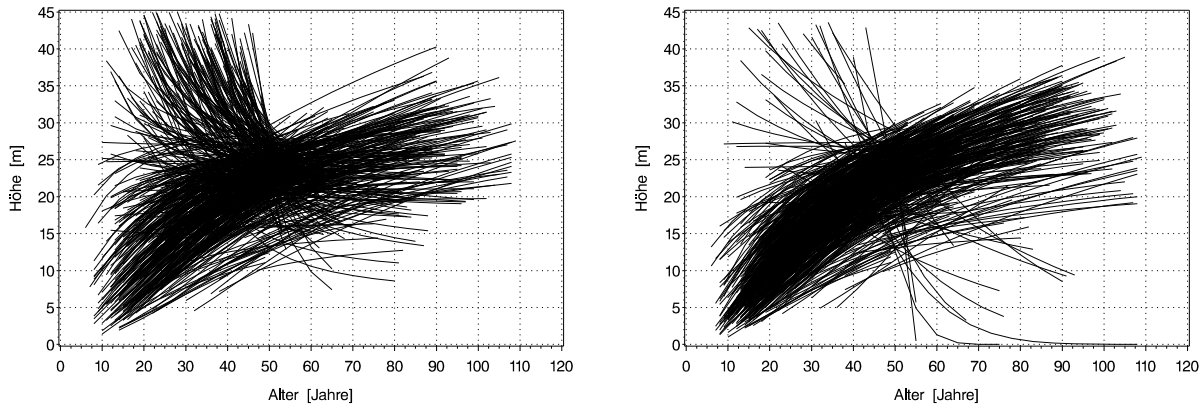


Abbildung 4.4: *Unplausible Kurvenverläufe des einstufigen Zufallsmodells. Modellanpassung mit SAS-Routine Proc NLMIXED. Zur Übersichtlichkeit sind nur die Modellkurven für die Mitscherlich-Daten aus dem Odenwald dargestellt. Links: Zufallseffekte auf* $\psi_i^{(1)}$, $\psi_i^{(3)}$ *und* $\psi_i^{(4)}$. *Rechts: Zufallseffekte auf* $\psi_i^{(2)}$, $\psi_i^{(3)}$ *und* $\psi_i^{(4)}$.

der Routine `Proc NLMIXED` von `SAS` ergaben sich unplausible Kurvenverläufe. Die unplausiblen Kurvenverläufe sind exemplarisch für die Mitscherlich-Daten aus der Region Odenwald in Abbildung 4.4 dargestellt. Das Modell scheint mit Zufallskomponenten in drei Parametern überparametrisiert zu sein. Aus diesem Grund werden für die weitere Modellkonstruktion Zufallsparameter nur auf $\psi_i^{(2)}$ und $\psi_i^{(4)}$ zugelassen. Damit wäre die erste Frage beantwortet, welche Funktionsparameter Zufallseffekte enthalten sollen.

Die Konstruktion des einstufigen Zufallsmodells (Modell 1) ist schließlich:

$$
\begin{aligned}
\phi_i &= \left(\psi_i^{(2)}, \psi_i^{(3)}, \psi_i^{(4)}\right)^T \\
\psi_i^{(2)} &= \iota^{(2)} + z_i^{(2)} \\
\psi_i^{(3)} &= \iota^{(3)} \\
\psi_i^{(4)} &= \iota^{(4)} + z_i^{(4)} \\
z_i^{(2)} &\sim N\left[0, var\left(z_i^{(2)}\right)\right] \qquad z_i^{(4)} \sim N\left[0, var\left(z_i^{(4)}\right)\right] \\
\Delta_1 &= \begin{bmatrix} Var\left[z_i^{(2)}\right] & Cov\left[z_i^{(2)}, z_i^{(4)}\right] \\ Cov\left[z_i^{(2)}, z_i^{(4)}\right] & Var\left[z_i^{(4)}\right] \end{bmatrix}
\end{aligned}
\tag{4.4}
$$

Die geschätzten Parameterwerte, sowie deren Standardabweichungen und p-Werte können aus der Spalte „Modell 1“ der Tabelle 8.1 des Anhangs entnommen werden.

4.1.2 Zweistufiges Zufallsmodell

Im einstufigen Modell haben sich Zufallskomponenten auf den Parametern $\psi_i^{(2)}$ und $\psi_i^{(4)}$ bewährt. Nun soll der zweiten Frage nachgegangen werden, auf welchen Hierarchieebenen diese Parameter Zufallseffekte enthalten sollen. Mit Zufallseffekten auf den Ebenen der Probekreise würde beschrieben werden, wie die mittleren Responsekurven für die Probeflächen voneinander abweichen. Zufallseffekte auf der darunterliegenden Ebene der Bäume würden aufklären, wie die Response-Variable zwischen den Bäumen an einer Probefläche variiert. Mit einem zweistufigen Modell kann außerdem der unbalancierten Datenstruktur Rechnung getragen werden.

Bei den Daten aus der IWE stellte sich die Identifizierung des Merkmals der oberen Gruppierungsebene als schwierige Fragestellung heraus. Die vorliegenden Daten der IWE stammen aus den Jahren 1983, 1988, 1994 und 2001. In der Umgebung eines Stichprobenpunktes wurden zu einem Aufnahmeterminen 1983, 1988 und 1994 jeweils zwei Bäume und im Jahr 2001 jeweils ein Baum zur Untersuchung gefällt. In einzelnen Fällen ist es möglich gewesen, dass die Probebäume an einem Stichprobenpunkt zu unterschiedlichen Aufnahmeterminen aus unterschiedlichen Waldbeständen stammen. Wenn im Waldbestand, in dem die Fällungen der Voraufnahme unternommen wurden, keine geeigneten Kandidaten mehr zur Verfügung standen, wurde auf Bäume im benachbarten Bestand zurückgegriffen. Die Bäume können dann aus Beständen stammen, die zu verschiedenen Zeitpunkten begründet wurden oder unterschiedlich behandelt worden sind. Die mittlere Abbildung 4.5 zeigt exemplarisch die Höhenmessungen an Bäumen eines Stichprobenpunktes der IWE, bei denen die Bäume zu den verschiedenen Inventurterminen wegen ihres Altersunterschiedes offensichtlich aus unterschiedlichen Beständen entnommen worden sein müssen.

Leider wird bei der IWE nicht dokumentiert, ob die bei verschiedenen Aufnahmezeitpunkten gefällten Bäume aus unterschiedlichen Beständen entnommen wurden. Da die Bestände selbst vielerorts ungleichaltrig waren, können sich auch die Bäume aus einem Bestand zu einem Inventurtermin recht stark im Alter unterscheiden. In einigen Fällen, wie z.B. in der unteren Abbildung 4.5, deutet manches darauf hin, dass selbst die Bäume

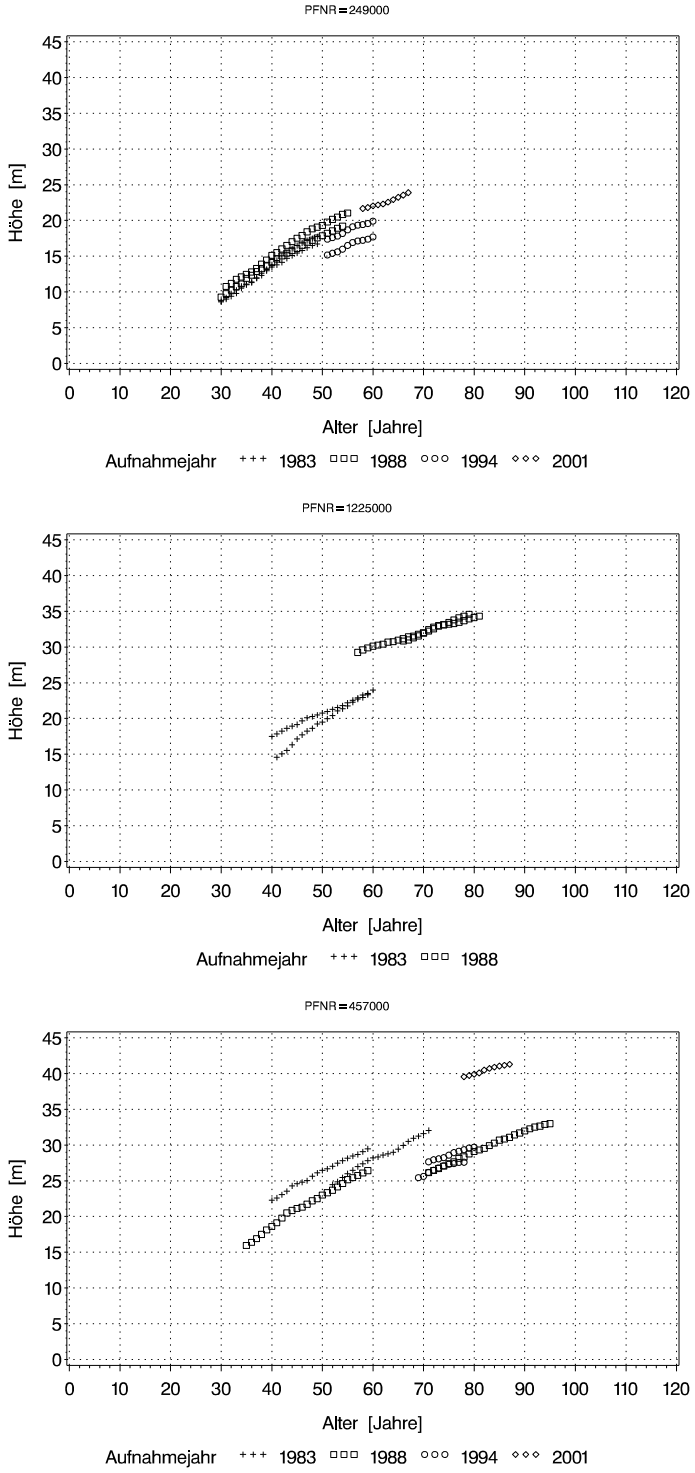


Abbildung 4.5: *Gruppierungsmerkmal für die Bäume der IWE. Oben: Aufnahmen zu vier Zeitpunkten. Alle Bäume stammen offensichtlich aus dem gleichen Bestand. Mitte: Die Bäume der beiden verschiedenen Inventurtermine stammen vermutlich aus verschiedenen Beständen. Unten: Im Jahr 1988 wurden die Bäume vermutlich aus unterschiedlichen Beständen entnommen.*

eines Stichprobenpunktes bei einem Aufnahmetermin aus unterschiedlichen Beständen stammen könnten. Es kann im Nachhinein nicht zweifelsfrei geklärt werden, an welchen Stichproben die Bäume aus unterschiedlichen Beständen kamen oder aus dem gleichen Bestand entnommen wurden.

Weder die Bäume eines Stichprobenpunktes, noch die zu einem Aufnahmezeitpunkt an einem Stichprobenpunkt sind stets mit absoluter Sicherheit eindeutig einem Waldbestand zuzuordnen. Nach visueller Kontrolle aller Baumhöhenbeobachtungen an den einzelnen Stichprobenpunkten stammen diese jedoch in der überwiegenden Mehrheit aller Stichproben vermutlich aus dem selben Bestand. Für die weitere Konstruktion des Modells wird das Merkmal Stichprobenpunkt deshalb als Gruppierungsmerkmal der oberen Ebene betrachtet. Es wird also nicht die annähernd gleiche punktgenaue Lokalisation der Bäume oder die Zugehörigkeit zum selben Bestand als räumliches Gruppenmerkmal angenommen, statt dessen wird die weitere räumliche Umgebung am Stichprobenpunkt als Gruppierungsmerkmal aufgefasst. Es wird damit angenommen, dass für Bäume in der Nähe des selben Punktes in ungünstigen Fällen zumindest ähnliche Wuchsbedingungen hinsichtlich der Höhenlage, des Klimas oder der groben Standortsverhältnisse vorherrschen. Da in der überwiegenden Zahl der Fälle die Bäume an einem IWE-Punkt im selben Bestand standen, liegen in diesen günstigeren Fällen noch weitere Gemeinsamkeiten vor. Wenn nun die Umgebung eines Stichprobenpunktes und nicht der bestimmte Aufnahmezeitpunkt an einem Punkt als Gruppenmerkmal dient, dann können auch die Höhenmesswerte der IWE aus dem Jahre 2001 in die Studie mit einbezogen werden. Bei der IWE im Jahre 2001 wurde an einem Punkt jeweils nur ein Baum gefällt. Mindestens zwei Bäume sind aber nötig, damit überhaupt Variation um die Mittelwertkurve einer Gruppe zu beobachten ist. Indem die Messwerte der Bäume zu allen Inventurzeitpunkten eines Stichprobenpunktes als eine Gruppe begriffen werden, können auch die Höhenmesswerte der IWE 2001 berücksichtigt werden, bei der jeweils nur ein Baum untersucht wurde.

Die Kennung der Probeflächen für die Bäume aus Mitscherlichs Erhebung und die Stichprobennummer der IWE wurden fortan als Gruppierungsmerkmal der oberen Hierarchieebene begriffen. Für die Beschreibung des zweistufigen Modells wird der Notation

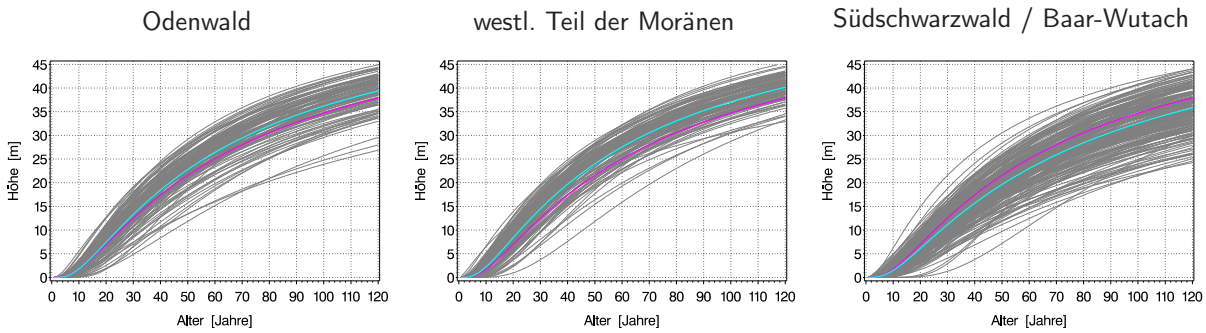


Abbildung 4.6: *Stichproben-Mittelwertkurven (Graue Linien) aus dem 2-Stufen-Zufallsmodell, getrennt für die drei Regionen dargestellt. Magenta-farbene Linien: Populationsmittelwertkurve; Cyan-farbene Linien: Mittelwertkurven für die Regionen, bestehend aus Interzepten für die Population und Mittelwerten der Zufallparameter in den jeweiligen Regionen.*

aus 3.8 gefolgt. Beim einstufigen Modell hatten sich Zufallseffekte auf der individuellen Baum-Ebene in den Parametern $\psi_i^{(2)}$ und $\psi_i^{(4)}$ bewährt. Deshalb wurden für die Parameter $\psi_{ij}^{(2)}$ und $\psi_{ij}^{(4)}$ des zweistufigen Modells auch auf der darüber liegenden Stichproben-Ebene Zufallseffekte angenommen. Für alle Bäume konnten damit plausible Kurvenverläufe geschätzt werden.

Das zweistufige Zufallsmodell (Modell 2) ist wie folgt konstruiert:

$$
\begin{aligned}
\phi_i &= \left(\phi_{i1}^T, \ldots, \phi_{ij}^T, \ldots, \phi_{in_i}^T\right)^T \qquad (4.5)\\
&= \left(\psi_{i1}^{(2)}, \psi_{i1}^{(3)}, \psi_{i1}^{(4)}, \ldots, \psi_{in_i}^{(2)}, \psi_{in_i}^{(3)}, \psi_{in_i}^{(4)}\right)^T\\
\psi_{ij}^{(2)} &= \iota^{(2)} + z_i^{(2)} + z_{ij}^{(2)}\\
\psi_{ij}^{(3)} &= \iota^{(3)}\\
\psi_{ij}^{(4)} &= \iota^{(4)} + z_i^{(4)} + z_{ij}^{(4)}\\
z_i^{(2)} &\sim N\left[0, var\left(z_i^{(2)}\right)\right] \qquad z_i^{(4)} \sim N\left[0, var\left(z_i^{(4)}\right)\right]\\
z_{ij}^{(2)} &\sim N\left[0, var\left(z_{ij}^{(2)}\right)\right] \qquad z_{ij}^{(4)} \sim N\left[0, var\left(z_{ij}^{(4)}\right)\right]\\
\Delta_1 &= \begin{bmatrix} Var\left[z_i^{(2)}\right] & Cov\left[z_i^{(2)}, z_i^{(4)}\right] \\ Cov\left[z_i^{(2)}, z_i^{(4)}\right] & Var\left[z_i^{(4)}\right] \end{bmatrix}\\
\Delta_2 &= \begin{bmatrix} Var\left[z_{ij}^{(2)}\right] & Cov\left[z_{ij}^{(2)}, z_{ij}^{(4)}\right] \\ Cov\left[z_{ij}^{(2)}, z_{ij}^{(4)}\right] & Var\left[z_{ij}^{(4)}\right] \end{bmatrix}
\end{aligned}
$$

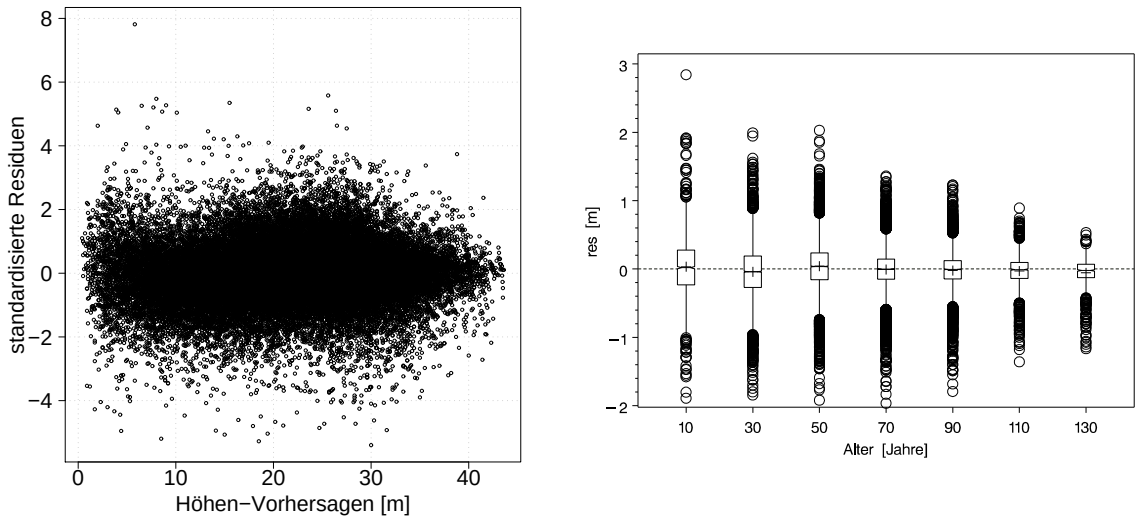


Abbildung 4.7: *Validität des zweistufigen Zufallsmodell. Links: Standardisierte Residuen über den Vorhersagewerten. Rechts: Boxplots der Residuen über der Prädiktorvariablen Baumalter in Altersklassen.*

Abbildung 4.7 deutet auf einen straffen und gleichmäßigen Zusammenhang zwischen den Schätzwerten und den Beobachtungen hin. In Abbildung 4.6 sind die resultierenden Mittelwertkurven für die einzelnen Stichproben getrennt für die drei Regionen dargestellt. Dabei liegt ein zweistufiges Zufallsmodell mit Zufallsparametern für Baum- und Stichproben-Ebene in $\psi_{ij}^{(2)}$ und $\psi_{ij}^{(4)}$ zu Grunde. Das Ausmaß der Variation zwischen den Mittelwertkurven für die einzelnen Probeflächen ist insgesamt recht hoch. Die Varianz des q-ten von den p Parametern ergibt sich wegen Unabhängigkeit der Zufallseffekte über die Ebenen hinweg ($Cov\left[z_i^{(q)}, z_{ij}^{(q)}\right] = 0$) für das zweistufige Zufallsmodell über die Summe der Varianzen der Zufallsparameter auf den beiden Ebenen:

$$Var\left[\psi_{ij}^{(q)}\right] = Var\left[z_i^{(q)} + z_{ij}^{(q)}\right] = Var\left[z_i^{(q)}\right] + Var\left[z_{ij}^{(q)}\right] \tag{4.6}$$

Von Interesse ist jetzt zum einen, wie hoch die Variation auf Ebene der Bäume ist, d.h. wie stark variieren die Bäume eines Stichprobenpunktes untereinander und wie stark weichen sie von der Mittelwertkurve für den Stichprobenpunkt ab, und zum anderen, wie hoch die Variation auf Ebene der Stichprobenpunkte ist; d.h. wie stark unterscheiden sich die

Mittelwertkurven der Punkte voneinander und wie weichen diese von der Mittelwertkurve für die Population ab.

Der %-Anteil der Varianz, der auf die Zufallseffekte einer Ebene, z.B. der oberen Ebene, entfällt ist:

$$\hat{R_i\%} = \frac{Var\left[z_i^{(q)}\right]}{Var\left[z_i^{(q)}\right] + Var\left[z_{ij}^{(q)}\right]} \cdot 100 \tag{4.7}$$

Bei dem 2-Stufen-Zufallsmodell wird durch die Zufallseffekte auf der Ebene Probekreise rund $3,9998^2/\left(3,9998^2 + 2,0191^2\right) \cdot 100\% = 79,7\%$ der gesamten Zufallsparameter-Varianz in $\psi_{ij}^{(4)}$ hervorgerufen und in Höhe von
$0,0759^2/\left(0,0759^2 + 0,0630^2\right) \cdot 100\% = 54,6\%$ in $\psi_{ij}^{(2)}$. Der Rest der Varianz wird durch Zufallseffekte auf der Baumebene hervorgerufen. Das bedeutet die Mittelwertkurven der unterschiedlichen Probekreise unterscheiden sich stärker voneinander als die individuellen Kurven an einem Probekreis. Das Wachstum von Bäumen unterscheidet sich zwischen Waldbeständen stärker als innerhalb eines Waldbestandes. Die Merkmale, die an einem Probekreis für alle Bäume gleich sind, wie z.B. genetische Herkunft, Standort oder Klima bewirken offensichtlich höhere Unterschiede im Baumwachstum als Merkmale, die von Baum zu Baum am selben Probekreis differieren können, wie z.B. individuelle Konkurrenzvorteile. Mit dem Wissen darüber, wie die Variation anteilsmäßig den einzelnen Hierarchieebenen zuzuordnen ist, kann der weitere Modellbildungsprozess effizienter gestaltet werden. So sollten verstärkt Bemühungen investiert werden, die Variation auf der oberen Ebene der Stichprobenpunkte mit Hilfe von Kovariablen als fixe Effekte zu beschreiben.

In Grafik 4.6 ist neben den Stichproben-Mittelwertkurven noch die Populationsmittelwertkurve (magenta) ohne Zufallsparameter mit lediglich fixen Parametereffekten eingezeichnet. Außerdem sind cyan-farbene Kurven dargestellt, bei denen zu den festen Interzepten die Mittelwerte der Zufallsparameter in den jeweiligen Regionen hinzu addiert wurden. Diese verdeutlichen den durchschnittlichen Wachstumsverlauf in den einzelnen Regionen. Diese regionalen Mittelwertkurven variieren in den drei Regionen auf unterschiedlichen Niveaus um die Populationsmittelwertkurve (magenta). Die Fichten-Baumhöhen aus dem Odenwald (Region 1) und auf dem westlichen Teil der Moränen

(Region 2) scheinen bei gleichem Alter höher zu sein als die im Südschwarzwald / Baar-Wutach (Region 3). Es hat zudem den Anschein, dass die Fichten auf den Moränen ein schnelleres Jugendwachstum haben als im Odenwald.

In Abbildung 4.8 sind die Höhenmesswerte aller Bäume an zwei Stichprobenpunkten dargestellt. Es wird hier die hohe Variationsbreite verdeutlicht, die das Höhenwachstum der Fichte an einem Ort aufweisen kann. Mit einer roten Linie ist jeweils die geschätzte mittlere Höhenwachstumskurve für diese Stichprobenpunkte eingezeichnet. Die Funktionsparameter dieser Kurven enthalten die fixen Effekte auf Populations-Ebene ι und die über BLUPs nach Gleichung 3.24 geschätzten Zufallseffekte auf der Stichproben-Ebene. Die schwarzen Linien stellen die Ausgleichskurven für die Höhenmesswerte der einzelnen Bäume dar. Bei diesen Kurven bestehen die Funktionsparameter aus den fixen Effekten auf Populations-Ebene und den vorhergesagten Zufallsparametern auf beiden Ebenen, der Stichproben- und der Baumebene. Die Höhenmesswerte des Baumes mit der Nr. 3 am Stichprobenpunkt mit der Nr. 202 (s. Abb. 4.8 unten) lagen in jungen Jahren unter denen des Baumes mit der Nummer 2 und ab einem Baumalter von 80 Jahren kehrte sich diese Situation um. Der Baum mit der Nummer 2 wurde also im Laufe der Zeit vom Baum mit der Nummer 3 in seiner Höhe überholt. Bemerkenswerter Weise kann dieser Sachverhalt durch das auf Zufallsparametern basierende Modell nachgebildet werden.

4.1.3 Zweistufiges gemischtes Modell

4.1.3.1 Überlegungen zur Beschreibung räumlicher Variation

In den beiden bisher dargestellten Zufallsmodellen repräsentieren die Zufallseffekte die Abweichung der Baum-individuellen Funktionsparameter von den festen Parameterwerten für den Populationsdurchschnitt. Durch die weitere Modellkonstruktion soll untersucht werden, ob sich diese Abweichungen zwischen den Datengruppen der Stichproben und zwischen den Gruppen der Bäume in einer Stichprobe erklären lassen. Zur Erklärung könnten unterschiedliche Werte von weiteren beobachtbaren Variablen (Kovariablen) dienen. Im günstigen Fall korrespondieren die Abweichungen der einzelnen

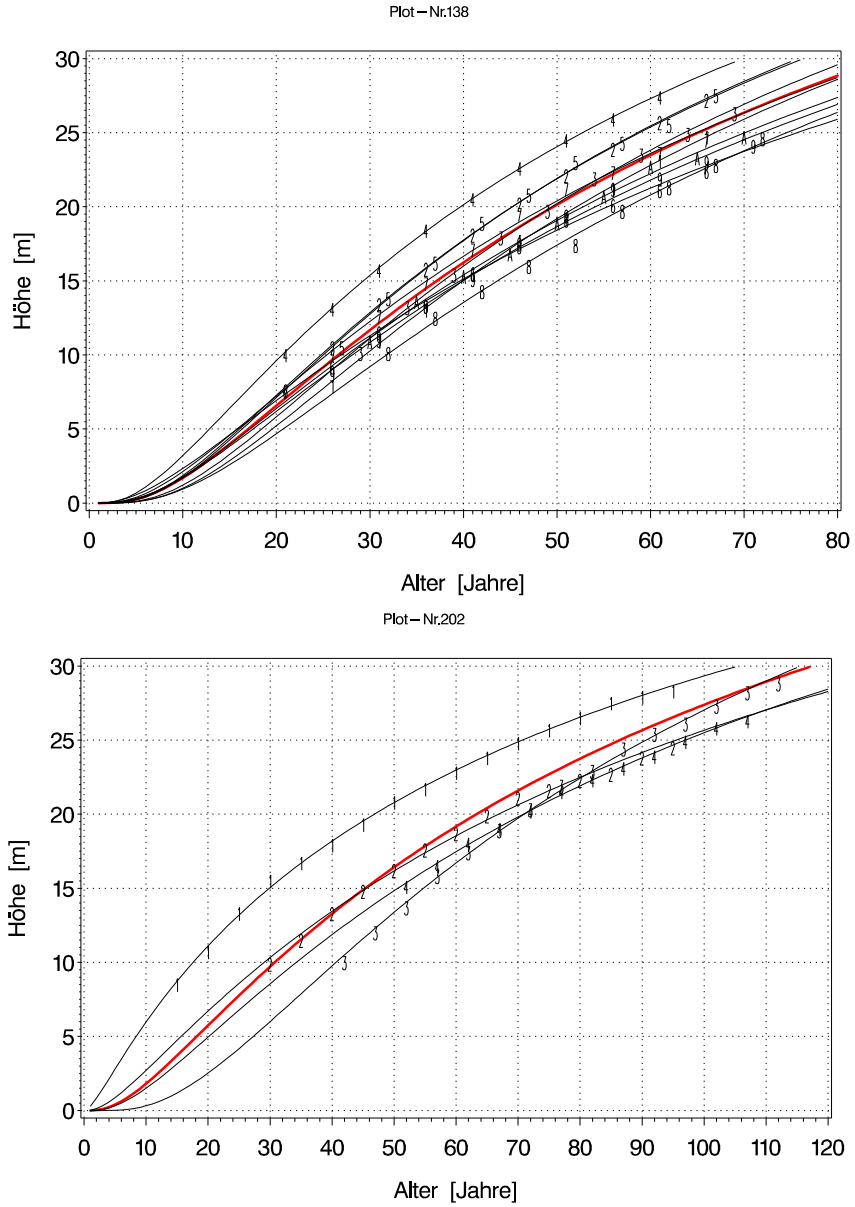


Abbildung 4.8: *Variation zwischen Bäumen. Darstellung der Variation zwischen den Bäumen an zwei ausgewählten Stichprobenpunkten. Rote Linie: Mittelwertkurve für den Stichprobenpunkt mit vorhergesagten Zufallsparametern auf der oberen Stichproben-Ebene. Schwarze Ziffern: Höhenmesswerte der einzelnen Bäume. Schwarze Linien: Individuelle Kurven mit vorhergesagten Zufallsparametern auf Stichproben- und Baum-Ebene.*

Stichproben-Mittelwertkurven vom Populationsdurchschnitt mit vom Durchschnitt abweichenden Ausprägungen von Merkmalen, die den Stichprobenpunkten zugeordnet werden können. Und entsprechend könnten Abweichungen der einzelnen Bäume von ihrer Stichproben-Mittelwertkurve durch eine vom Durchschnitt verschiedene Ausprägung eines bestimmten Merkmals erklärt werden. Nach Pinheiro und Bates (2002, S. 367) sollten die folgenden Fragestellungen bei der Modellkonstruktion mit Kovariablen bearbeitet werden:

1. Welche Kovariablen können Variation von Zufallsparametern teilweise erklären?
2. Welche Zufallsparameter werden hinsichtlich ihrer Variation am besten durch die Kovariablen erklärt?
3. Sollten alle Zufallseffekte im Modell verbleiben oder kann auf einige verzichtet werden?

Bisher haben wir den Stichprobenpunkt mit seinen Bäumen näherungsweise als Repräsentant eines Bestandes angesehen. Die Stichprobenpunkte galten als Zufallsauswahl aus dem Untersuchungsgebiet. Die Ausprägungen des Wachstumsverhaltens an den Stichprobenpunkten sahen wir bis zum jetzigen Zeitpunkt als von stochastischer Natur an. Den Einfluss, den ein bestimmter Stichprobenpunkt auf das Wachstum zu haben schien, nahmen wir als Zufallsgröße an. Es lag deshalb auch nahe, dass wir bis jetzt mit dem Stichprobenpunkt ausschließlich zufällige Effekte verbunden haben. Von nun an sollen deterministische Komponenten gesucht werden, die ein bestimmtes Baumwachstum direkt hervorrufen.

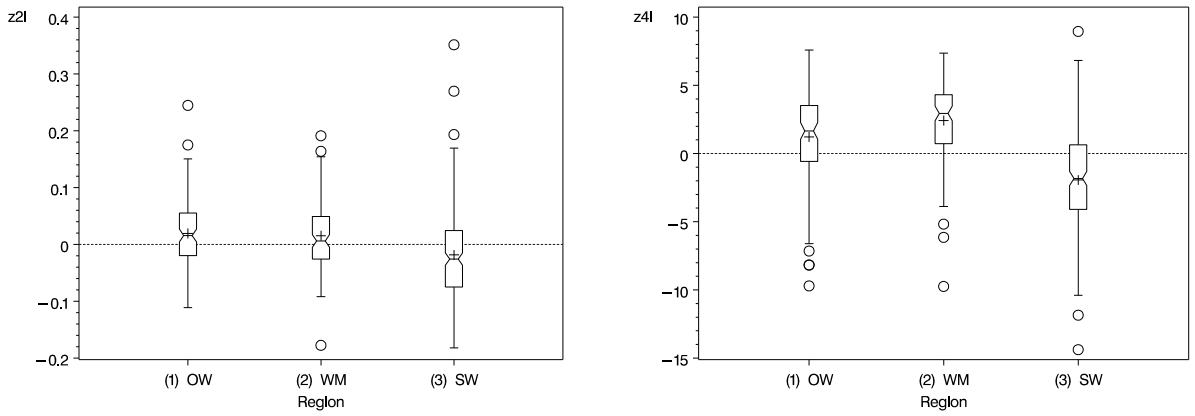
Zunächst sollen Kovariablen aufgespürt werden, mit denen die Variation der Zufallseffekte auf den Parametern erklärt werden kann. Dazu eignet sich das Vorgehen von Pinheiro und Bates (2002), bei dem die vorhergesagten Werte der Zufallsparameter in Abhängigkeit von den Kovariablen grafisch aufgetragen werden. Wie aus der Abbildung 4.6 ersichtlich ist, liegen die Mittelwertkurven für die Stichprobenpunkte in den drei verschiedenen Regionen offensichtlich im Durchschnitt jeweils unterschiedlich zueinander und zur Mittelwertkurve für die Population. Daher liegt die Vermutung nahe, dass sich die

Zufallseffekte auf Ebene der Stichproben $z_i^{(2)}$ und $z_i^{(4)}$ je nach Regionen unterscheiden. Um dieser Vermutung nachzugehen, werden die geschätzten Werte für die Zufallseffekte $z_i^{(2)}$ und $z_i^{(4)}$ durch Box-Plots in Abbildung 4.9 dargestellt. Die Tabelle unter dieser Abbildung enthält die Verteilungsmaße der Zufallseffekte. Mit den Einkerbungen der Boxplots können die Mediane der Verteilungen zwischen den Regionen nach McGill et al. (1978) auf Signifikanz geprüft werden. Mit Median plus bzw. minus $1,58 \cdot ((\mathrm{Q3} - \mathrm{Q1}) / \sqrt{n})$ ergeben sich die obere (K_o) und die untere Grenze (K_u) des Vertrauensintervalls auf Basis des 0,05-Niveaus. Dem Tabellenteil unter der Abbildung 4.9 ist zu entnehmen, dass die obere Grenze für die Verteilung der $z_i^{(4)}$-Zufallseffekte aus dem Odenwald niedriger ist als die untere Grenze für die Verteilung aus der Moränen-Region und dass die untere Grenze für den Odenwald höher ist als die obere Grenze für die Verteilung aus der Region Schwarzwald / Baar-Wutach. Da der Funktionsparameter $\psi_{ij}^{(4)}$ den Erwartungswert für die Baumhöhe im Alter von 50 Jahren darstellt, kann daraus geschlossen werden, dass die Fichten mit einem Alter von 50 Jahren auf dem Westteil der Moränen höher gewesen sind als die Fichten aus den anderen beiden Regionen. Eine 50 jährige Fichte aus der Odenwald-Region ist zudem erwartungsgemäß höher gewesen als eine gleichen Alters aus der Region Südschwarzwald / Baar -Wutach.

Nach Abbildung 4.9 unterscheiden sich die Mediane der Verteilungen für den $z_i^{(2)}$-Zufallsparameter zwischen der Region Odenwald und der Moränen-Region nicht signifikant. Die empirischen Vorhersagewerte des $z_i^{(2)}$-Zufallsparameters sind lediglich für die Fichten aus dem Südschwarzwald im Mittel signifikant kleiner als für die Fichten aus den anderen beiden Regionen.

Aus den im Mittel unterschiedlichen Schätzwerten für die Zufallseffekte $\hat{z}_i^{(4)}$ in den einzelnen Regionen resultieren im Mittel unterschiedliche Parameterwerte für $\psi_{ij}^{(4)}$. Während die mittleren Parameterwerte für $\psi_{ij}^{(2)}$ in den Regionen Odenwald und Moränen recht ähnlich sind, sind die Werte dieses Parameters für die Fichten des Südschwarzwalds im Mittel vergleichsweise kleiner.

Nun soll versucht werden, die erste Frage zu beantworten: Welche Kovariablen können Variation von Zufallseffekten zu erklären? Dabei liegt die Vermutung nahe, dass sich



	$\hat{z}_i^{(2)}$			$\hat{z}_i^{(4)}$		
	(1) OW	(2) WM	(3) SW	(1) OW	(2) WM	(3) SW
N	111	112	210	111	112	210
Mittelwert	0.0194	0.0153	-0.0184	1.217	2.424	-1.936
100% Max	0.2448	0.1913	0.3514	7.593	7.358	8.956
99%	0.1752	0.1640	0.1934	6.887	6.873	5.969
95%	0.1076	0.1292	0.1123	6.465	6.443	3.809
90%	0.0746	0.0845	0.0860	5.362	5.803	2.597
75% Q3	0.0551	0.0492	0.0244	3.513	4.317	0.644
K_o	0.0270	0.0172	-0.0150	2.273	3.473	-1.312
50% Median	0.0158	0.0060	-0.0259	1.659	2.936	-1.828
K_u	0.0046	-0.0051	-0.0367	1.045	2.400	-2.344
25% Q1	-0.0198	-0.0256	-0.0749	-0.582	0.724	-4.090
10%	-0.0407	-0.0580	-0.1149	-2.949	-1.320	-6.793
5%	-0.0659	-0.0673	-0.1305	-5.845	-2.712	-8.334
1%	-0.0786	-0.0918	-0.1527	-8.179	-6.141	-10.398
0% Min	-0.1112	-0.1774	-0.1820	-9.704	-9.741	-14.376

Abbildung 4.9: *Verteilungen für die Vorhersagewerte der Zufallsparameter $z_i^{(2)}$ und $z_i^{(4)}$ auf Stichprobenebene getrennt nach Regionen aus dem zweistufigen Zufallsmodell. OW: Odenwald, WM: westliche Moränen, SW: Südschwarzwald / Baar-Wutach.*

die Regionen als fixe Effekte auf der Ebene der Stichproben in den Parametern $\psi_{ij}^{(2)}$ und $\psi_{ij}^{(4)}$ implementieren lassen. Der Vektor κ für die Parameterwerte der fixen Effekte auf Stichproben-Ebene enthält dann die Interzepte für die Region 1 (Odenwald) mit $k_1^{(2)}$ und $k_1^{(4)}$ und für die Region 3 (Südschwarzwald / Baar-Wutach) mit $k_2^{(2)}$ und $k_2^{(4)}$. Ab diesem Schritt der Modellkonstruktion werden die Zufallseffekte mit $Cov\left[z_i^{(2)}, z_i^{(4)}\right] = Cov\left[z_{ij}^{(2)}, z_{ij}^{(4)}\right] = 0$ als unabhängig voneinander betrachtet und für die Kovarianz-Matrizen der Zufallseffekte Diagonalmatrizen angenommen. Andernfalls kommt es zu Konvergenzproblemen. Sowohl Pinheiro und Bates (2002, S. 388), als auch Hall und Clutter (2004) wenden diese Vereinfachung für eine numerisch stabile Anpassung eines mehrstufigen, gemischten Modells mit mehreren Zufallseffekten an. Zum Vergleich mit diesem neuen Modell (Modell 4) wurde deshalb Modell 2 ohne Kovarianzen zwischen den Zufallsparametern auf den Hierarchieebenen angepasst und als Modell 3 bezeichnet. Wie zu erwarten war, vergrößerte sich der Log-Likelihood-Wert betragsmäßig im Vergleich zu dem von Modell 2. Dieses musste in Kauf genommen werden, wenn neben den Zufallsparametern auch noch von Kovariablen abhängige fixe Effekte installiert werden sollten. Mit den regional

abhängigen fixen Effekten und ohne Kovarianzen für die Zufallsparameter ergibt sich das Modell 4 durch diese Gleichungsfolge:

$$
\begin{aligned}
&\phi_i = \left(\phi_{i1}^T, \ldots, \phi_{in_i}^T\right)^T \\
&= \left(\psi_{i1}^{(2)}, \psi_{i1}^{(3)}, \psi_{i1}^{(4)}, \ldots, \psi_{in_i}^{(2)}, \psi_{in_i}^{(3)}, \psi_{in_i}^{(4)}\right)^T \\
&\psi_{ij}^{(2)} = \iota^{(2)} + k_1^{(2)} I\left\{r = 1\right\} + k_2^{(2)} I\left\{r = 3\right\} + z_i^{(2)} + z_{ij}^{(2)} \\
&\psi_{ij}^{(3)} = \iota^{(3)} \\
&\psi_{ij}^{(4)} = \iota^{(4)} + k_1^{(4)} I\left\{r = 1\right\} + k_2^{(4)} I\left\{r = 3\right\} + z_i^{(4)} + z_{ij}^{(4)} \\
&z_i^{(2)} \sim N\left[0, var\left(z_i^{(2)}\right)\right] \qquad z_i^{(4)} \sim N\left[0, var\left(z_i^{(4)}\right)\right] \\
&z_{ij}^{(2)} \sim N\left[0, var\left(z_{ij}^{(2)}\right)\right] \qquad z_{ij}^{(4)} \sim N\left[0, var\left(z_{ij}^{(4)}\right)\right] \\
&\Delta_1 = \begin{bmatrix} Var\left[z_i^{(2)}\right] & 0 \\ 0 & Var\left[z_i^{(4)}\right] \end{bmatrix} \\
&\Delta_2 = \begin{bmatrix} Var\left[z_{ij}^{(2)}\right] & 0 \\ 0 & Var\left[z_{ij}^{(4)}\right] \end{bmatrix}
\end{aligned}
\tag{4.8}
$$

Dabei ist $I\left\{C\right\}$ eine Indikatorvariable, die den Wert 1 annimmt, wenn die Bedingung wahr ist und andernfalls den Wert 0 (vgl. Hall und Bailey, 2001, bezüglich der Notation). Die Bedingung $r = 1$ bedeutet, dass die betreffende Stichprobe aus dem Odenwald stammt und $r = 3$, dass sie in der Region Südschwarzwald / Baar-Wutach lag.

Die Parameterschätzwerte für dieses zweistufige gemischte Modell (Modell 4) sind der Übersichtstabelle 8.1 des Anhangs zu entnehmen. An dieser Stelle soll dargelegt werden, wie die Brauchbarkeit von Kovariablen als fixe Effekte beurteilt werden kann. Die Signifikanz eines fixen Effektes kann mit dem Wald-Test beurteilt werden (Pinheiro und Bates, 2002, S. 368). Mit $H_0 : \beta_j = 0$ wird die Hypothese aufgestellt, dass das j-te Element des fixen-Effekt-Vektors β den Wert 0 hat. Der Parameter-Schätzwert wird durch seinen geschätzten Standardfehler (ese) dividiert:

$$
t_W = \frac{\hat{\beta}_j^2}{\hat{Var}\left[\hat{\beta}_j\right]} = \left(\frac{\hat{\beta}_j}{ese\left(\hat{\beta}_j\right)}\right)^2 \tag{4.9}
$$

Die Prüfgröße t_W hat näherungsweise eine asymptotische χ^2-Verteilung mit einem Freiheitsgrad. Die Nullhypothese kann bei einer Irrtumswahrscheinlichkeit von 1% verworfen werden, wenn der t_W-Wert größer als 6,635 ist oder bei einer Irrtumswahrscheinlichkeit von 5%, wenn er größer als 3,841 ist. Wegen der schwachen Signifikanz des fixen Effektes $k_1^{(2)}$ auf $\psi_{ij}^{(2)}$ sollte dieser in der weiteren Modellbildung nicht mehr berücksichtigt und fortan angenommen werden, dass die Höhenwachstumskurven von Fichten in den Regionen Odenwald und Moränen im Mittel mit den gleichen $\psi_{ij}^{(2)}$-Parameterwerten beschrieben werden können. Nachdem die Regionen als fixe Effekte aufgefasst sind, dürften die geschätzten Zufallseffekte in den einzelnen Regionen jeweils im Mittel nicht mehr von 0 verschieden sein. Aus Abbildung 4.10 ist zu erkennen, dass dieses zutrifft, indem die Zufallseffekte für die einzelnen Regionen um 0 verteilt sind. In Abbildung 4.11 sind die

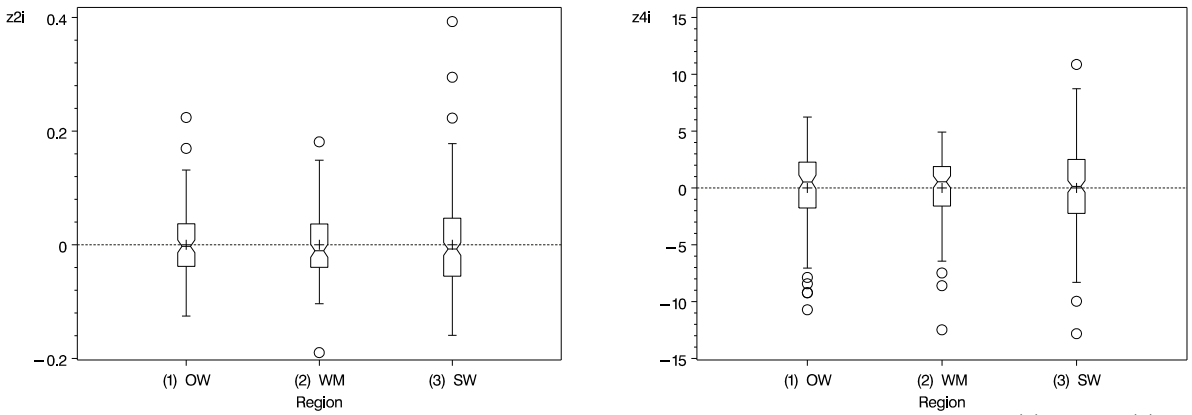


Abbildung 4.10: *Verteilungen für die Vorhersagewerte der Zufallsparameter $z_i^{(2)}$ und $z_i^{(4)}$ auf Stichprobenebene nach Regionen aus dem zweistufigen gemischten Modell 4 mit Regionen als fixe Effekte.*

mittleren Höhenwachstums-Verläufe für die drei Regionen dargestellt, die aufgrund der fixen Effekte zu erwarten sind. Es ist zu erwarten, dass eine Fichte auf den Moränen im Referenzalter von 50 Jahren 24,1 m hoch ist ($\iota^{(4)}$). Fichten aus dem Odenwald sind im gleichen Alter rund 1,3 m niedriger ($k_1^{(4)}$) und die aus der Region Südschwarzwald / Baar-Wutach rund 4,5 m niedriger ($k_2^{(4)}$). Mit den fixen Effekten, die mit der Region verbunden sind, beschreibt das Modell nun räumliche Variation im Höhenwachstum der Fichten in Baden-Württemberg.

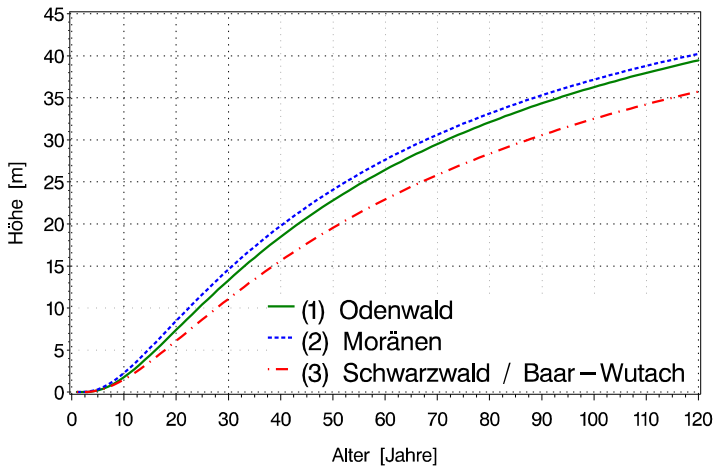


Abbildung 4.11: *Mittelwertkurven für Regionen. Zweistufiges Modell mit Regionen als fixe Effekte auf Parametern* $\psi_{ij}^{(2)}$ *und* $\psi_{ij}^{(4)}$.

4.1.3.2 Überlegungen zur Beschreibung zeitlicher Variation

Die Ergebnisse aktueller waldwachstumskundlicher Untersuchungen aus dem Kompendium von Spiecker et al. (1996) zeigen, dass sich das Wachstum der Waldbäume in der Vergangenheit geändert hat. Folgende Ursachen werden dafür als verantwortlich angesehen: atmosphärische Deposition von Nährstoffen aus der Landwirtschaft oder Verbrennung, klimatische Veränderungen, CO_2-Anstieg in der Atmosphäre, Düngung von Waldbeständen, verbesserte genetische Ressourcen der Waldbäume und Abkehr von historischen, devastierenden Waldnutzungsformen wie Streunutzung und Waldweide. Die longitudinalen Beobachtungen von Höhen zu bestimmten Baumaltern liegen mit den Daten dieser Studie für eine weit zurückreichende Zeitspanne von Kalenderjahren vor. Zum einen bestehen die Daten aus aktuellen Höhenmessreihen der IWE und zum anderen aus aus heutiger Sicht historischen Messungen von Mitscherlich. Es liegt daher nahe zu untersuchen, ob das Höhenwachstum mit der Zeit Veränderungen unterworfen war. Es soll deshalb geklärt werden, ob sich anhand der Daten neben der räumlichen Variation auch eine zeitliche Variation aufspüren lässt. Die gerichtete zeitliche Variation wird in der waldwachstumskundlichen Forschung auch als Wachstumstrend bezeichnet. Dieser ist messbar als eine langfristige Veränderung eines mittleren Levels (Spiecker et al., 1996). Durch den nichtlinearen und

gemischten Ansatz könnte ein zeitlicher Trend wirksam in das Wachstumsmodell implementiert werden.

In einer Fallstudie ging Untheim (1996) der Frage nach, ob das Höhenwachstum in der Vergangenheit einem zeitlichen Trend unterworfen war. Die Datengrundlage seiner Studie waren Höhenmessungen an gefällten Fichten und Buchen auf drei verschiedenen Standortstypen aus der Ostalb-Region in Baden-Württemberg. Da die gefällten Probebäume unterschiedlich alt waren, konnte Untheim die Baumhöhenmesswerte zu bestimmten Altern miteinander vergleichen. Mittels linearer Einfachregressionen konnte er zeigen, dass beispielsweise die Höhe von 50 jährigen Fichten in den Jahren von 1940 bis 1990 im Mittel um rund 10 bis 12 cm pro Jahr angewachsen ist. Das bedeutete einen Anstieg von 5 bis 6 m in 50 Jahren. Das Untersuchungsgebiet aus Untheims Studie war auf die Ostalb begrenzt. Die Resultate konnten mangels zweckdienlicher, moderner Verfahren zur longitudinalen Datenanalyse noch keine Berücksichtigung in Wachstumsmodellen finden. Es wird deshalb mit stochastischen Modellansätzen in dieser Studie an die Vorarbeiten Untheims angeknüpft. Es sollen diejenigen Sachverhalte waldwachstumskundlich modelliert werden, die Untheim in seinem Untersuchungsgebiet beobachten konnte.

Die zeitliche Variation der Baumhöhe in einem bestimmten Alter kann zweckmäßig über das Kalenderjahr der Keimung als Regressorvariable beschrieben werden. Es wäre denkbar, dass das Jahr der Keimung ein fixer Effekt auf den Funktionsparametern ist. Das Jahr der Keimung bzw. die seitdem verstrichene Zeit wäre ein Surrogat für die mögliche Dauer und Intensität von veränderten Umweltbedingungen und Einflüssen am Ort der Stichprobennahme. Das durchschnittliche Keimjahr der Bäume am Stichprobenpunkt i ist gegeben durch das arithmetische Mittel der Keimjahreswerte der einzelnen Bäume des Plots über $KJ_i = 1/n_i \sum_{j=1}^{n_i} KJ_{ij}$. Dabei ergibt sich das Keimjahr des einzelnen Baumes durch das Kalenderjahr des Aufnahmezeitpunktes minus das Baumalter zum Aufnahmezeitpunkt. Zu beachten war, dass das Kalenderjahr des Aufnahmezeitpunktes bei den Stichprobengruppen der oberen Hierarchieebene für die Daten der IWE nicht für alle Bäume gleich ist, sofern diese Punkte mehrmals inventarisiert worden sind (siehe Abschnitt 4.1.2). Damit die Parameterwerte leichter zu deuten sind, wurde zum ersten

das durchschnittliche Keimjahr umformuliert zum Kalenderjahr, in dem die Bäume im Durchschnitt das Alter von 50 Jahren erreicht haben ($yr_{50} = KJ_i + 50$) und zum zweiten wurde dieser Wert als Differenz zum Referenzjahr von 1950 gesetzt, da wir annehmen, dass sich die Bedingungen für das Waldwachstum in der jüngeren Vergangenheit verändert hätten. Deshalb könnten Fichten, die im Kalenderjahr 1950 gekeimt waren im Alter von 50 Jahren eine andere Baumhöhe aufweisen als Fichten die im Jahr 1800 gekeimt waren. Dieses könnte durch unterschiedliche Parameterwerte von $\psi_{ij}^{(4)}$ formuliert werden.

Nach längeren Studien waren derartig eindeutige Trends weder für die gesamte Population, noch für die einzelnen Regionen aus den vorhergesagten Werten für die Zufallsparameter zu lesen. Das liegt daran, dass sich die Keimjahreswerte der Stichprobenpunkte aus den verschiedenen Aufnahmezeitpunkten überlappen. So hatten beispielsweise eine 100 jährige Fichte aus der IWE im Jahre 1994 und eine 60 jährige Fichte aus Mitscherlichs Erhebung im Jahr 1954 das gleiche Keimjahr, nämlich 1894. Angenommen der 100 jährige Baum der IWE hätte in den letzten vier Jahrzehnten einen, im Vergleich zum sonst üblichen Maß, deutlich gesteigerten Höhenzuwachs gezeigt und sein geleistetes Wachstum läge dadurch in bestimmten Altern höher als es erwartet worden wäre, dann kann das Keimjahr diesen Niveauanstieg gegenüber dem 60 jährigen Baum aus Mitscherlichs Erhebung nicht erklären, denn beide Bäume sind im gleichen Jahr gekeimt.

In Abbildung 4.12 sind die Häufigkeiten der Keimjahreswerte aller Bäume getrennt nach den drei Regionen und der Erhebung dargestellt. Die Verteilungen für die IWE-Daten und für Mitscherlichs Daten überlappen sich stark, und zwar in einem Maß, wie es nicht erwartet wurde. Die Ursache dafür ist, dass bei der IWE im Mittel ältere Bäume gefällt wurden als bei Mitscherlichs Erhebung. Das durchschnittliche Alter zum Fällzeitpunkt war für die IWE-Bäume aus dem Odenwald um rund 6 Jahre höher als für die Bäume aus Mitscherlichs Erhebung; die IWE-Bäume von den Moränen waren im Durchschnitt sogar 18 Jahre älter als Mitscherlichs Probestämme aus dieser Region und die Bäume aus dem Südschwarzwald waren um 15 Jahre älter (vgl. dazu Tabelle 8.3 des Anhangs). Sofern das Keimjahr lediglich für die Population oder die Regionen betrachtet würde, ist es deshalb kein geeigneter Regressor, um unterschiedliches Wachstum zu erklären. Die

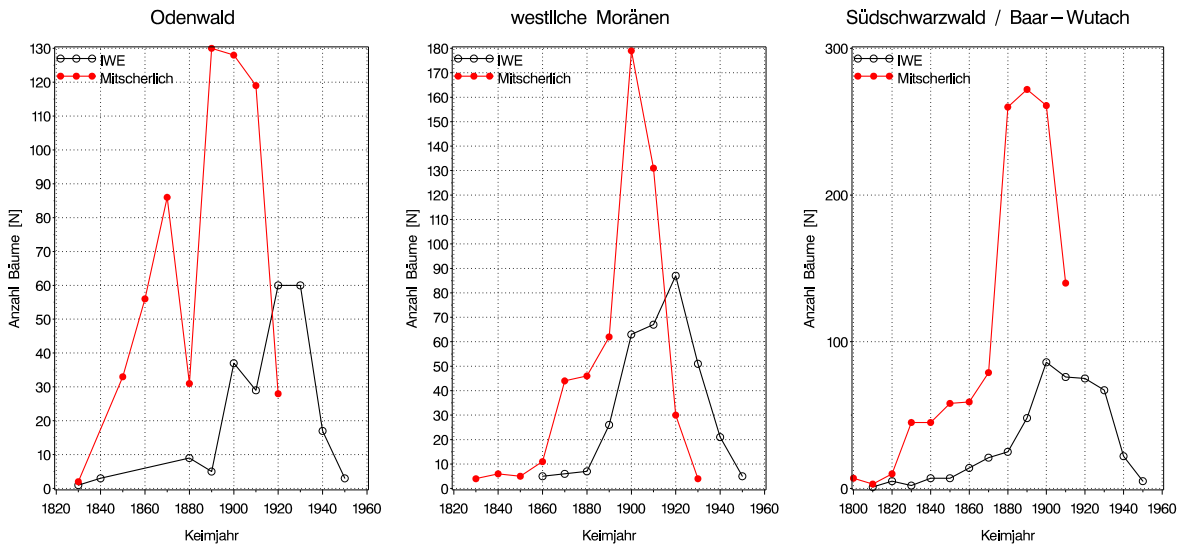


Abbildung 4.12: *Verteilung der Keimjahreswerte aller Probebäume getrennt nach Regionen und Erhebung, Werte gruppiert in Klassen mit einer Breite von 10 Jahren.*

Wachstumsverläufe über das Keimjahr müssen daher gesondert für die IWE mit den Inventurterminen aus der jüngeren Vergangenheit und für Mitscherlichs Aufnahmen aus den weiter zurückliegenden Aufnahmen untersucht werden. In den Abbildungen 4.13 ist die Summe aus Interzept $\iota^{(4)}$ und den geschätzten $\hat{z}_i^{(4)}$ des zweistufigen Zufallsmodells für jeden Stichprobenpunkt nach Regionen und Aufnahmezeiträumen dargestellt. Aufgrund dieser Scatterplot-Streubilder wird vermutet, dass der Höhenerwartungswert im Alter von 50 Jahren in der jüngeren Vergangenheit (IWE-Daten) in den Regionen Schwarzwald und Odenwald angestiegen ist. Außerdem scheint es, als ob die Fichten auf den Moränen in der weiter zurückliegenden Vergangenheit (Mitscherlich-Daten) eine Bonitätsverbesserung erfahren haben. Zunächst soll aber für alle sechs Merkmalsgruppen eine zeitliche Variation der Bonität ($\psi_{ij}^{(4)}$) angenommen werden, die durch das Keimjahr zum Ausdruck kommen soll. Über Wald-Tests für die Parameterschätzwerte aus der Likelihood-Maximierung können dann die obigen Vermutungen statistisch geprüft werden.

Wenn die Zeittrends als fixe Effekte für einzelne Merkmalskombinationen von Region und Datenherkunft (Erhebung im Zuge der IWE oder Mitscherlichs Studie) angenommen werden, dann resultiert aus der bisherigen Modellspezifikation, dass die mittleren Niveaus der Parameterschätzwerte zwischen diesen Merkmalsgruppen variieren könnten. Um

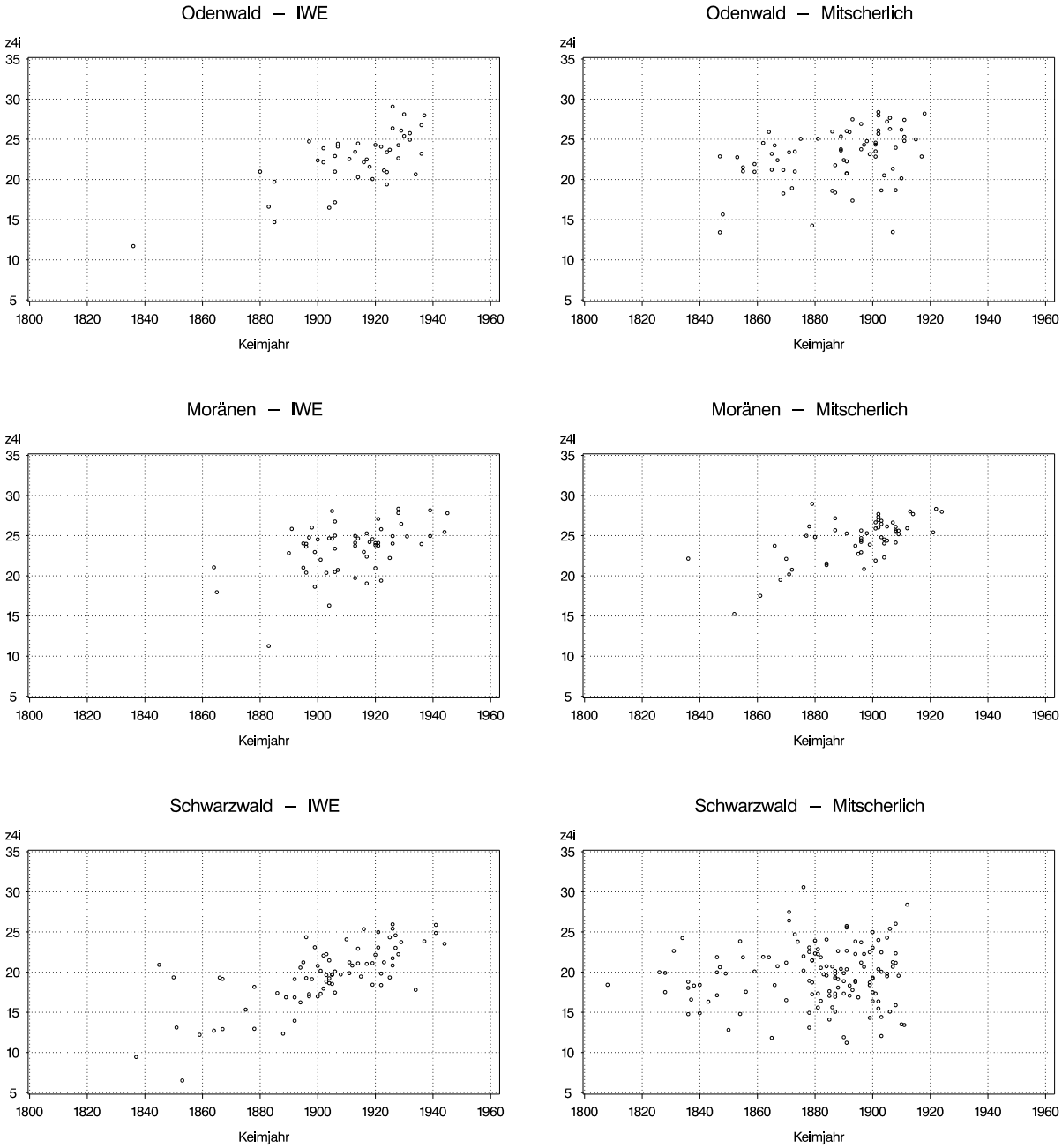


Abbildung 4.13: *Schätzungen für den Höhenerwartungswert-Parameter differenziert nach Region und Erhebung.*

diesen Umstand zu berücksichtigen, müssen anstatt für zwei von drei Regionen, nun für 5 von 6 Merkmalsgruppen fixe Effekte als Interzepte auf den Parametern angenommen werden.

Bevor das Keimjahr überhaupt eine Berücksichtigung im Modell erfährt, soll zunächst ein Modell erstellt werden, das zunächst nur die Interzeptparameter für die Merkmalsgruppen schätzt. Nachdem die Merkmalsgruppen als fixe Effekte angenommen wurden (Modell 5), soll untersucht werden, ob die Verteilungen der vorhergesagten Zufallsparameter Inkonsistenzen aufweisen. Wir fordern, dass die Zufallsparameter in den einzelnen Merkmalsgruppen um den Mittelwert von 0 verteilt sind. Träfe diese Annahme nicht zu, dann wäre das Modell unbrauchbar; denn die Parameter der fixen Effekte für die Merkmalsgruppen aus Region und Datenherkunft hätten keine nutzbare Aussage. Aus Abbildung 4.14 ist zu entnehmen, dass die vorhergesagten Zufallsparameter auf Stichproben-Ebene in den Merkmalsgruppen jeweils einen Mittelwert von 0 haben.

In Abbildung 4.15 sind die Mittelwertkurven für die Regionen getrennt nach dem Ursprung der Daten dargestellt. Bemerkenswert ist, dass die Kurven für die aktuelleren IWE-Daten aus den Regionen Odenwald und Moränen etwas tiefer verlaufen, als diejenigen für die Mitscherlich-Daten. Die Formen der Kurven unterscheiden sich jeweils nicht wesentlich, da die ohnehin nicht hoch signifikant von 0 verschiedenen fixen Effekte auf dem Formparameter $\psi_{ij}^{(2)}$ betragsmäßig zu vernachlässigen sind. Wie zuvor erläutert wurde, bewirken sie aber, dass die vorhergesagten empirischen Zufallsparameter in den Merkmalsgruppen im Mittel um den Wert von 0 verteilt sind. Die Mittelwertkurve für die Fichten der IWE aus der Region Südschwarzwald / Baar-Wutach liegt bis zum Baumalter von rund 52 Jahren zunächst unter der Kurve für die Bäume aus der gleichen Region aus Mitscherlichs Studie; in höheren Altern liegt sie dann darüber. Die Ursache dafür ist der betragsmäßig große und hoch signifikante fixe Effekt auf den Formparameter für die Bäume aus den Mitscherlich-Daten.

Es scheint zunächst nicht plausibel, dass die Mittelwertkurven des Wachstums für die historischen Mitscherlich-Daten höher verlaufen als für die aktuellen IWE-Daten. Dieser Sachverhalt soll im folgenden geklärt werden. Zunächst liegt die Vermutung nahe,

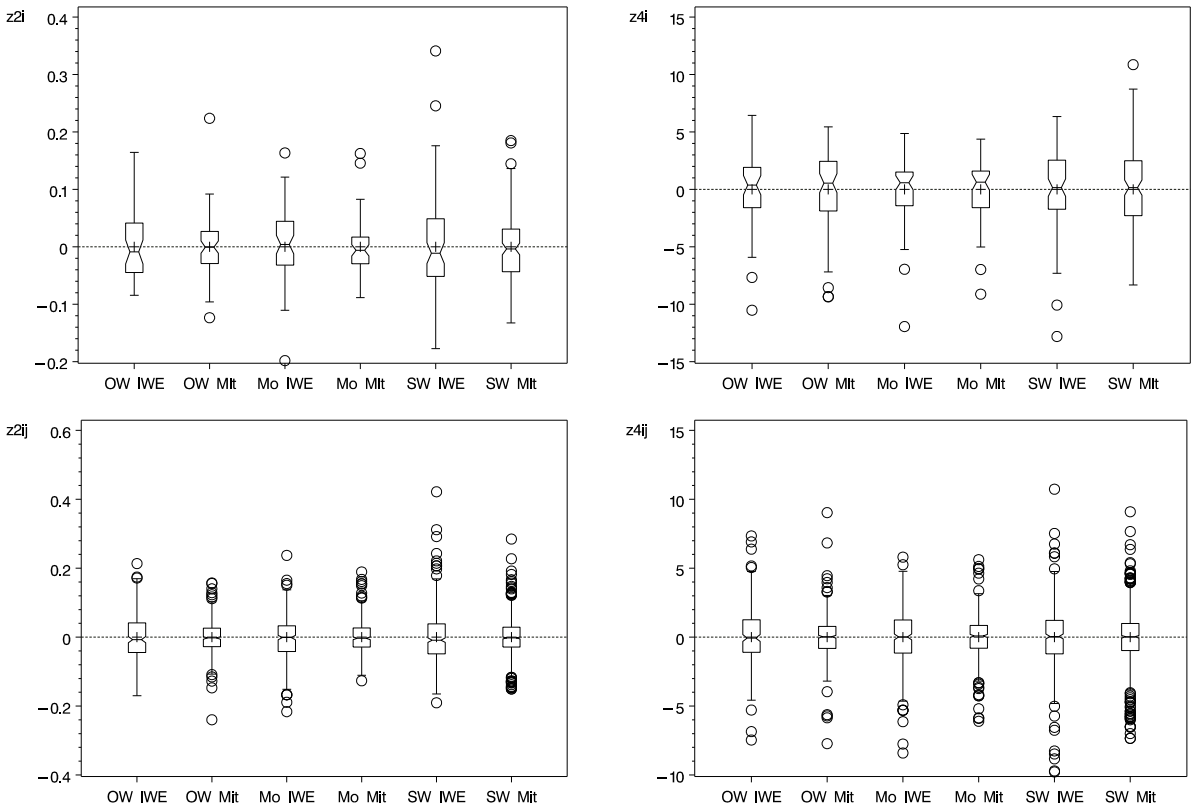


Abbildung 4.14: *Überprüfung der Annahmen für die Zufallseffekte. Box-Plots für vorhergesagte Zufallsparameter über die Merkmalskombinationen von Region und Datenherkunft. Plus-Symbole kennzeichnen die Mittelwerte der Verteilungen. Oben: Vorhergesagte Zufallsparameter auf Stichproben-Ebene. Unten: Vorhergesagte Zufallsparameter auf Baum-Ebene. Links: Zufallsparameter in* $\psi_{ij}^{(2)}$*. Rechts: Zufallsparameter in* $\psi_{ij}^{(4)}$*.*

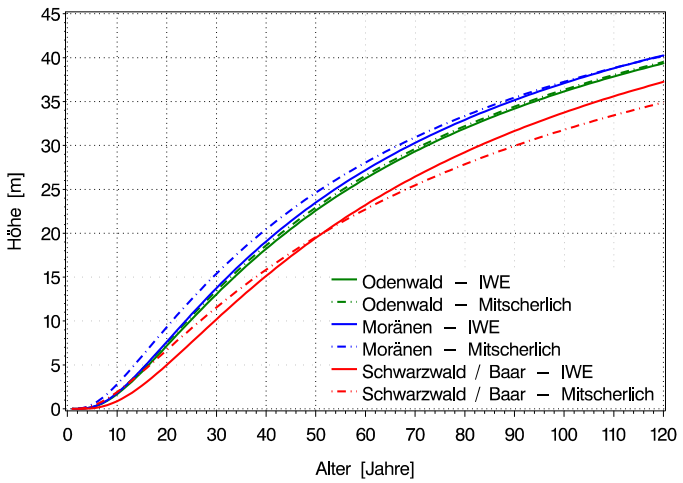


Abbildung 4.15: *Mittelwertkurven nach Regionen und Datenherkunft aus dem zweistufigen Modell mit fixen Effekten für die Merkmalskombinationen aus Region und Datenherkunft.*

dass die Modellanpassung unlogische Parameterwerte liefert und deshalb als gescheitert zu betrachten ist. Wie schon oben geschildert wurde, treffen aber die Annahmen für die Zufallsparameter zu. Die Verteilungen der Zufallsparameter aus Abbildung 4.14 liefern keinen Grund, die Mittelwertkurven anzuzweifeln. In Abbildung 4.16 sind hinter die Mittelwertkurven noch einmal die Beobachtungen hinterlegt.

Zuerst ist zu erkennen, dass die tatsächlichen Höhenwachstumsverläufe der einzelnen Bäume stark um die Mittelwertkurven für die einzelnen Gruppen variieren. Bildlich gesprochen scheinen sich die Mittelwertkurven aber gut einzufügen. Dann wird vor allem deutlich, dass bei den Aufnahmen der IWE vergleichsweise viele alte und mattwüchsige Bäume erfasst wurden. Dieses zeigt auch die Darstellung der Datengrundlage mit Tabelle 8.4 im Anhang.

Bei den Aufnahmen von Mitscherlich wurden zwar auch alte, niedrige Bäume aufgenommen, doch nur in einem vergleichsweise geringeren Ausmaß. Aus Abbildung 4.16 und 4.17 wird ersichtlich, dass durch die Probebäume aus alten Beständen niederer Bonitäten die Schwerpunkte der Stichproben-Mittelwertkurven abgesenkt werden.

Neben diesem Umstand gibt es weitere Anhaltspunkte, die dagegen sprechen, dass die Daten aus den beiden Inventuren einfach zusammengeführt werden können, um gemein-

same Aussagen zu treffen. Sollten bei den Erfassungen der IWE im Vergleich zu der von Mitscherlich die Bäume aus verschiedenen soziologischen Stellungen stammen, dann könnte dieses ebenfalls das scheinbar unterschiedliche Bonitätsniveau erklären. Der Arbeit von Mitscherlich (1957) ist auf S. 8 der anhängenden Ausführungen „Grundsätze und Methode der einmaligen Aufnahme der Fichtenprobeflächen“ zu entnehmen, dass die Probestämme „beliebig aber unter Wahrung der notwendigen waldbaulichen Rücksichten aus den vorherrschenden Stämmen ausgewählt“ wurden. Auch die Probestämme der IWE sind nach Schöpfer et al. (1994) „Höhen vorherrschender Probebäume für einen nicht exakt definierten Mittelstamm einer Spitzenhöhe“. In der Arbeitsanleitung (Schöpfer und Hradetzky, 1983) heißt es: „Es sollen nach Möglichkeit vorherrschende Bäume (KRAFT'sche Klasse 1) mit gut entwickelten Kronen, ohne Wipfelbrüche und möglichst ohne andere Stammbeschädigungen sein“. Erfassungskräfte gaben jedoch Auskunft, dass die Bäume bei den Inventuren der IWE auch niederen soziologischen Stellungen angehört haben konnten. Das liegt zum einen daran, dass aus wirtschaftlichen Gründen nicht immer die höchsten und dicksten Bäume am Stichprobenpunkt für die IWE „geopfert“ werden durften. Andererseits sind zu den Inventurterminen der IWE in den Jahren 1983, 1988 und 1994 an einem Punkt jeweils gezielt ein sogenannter „Plus-Baum“ und ein „Minus-Baum“ gefällt worden. Die „Plus-Bäume“ sollten aus den „relativ gesündesten“ und die „Minus-Bäume“ aus den „stärker erkrankten Bäumen“ ausgewählt werden. Tatsächlich wurden bis einschließlich zur IWE im Jahr 2001 insgesamt 8010 Bäume gefällt, davon stammten nach den Aufschrieben der Erfassungskräfte 5559 aus der Kraft'schen Klasse 1, 2054 Bäume aus Klasse 2, 17 aus Klasse 3, 2 aus Klasse 5 und 378 wurden nicht zugeordnet.

Diese unterschiedlichen Rahmenbedingungen für die Erfassung könnten neben der überproportional häufigen Erfassung alter mattwüchsiger Bäume bei der IWE dafür verantwortlich sein, dass die geschätzten Wachstumsverläufe für die Fichten aus den Aufnahmen der IWE niedriger sind als für die Fichten aus Mitscherlichs Daten.

Eine ursprüngliche Intention dieser Studie war es, durch den Vergleich von IWE- und Mitscherlich-Daten eine zeitliche Variation im Baumhöhenwachstum nachzuweisen. Schöpfer et al. (1994) unternahmen schon einen Versuch, die Höhen der IWE mit denen

aus der Erhebung von Mitscherlich (1957) zu vergleichen. Das komplette Datenmaterial aus Mitscherlichs Erhebung lag jedoch jener Studie nicht zugrunde, sondern nur gemittelte Werte für die einzelnen Probeflächen. Schöpfer et al. (1994) verglichen damals die Bonitäten aus beiden Datengrundlagen für das Wuchsgebiet Schwarzwald. Da jedoch bei der Bonitierung Mittelhöhen in die GWL-Funktion eingesetzt werden, mussten die Höhen aus IWE und Mitscherlichs Erhebung, die als Spitzenhöhen angenommen wurden, auf Mittelhöhen transformiert werden. Für die Höhen der IWE wurden mittlere Differenzen aus den Hilfstabellen für die Forsteinrichtung (Landesforstverwaltung Baden-Württemberg, 1993) verwendet. Für die aggregierten Beobachtungen der Mitscherlich-Studie konnten für jede Probefläche individuelle Abzugstabellen hergeleitet werden, da die aggregierten Tabellen im Anhang der Arbeit von Mitscherlich (1957) auch Angaben zur Höhe des Grundflächenmittelstammes enthielten. Nach diesen Transformationen zeigte sich für Schöpfer et al. (1994), dass für Baumalter zwischen 20 und 40 Jahren die Mittelhöhen aus den IWE-Daten zu dGz-Bonitäten führten, die um 3 bis 4 Vfm/(ha*a) höher waren als diejenigen, die von den Mittelhöhen der Mitscherlich-Daten erzielt wurden. Aus den Mittelwertkurven des hierarchischen, nichtlinearen und gemischten Modell (HNLMM) in Abbildung 4.15 kann ein derartiger Trend nicht abgeleitet werden. In der Studie von Schöpfer et al. (1994) sind die Bonitäten der IWE- und Mitscherlich-Daten getrennt nach Altersklassen verglichen worden; in meiner Studie wird jedoch ein Wachstumsmodell aufgestellt. Aus den zugrunde liegenden longitudinalen Daten werden die Funktionsparameter für die gesamte Population und nicht für einzelne Altersklassen geschätzt. Die Mittelwert-Kurven für die Merkmalskombinationen aus Region und Datenherkunft beschreiben das durchschnittliche Wachstum aller Bäume an allen Stichproben. Die Bäume aus alten Beständen mit schlechten Bonitäten beeinflussen genauso die Interzepte für den Höhenerwartungswert-Parameter $\psi_{ij}^{(4)}$ wie die jungen Bestände, bei denen die schlechten Bonitäten noch nicht so stark ins Gewicht fallen.

Schließlich bewirkt die unterschiedliche Zusammensetzung der Aufnahmekollektive, dass die Daten aus den beiden Inventuren nicht einfach zusammengeführt werden können. Vielmehr müssen die Daten aus den unterschiedlichen Inventuren gesonderte Beachtung

finden. Bei dem bisherigen Vorgehen der Modellbildung haben diese Überlegungen schon besondere Berücksichtigung gefunden, indem für die Merkmalskombinationen aus Region und Datenherkunft fixe Effekte implementiert wurden. Auf diesem Modell kann folglich weiter aufgebaut werden. Wegen der dargelegten Gründe ist eine ertragskundliche Interpretation der unterschiedlichen Mittelwertkurven *zwischen* den beiden Inventuren leider nicht zulässig. Streng genommen beschreibt das Modell 5 einerseits die Höhenentwicklung von Oberhöhenstämmen der Mitscherlich-Daten und andererseits das Wachstum eines eher indifferenten Kolletives aus den IWE-Daten. Vermutlich dürfen wir jedoch unterstellen, dass die erfassten Kollektive *innerhalb* der beiden Inventuren die gleichen Charakteristika aufweisen. Aus den Kurvenverläufen in Abbildung 4.15 können wir dann schließen, dass das Wachstumsniveau auf den Moränen zum Zeitpunkt Mitscherlichs Aufnahme am höchsten war und es gegenwärtig auch noch ist. Etwas niedriger verlaufen die mittleren Kurven für die Fichten aus dem Odenwald, und deutlich geringeres Wachstum zeigten und zeigen die Fichten in der Region Südschwarzwald / Baar-Wutach.

4.1.3.3 Beschreibung zeitlicher Variation mit dem Keimjahr

Im folgenden Schritt der Modellkonstruktion wird nun versucht, die zeitliche Variation zu beschreiben, indem das durchschnittliche Keimjahr als fixer Effekt auf Stichproben-Ebene formuliert wird. Aufgrund der Streubilder der Abbildung 4.13 wurde angenommen, dass die zeitliche Variation durch fixe Effekte in Abhängigkeit vom durchschnittlichen Keimjahr am Stichprobenpunkt abgebildet werden kann. Wald-Test-Statistiken erwiesen einen signifikanten zeitlichen Trend für die Veränderung des Höhenerwartungswert-Parameters $\psi_{ij}^{(4)}$ in den jüngst während der IWE aufgenommenen Fichten aus dem Odenwald und

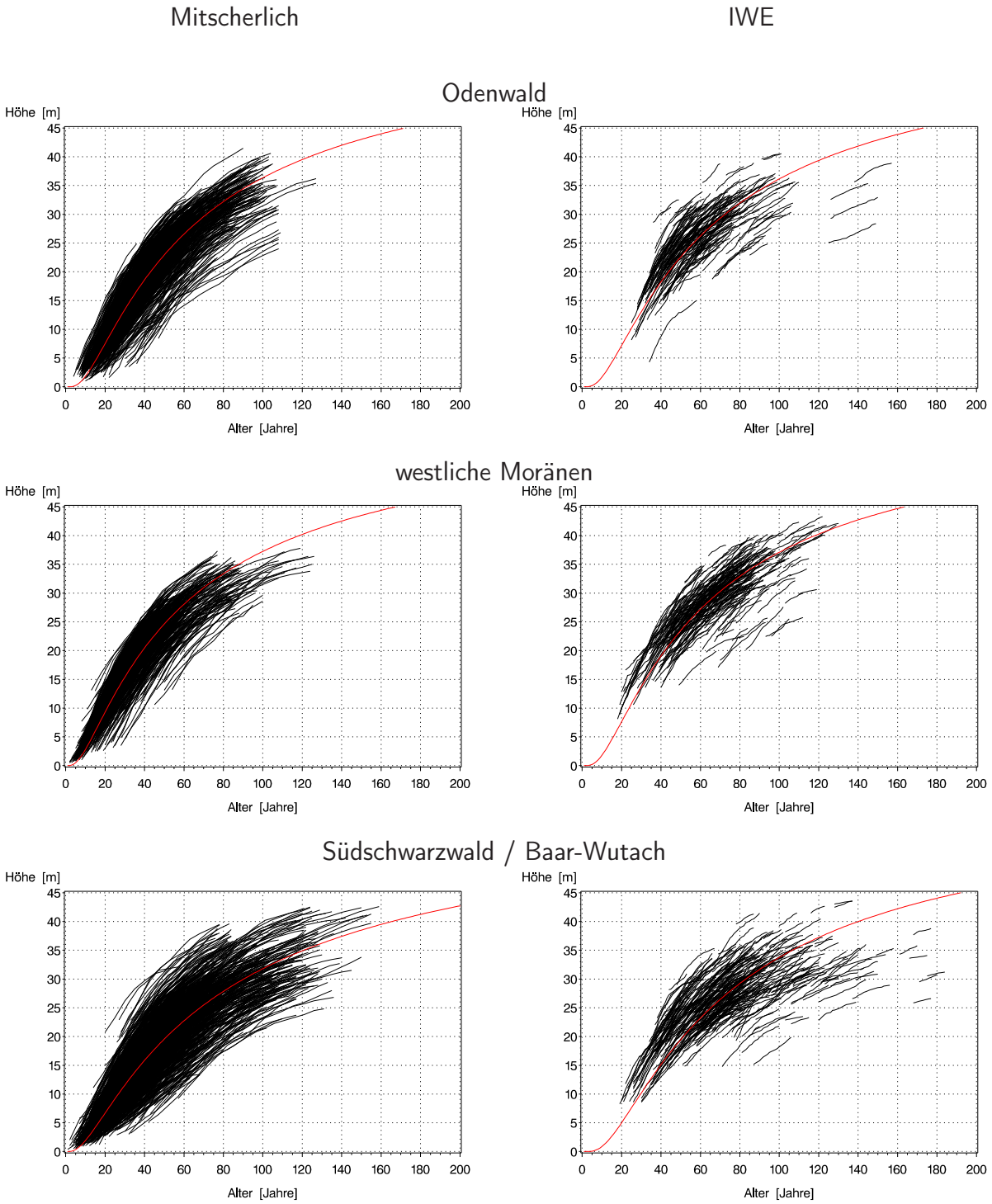


Abbildung 4.16: *Mittelwertkurven und Beobachtungen für die Merkmalsgruppen. Zweistufiges gemischtes Modell mit Merkmalskombination aus Datenherkunft und Region als fixer Effekt auf Stichproben-Ebene.*

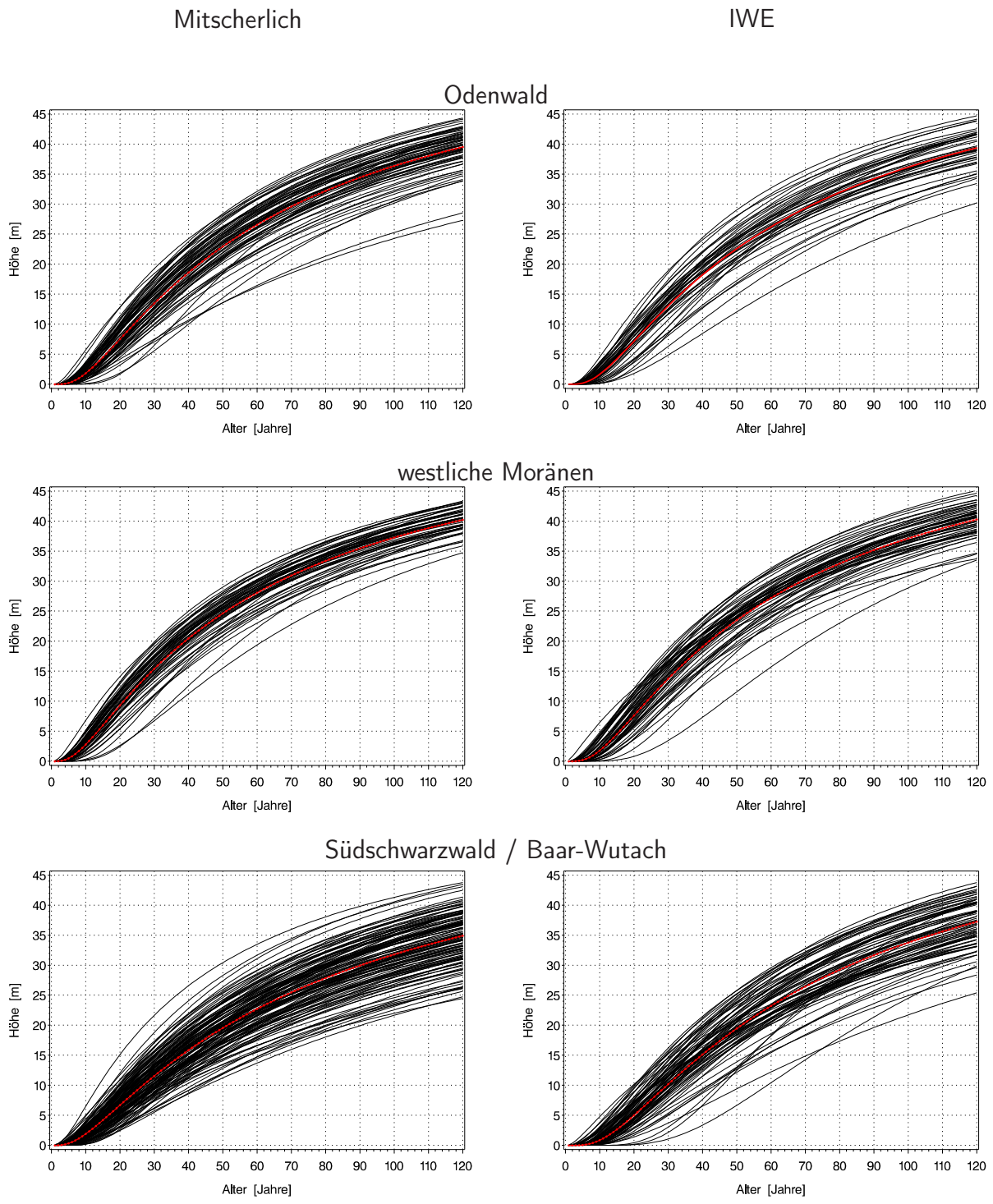


Abbildung 4.17: *Stichproben-Mittelwertkurven nach Merkmalsgruppen. Zweistufiges gemischtes Modell mit Merkmalskombination aus Datenherkunft und Region als fixer Effekt auf Stichproben-Ebene. Schwarz: Stichproben-Mittelwertkurven mit fixen Effekten für Merkmalsgruppen und vorhergesagten Zufallsparametern auf Stichproben-Ebene. Rot: Mittelwert-Kurven für Merkmalsgruppen mit ausschließlich fixen Effekten.*

dem Schwarzwald und für die Fichten von den Moränen aus Mitscherlichs Studie. Das resultierende Modell 6 kann beschrieben werden durch:

$$
\begin{aligned}
\phi_i &= \left(\phi_{i1}^T, \ldots, \phi_{in_i}^T\right)^T \qquad (4.10)\\
&= \left(\psi_{i1}^{(2)}, \psi_{i1}^{(3)}, \psi_{i1}^{(4)}, \ldots, \psi_{in_i}^{(2)}, \psi_{in_i}^{(3)}, \psi_{in_i}^{(4)}\right)^T\\
\psi_{ij}^{(2)} &= \iota^{(2)} + k_1^{(2)} I\left\{r=1, h=1\right\} + k_2^{(2)} I\left\{r=1, h=2\right\}\\
&\quad + k_3^{(2)} I\left\{r=2, h=2\right\} + k_4^{(2)} I\left\{r=3, h=1\right\} + k_5^{(2)} I\left\{r=3, h=2\right\}\\
&\quad + z_i^{(2)} + z_{ij}^{(2)}\\
\psi_{ij}^{(3)} &= \iota^{(3)}\\
\psi_{ij}^{(4)} &= \iota^{(4)} + k_1^{(4)} I\left\{r=1, h=1\right\} + k_2^{(4)} I\left\{r=1, h=2\right\}\\
&\quad + k_3^{(4)} I\left\{r=2, h=2\right\} + k_4^{(4)} I\left\{r=3, h=1\right\} + k_5^{(4)} I\left\{r=3, h=2\right\}\\
&\quad + (yr_{50} - 1950)\\
&\quad \cdot \left[k_6^{(4)} I\left\{r=1, h=1\right\} + k_7^{(4)} I\left\{r=2, h=2\right\} + k_8^{(4)} I\left\{r=3, h=1\right\}\right]\\
&\quad + z_i^{(4)} + z_{ij}^{(4)}\\
z_i^{(2)} &\sim N\left[0, var\left(z_i^{(2)}\right)\right] \qquad z_i^{(4)} \sim N\left[0, var\left(z_i^{(4)}\right)\right]\\
z_{ij}^{(2)} &\sim N\left[0, var\left(z_{ij}^{(2)}\right)\right] \qquad z_{ij}^{(4)} \sim N\left[0, var\left(z_{ij}^{(4)}\right)\right]\\
\Delta_1 &= \begin{bmatrix} Var\left[z_i^{(2)}\right] & 0 \\ 0 & Var\left[z_i^{(4)}\right] \end{bmatrix}, \quad \Delta_2 = \begin{bmatrix} Var\left[z_{ij}^{(2)}\right] & 0 \\ 0 & Var\left[z_{ij}^{(4)}\right] \end{bmatrix}
\end{aligned}
$$

Dabei hat r die selbe Bedeutung wie in Gleichungsfolge 4.8. Wenn die Bedingung $h = 1$ zutrifft, kommt der Stichprobenpunkt aus der IWE und bei $h = 2$ handelt es sich um eine Aufnahme von Mitscherlich. Die Parameterwerte für die IWE-Stichproben von den Moränen werden folglich als Basis-Interzepte benutzt.

Die Abbildung 4.18 zeigt die Stichproben-Mittelwertkurven, die für die Regionen Odenwald und Südschwarzwald / Baar-Wutach aufgrund der IWE-Daten und für die Moränen aufgrund der Mitscherlich-Daten zu erwarten gewesen sind. Den Parameterwerten des Modells 6 zufolge stieg im Odenwald in jüngerer Vergangenheit der Höhenerwartungswert für Fichten im Alter von 50 Jahren ($\psi_{ij}^{(4)}$) um rund 0,14 m pro Jahr an ($k_6^{(4)}$) und in der Region Südschwarzwald / Baar-Wutach um 0,13 m ($k_8^{(4)}$). Auf den Moränen stieg

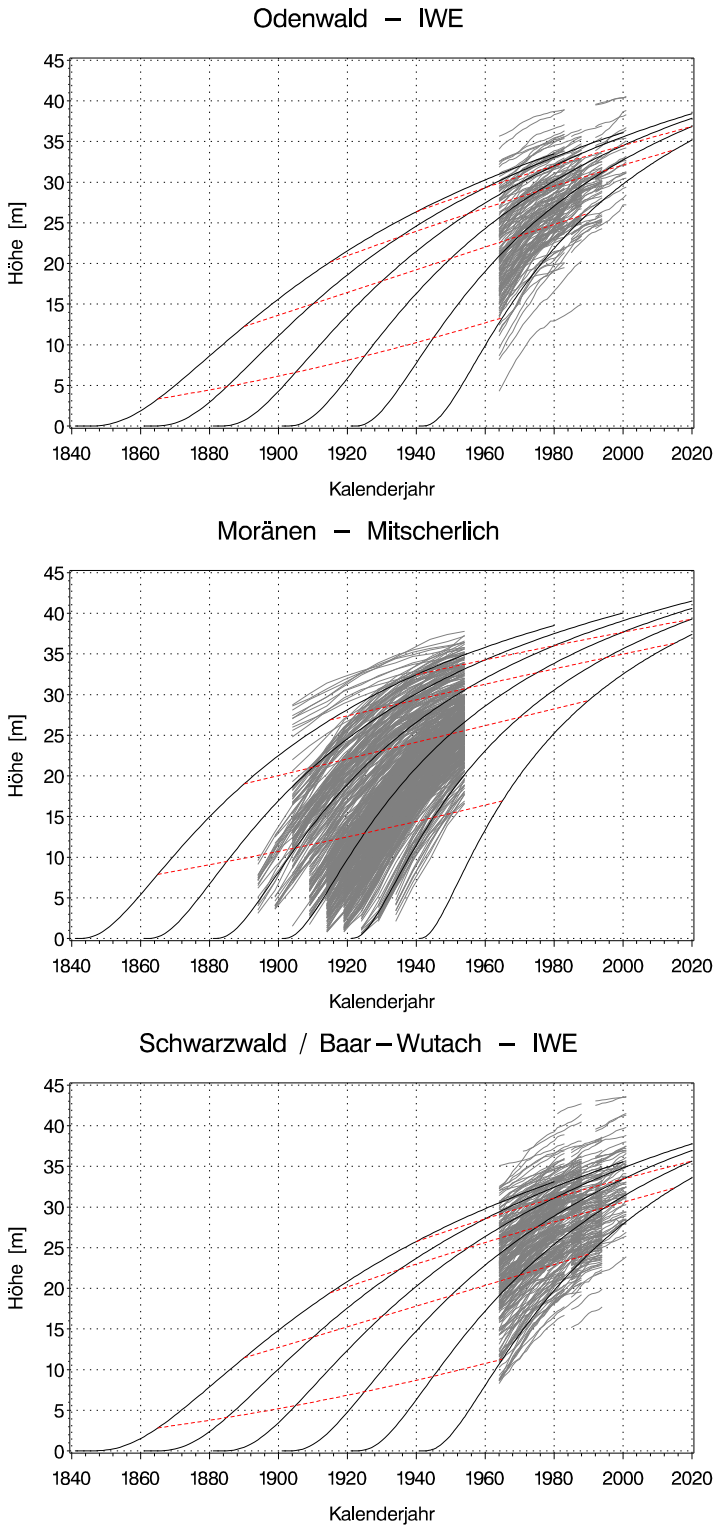


Abbildung 4.18: *Zeitliche Variation des Höhenwachstums von Fichten beschrieben durch das durchschnittliche Keimjahr am Stichprobenpunkt als fixen Effekt auf den Parameter* $\psi_{ij}^{(4)}$*. Die unterbrochenen roten Linien verbinden Höhen gleichen Baumalters (25, 50, 75 und 100 Jahre) bei unterschiedlichen Zeitpunkten der Keimung. Grau im Hintergrund: die Höhenverläufe der einzelnen Bäume aus den zugrunde liegenden Modellierungsdaten.*

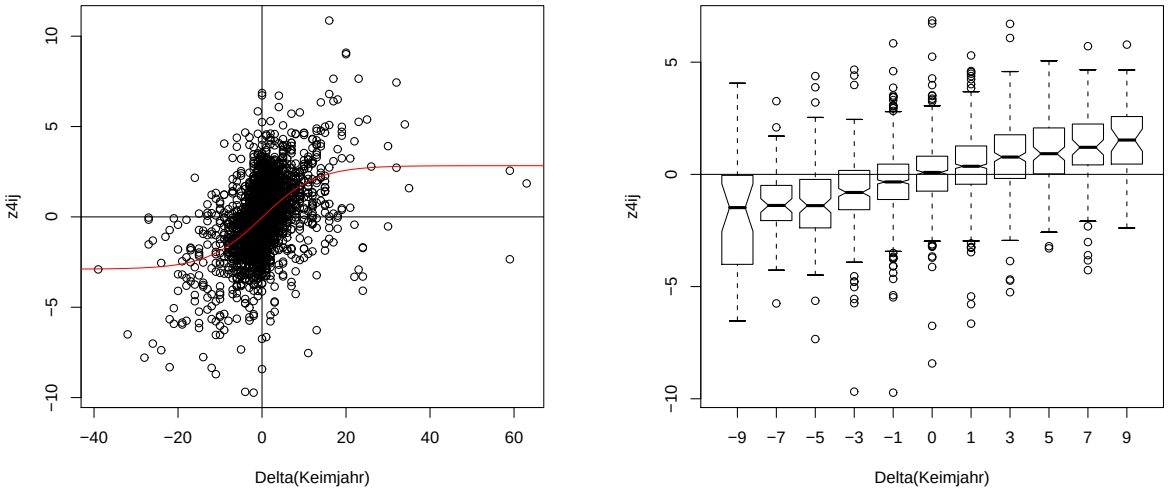


Abbildung 4.19: *Zufallseffekt auf Baumebene. Vorhersagewerte für den Zufallsparameter $z_{ij}^{(4)}$ auf Baumebene aufgetragen über die Abweichung des Keimjahres des Baumes j vom durchschnittlichen Keimjahr am Stichprobenpunkt i. Die rote Linie im linken Streubild ist die Modellkurve für den später eingeführten fixen Effekt auf Baumebene.*

dieser Wert in weiter zurückliegender Vergangenheit um rund 0,1 m pro Jahr an ($k_7^{(4)}$). Für den Parameter $\psi_{ij}^{(2)}$ konnten keine zeitlichen Trends entdeckt werden.

Die Bäume an einem Stichprobenpunkt sind in den meisten Fällen nicht genau gleich alt. Im gesamten, für die Modellanpassung genutzten, Datenbestand weicht bei 2704 von 3397 Bäumen das Alter vom mittleren Alter am Stichprobenpunkt ab. Da die Stichprobenpunkte der IWE verschiedene Aufnahmetermine hatten, muss die Ungleichaltrigkeit anhand der Differenz der Keimjahre der Bäume zueinander beurteilt werden. Die Spannweite zwischen dem ältesten und jüngsten Keimjahr an einem Stichprobenpunkt reicht von 1 bis 71 Jahren. Im Mittel liegen zwischen der Keimung des am frühesten und am spätesten gekeimten Baum 10 Jahre, bei einer Standardabweichung von rund 5,5 Jahren. Es soll nun untersucht werden, wie sich die Parameterwerte der einzelnen Bäume an einem Stichprobenpunkt unterscheiden; schließlich entfallen im aktuellen Modell rund 29% der gesamten Varianz der Zufallskomponenten im Höhenerwartungswert-Parameter auf die Baum-Ebene.

In Abbildung 4.19 sind die vorhergesagten Zufallsparameter $\hat{z}_{ij}^{(4)}$ auf Baum-Ebene über die Differenz ihres Keimjahres zum durchschnittlichen Keimjahr am Stichprobenpunkt $\delta(KJ)_{ij} = KJ_{ij} - 1/n_i \sum_{j=1}^{n_i} KJ_{ij}$ aufgetragen. Aufgrund der Vielzahl an Beobachtungen ist aus dem Streubild nur schwer ein Trend abzulesen. Deshalb wurde der nebenstehende Box-Plot für klassifizierte Abszissenwerte angefertigt. Daraus ist leichter ein Trend zu entnehmen. Je später ein einzelner Baum im Vergleich zu anderen Bäumen am Stichprobenpunkt gekeimt war, desto höher ist sein Zufallsparameter $z_{ij}^{(4)}$ und damit auch der Höhenerwartungswertparameter ($\psi_{ij}^{(4)}$) aus dem Wachstumsmodell. Dafür gäbe es zwei Erklärungen. Zum einen könnte das Versuchsdesign der IWE dafür verantwortlich sein. Nämlich dann, wenn in Ausnahmefällen bei verschiedenen Aufnahmezeitpunkten die Bäume aus verschiedenen Beständen mit unterschiedlichen Entstehungsjahren entnommen wurden. Die Bäume aus dem Bestand, der später gekeimt war, hätten analog zum Trend auf Stichprobenebene höhere Baumhöhen bei gleichem Baumalter. Eine andere mögliche Erklärung wäre, dass früher gekeimte Bäume länger unter dem Schirmdruck von Altbestandsresten oder Überhältern des Vorbestandes zu leiden gehabt hatten und später gekeimte Bäume wegen günstigerer Lichtverhältnisse eine höhere Wachstumsgeschwindigkeit in ihrer frühen Jugend zeigen konnten. Diese Erklärungen haben jedoch nur spekulativen Charakter und können durch die Datengrundlage nicht verifiziert werden. Das Phänomen soll aber durch das Modell beschrieben werden. Die Mittelwerte von $\hat{z}_{ij}^{(4)}$ beschreiben über Klassen von $\delta(KJ)_{ij}$ einen sigmoidalen Verlauf mit einem Wendepunkt im Ursprung und einer oberen und unteren Asymptote. Die Annahme von Asymptoten würde zur biologischen Plausibilität des Modells beitragen. Zur funktionalen Beschreibung wurde mit Gleichung 4.11 eine erweiterte logistische Funktion aus Ratkowsky (1989) (dort: Gleichung 5.4.1) verwendet. Der Parameter τ gibt die untere Asymptote an und $\tau + \alpha$ die obere. Der Parameter γ bestimmt die Steigung am Wendepunkt β/γ.

$$y = \tau + \alpha / \left[1 + \exp\left(\beta - \gamma\delta(KJ)_{ij}\right)\right] \tag{4.11}$$

Sobald das Keimjahr des Baumes j dem Mittelwert der Keimjahreswerte an der Stichprobe i entspricht, soll der fixe Effekt 0 sein. Mit $\beta = 0$ aus Gleichung 4.11 nehmen wir deshalb an, dass die Kurve durch den Ursprung geht und dort auch ihren Wendepunkt hat. Die

vorhergesagten Werte der Zufallsparameter auf Baumebene für $\hat{z}_{ij}^{(4)}$ wurden durch Kleinstquadratschätzung in Abhängigkeit von $\delta(KJ)_{ij}$ mit dem Gauß-Newton-Algorithmus ausgeglichen. Die Parameterschätzwerte sind in Tabelle 4.2 enthalten. Im weiteren Modell-

	Schätzung	Std. Fehler	t-Wert	Pr($>\|t\|$)
τ	-2.8926	0.2399	-12.057	<2e-16
α	5.7264	0.4750	12.056	<2e-16
γ	0.1535	0.0184	8.342	<2e-16

Tabelle 4.2: *Parameterwerte aus OLS-Schätzung für den fixen Effekt auf Baum-Ebene.*

bildungspozess soll diese Keimjahrdifferenz als fixer Effekt auf Baum-Ebene ins Modell eingefügt werden. Es ergaben sich Konvergenzschwierigkeiten, sofern alle Parameter der Kurve aus Gleichung 4.11 über die Likelihood-Maximierung geschätzt werden sollten; das Modell schien überparametrisiert zu sein. Bei einer ähnlichen Fragestellung fixierte Lappi (1997) den formbestimmenden Parameter mit Schätzwerten aus einer vorhergehenden OLS und ließ lediglich die asymptotischen Parameter Likelihood-basiert schätzen. In dieser Arbeit werden die asymptotischen Parameter fixiert und der Steigungsparameter γ im Modell-Fit geschätzt. In Abbildung 4.20 ist zu erkennen, dass das Ziel erfüllt wurde und durch den fixen Effekt der Trend der Zufallsparameter auf Baum-Ebene $\hat{z}_{ij}^{(4)}$ bereinigt wurde. Für die Zufallsparameter $\hat{z}_{ij}^{(2)}$ konnte kein deutlicher Trend beobachtet werden, ein fixer Effekt wurde deshalb nicht angenommen.

Die Modelldaten enthalten für die einzelnen Bäume keine weiteren individuellen Kovariablen. Als weitere fixe Effekte auf Baum-Ebene wären der Kronenansatz, oder Distanzen zu Nachbarbäumen denkbar; diese wurden jedoch nicht erfasst. Die Variation, die sich wegen dieser Charakteristika ergeben könnte, kann also nicht durch fixe Effekte des Modells beschrieben werden; sie verbleibt vielmehr in den Zufallsparametern.

Grundsätzlich könnten die Mittelwertkurven in Abhängigkeit von allen erdenklichen Standortsvariablen konstruiert werden. Es wurden Versuche unternommen, verschiedene Kovariablen als fixe Effekte auf Stichproben-Ebene zu gewinnen. Es lagen Beobachtungswerte für folgende Größen vor: Neigung und Exposition des Geländes, Höhe ü. NN, klimatische Wasserbilanz, Niederschlagssummen und Temperaturen für das Jahr und die

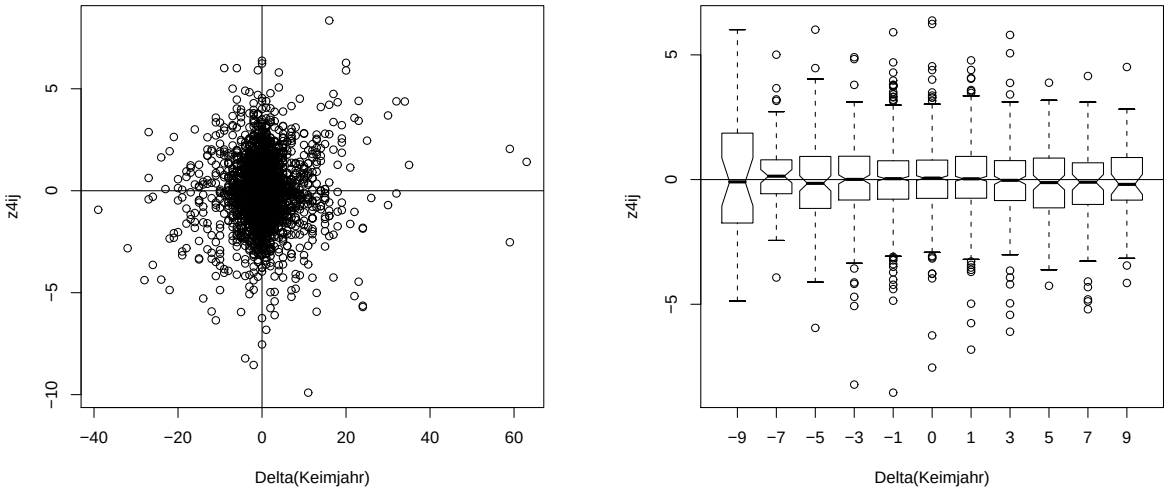


Abbildung 4.20: *Vorhersagen für den Zufallsparameter $\hat{z}_{ij}^{(4)}$ auf Baumebene über Abweichung des Keimjahres des Baumes j vom durchschnittlichen Keimjahr am Stichprobenpunkt i nach Implementierung des fixen Effektes.*

Vegetationszeit, Skelettanteil und Humusform (nur IWE). Diese Größen wurden in unterschiedliche Skalen transformiert, indiziert und kategorisiert. Die Bemühungen blieben jedoch fruchtlos. Die Koeffizienten waren selten signifikant und wenn doch, dann verringerte sich die Varianz der Zufallsparameter nicht merklich. Einzig die Höhe ü. NN schien zunächst für die gesamte Population ein nominell signifikanter fixer Effekt zu sein. Dabei muss aber beachtet werden, dass bereits für die Indikatorvariablen der drei Regionen (bzw. der Kombination von Region und Datenherkunft) als fixe Effekte Interzeptparameterwerte geschätzt wurden. Die Variable Höhenlage korreliert wiederum stark mit den regionalen Indikatorvariablen. Die Stichproben im Schwarzwald haben im Durchschnitt eine höhere Höhenlage als die von den Moränen und diese liegen im Schnitt wieder höher als die Punkte aus dem Odenwald (siehe Tabelle 4.3). Zur Vermeidung von Kollinearitätsproblemen

Region	N	Mittelwert	StdAbw	Min.	Max.
Odenwald	111	358.7	81.9	135.0	537.5
Moränen	112	621.7	84.1	422.2	808.5
Schwarzwald / Baar-Wutach	210	840.8	178.4	369.1	1207.0

Tabelle 4.3: *Höhenlage [m ü. NN] für die Stichprobenpunkte nach Regionen.*

wurde die Höhenlage deshalb nicht als fixer Effekt der Population berücksichtigt; als fixer Effekt in den einzelnen Gruppen von Region und Aufnahme hatte die Höhenlage darüber hinaus keine signifikanten Parameterwerte. Die weitere Suche nach möglichen fixen Effekten geriet dann ineffizient und musste schließlich abgebrochen werden. Wie im nächsten Abschnitt gezeigt wird, ist diese Erkenntnis jedoch kein Grund zur Entmutigung. Der Großteil der Variation kann mit Zufallsparametern beschrieben werden und diese können mit BLUPs anhand von Vorbeobachtungen für Fortschreibungen geschätzt werden. Wir sind zwar nicht auf die Gründe für die Variation gestoßen, aber wir können sie dennoch abschätzen und mit Zufallsparametern modellieren.

Damit erhalten wir das endgültige Modell 7:

$$
\begin{aligned}
\phi_i &= \left(\phi_{i1}^T, \ldots, \phi_{in_i}^T\right)^T \qquad (4.12)\\
&= \left(\psi_{i1}^{(2)}, \psi_{i1}^{(3)}, \psi_{i1}^{(4)}, \ldots, \psi_{in_i}^{(2)}, \psi_{in_i}^{(3)}, \psi_{in_i}^{(4)}\right)^T\\
\psi_{ij}^{(2)} &= \iota^{(2)} + k_1^{(2)} I\left\{r=1, h=1\right\} + k_2^{(2)} I\left\{r=1, h=2\right\}\\
&\quad + k_3^{(2)} I\left\{r=2, h=2\right\} + k_4^{(2)} I\left\{r=3, h=1\right\} + k_5^{(2)} I\left\{r=3, h=2\right\}\\
&\quad + z_i^{(2)} + z_{ij}^{(2)}\\
\psi_{ij}^{(3)} &= \iota^{(3)}\\
\psi_{ij}^{(4)} &= \iota^{(4)} + k_1^{(4)} I\left\{r=1, h=1\right\} + k_2^{(4)} I\left\{r=1, h=2\right\}\\
&\quad + k_3^{(4)} I\left\{r=2, h=2\right\} + k_4^{(4)} I\left\{r=3, h=1\right\} + k_5^{(4)} I\left\{r=3, h=2\right\}\\
&\quad + (yr_{50} - 1950)\\
&\quad \cdot \left[k_6^{(4)} I\left\{r=1, h=1\right\} + k_7^{(4)} I\left\{r=2, h=2\right\} + k_8^{(4)} I\left\{r=3, h=1\right\}\right]\\
&\quad + \left(-2.8926 + 5.7264 / \left[1 + \exp\left(-k_9^{(4)} \delta_{ij}^{(KJ)}\right)\right]\right)\\
&\quad + z_i^{(4)} + z_{ij}^{(4)}\\
z_i^{(2)} &\sim N\left[0, var\left(z_i^{(2)}\right)\right] \qquad z_i^{(4)} \sim N\left[0, var\left(z_i^{(4)}\right)\right]\\
z_{ij}^{(2)} &\sim N\left[0, var\left(z_{ij}^{(2)}\right)\right] \qquad z_{ij}^{(4)} \sim N\left[0, var\left(z_{ij}^{(4)}\right)\right]\\
\Delta_1 &= \begin{bmatrix} Var\left[z_i^{(2)}\right] & 0 \\ 0 & Var\left[z_i^{(4)}\right] \end{bmatrix}, \quad \Delta_2 = \begin{bmatrix} Var\left[z_{ij}^{(2)}\right] & 0 \\ 0 & Var\left[z_{ij}^{(4)}\right] \end{bmatrix}
\end{aligned}
$$

4.2 Modellevaluierung

Nachdem das nichtlineare, hierarchische und gemischte Modell für das Baumhöhenwachstum der Fichte konstruiert worden ist, soll es evaluiert werden. Dazu soll eine Prognose für Baumhöhen aus einem für die Modellbildung nicht verwendeten Datensatz vorgenommen werden. Die Daten zur Modellevaluierung stammen von aktuellen Höhenmessungen an gefällten Stämmen aus dem Südschwarzwald; sie sind im Abschnitt 2.2 beschrieben. Für die Prognose des Baumhöhenwachstums einzelner Bäume müssen die bekannten fixen Effekte berücksichtigt und die Zufallsparameter vorhergesagt werden. Als fixe Effekte auf Stichproben-Ebene enthält das Modell die Merkmalskombination von Region und Datenherkunft. Die Wertebereiche der durchschnittlichen Keimjahre an den Stichproben aus den aktuellen IWE-Aufnahmen und den historischen Aufnahmen von Mitscherlich überlappten sich und die Zusammensetzung der Kollektive unterschied sich bei beiden Aufnahmen. Diese beiden Gründe führten dazu, dass fixe Effekte für Merkmalskombinationen aus den Regionen und der Inventur gebildet wurden. Die Evaluierungsdaten wurden ausschließlich im Südschwarzwald erhoben; deshalb nehmen wir die Region Südschwarzwald / Baar-Wutach als einen Teil der Merkmalskombination der fixen Effekte an.

Mündlicher Überlieferungen zufolge durften die Erfassungskräfte bei den Aufnahmen zum ZF-Projekt ähnlich wie bei der IWE nicht ausschließlich die stärksten und dicksten Bäume fällen und ebenso wie die IWE sind die ZF-Aufnahmen aktuell und umfassen Baumhöhenbeobachtungen aus der jüngsten Vergangenheit. Daher werden die fixen Effekte für die Stichproben der Evaluierungsdaten aufgrund der Parameter für die Merkmalskombination „Südschwarzwald / Baar-Wutach - IWE“ aus der Modellanpassung geschätzt. Neben den fixen Effekten müssen auch die Zufallsparameter vorhergesagt werden. Die Vorhersage der Zufallsparameter wird von Lappi (1991) als „Kalibrierung“ bezeichnet; diese ist nicht mit der Modellanpassung zur Parameterschätzung zu verwechseln. Bei der Kalibrierung zeigt sich die Stärke des auf Zufallsparametern basierenden Modellansatzes. Damit die Zufallsparameter für jeden Baum vorhergesagt werden können, müssen eine oder mehrere Beobachtungen als bekannt angenommen werden. Anhand der Abwei-

chung der prognostizierten Werte von den tatsächlichen soll die Güte des Modells in der praktischen Anwendung beurteilt werden. Die Zufallsparameter werden über empirische Bayes'sche Schätzer nach Gleichung 3.24 für jeden Baum durch lineare Approximation vorhergesagt. Dieser Prädiktor kombiniert in optimaler Weise die verfügbaren Messungen (die als bekannt angenommen Beobachtungen) mit der Kenntnis über die Varianzen und Kovarianzen der Zufallsparameter aus der vorherigen Modellanpassung (Lappi und Bailey, 1988).

Ein wesentlicher Gesichtspunkt ist, dass durch die BLUPs der Zufallsparameter mehr Parameter geschätzt werden können, als Beobachtungen vorhanden sind. Wie hier gezeigt wird, kann die nichtlineare Wachstumsfunktion sogar anhand nur einer Beobachtung pro Individuum kalibriert werden. Für die Vorhersage der Zufallsparameter nehmen wir Beobachtungen eines Baumes als bekannt an. Diese Beobachtungen aller Bäume am Stichprobenpunkt i füllen den Vektor y_i^* aus Gleichung 3.23. Mit Hilfe der bekannten Kovariablen und der Parameterwerte aus der Modellanpassung werden die fixen Effekte beschrieben. Mit Funktions-Parametern, denen nur die fixen Effekte zugrunde liegen, werden durch $f_i\left(\theta_i^*, A_i^*, \hat{\beta}\right)$ Baumhöhen vorhergesagt, die man aufgrund der fixen Effekte erwarten würde. Mit den Residuen $\hat{e}_i$ werden die Zufallsparameter durch Gleichung 3.24 nacheinander für die Stichprobenpunkte mit ihren Bäumen vorhergesagt. Die Matrix-Algebra für diesen Prozess wurde in `IML (Interactive Matrix Language)` aus `SAS 9.1` programmiert.

Die Kalibrierung soll für diese zwei verschiedene Varianten von vorliegenden Informationen geschehen:

1. Die jeweils erste Beobachtung eines Baumes wird als bekannt angenommen. Die zweite und alle folgenden Beobachtungen sollen geschätzt werden.

2. Die jeweils erste und elfte Beobachtung eines Baumes werden als bekannt angenommen. Die zwölfte und alle folgenden Beobachtungen sollen geschätzt werden.

Das Verfahren der Kalibrierung soll exemplarisch anhand der Daten eines Stichprobenpunktes dargelegt werden. Am ausgewählten Stichprobenpunkt sind sechs Bäume erfasst worden. Die jeweils ersten Höhenbeobachtungen

$$y_i^* = (26,98 \quad 27,37 \quad 25,07 \quad 27,72 \quad 33,6 \quad 14,0)^T \tag{4.13}$$

werden als bekannt angenommen. Sie korrespondieren mit den beobachteten Werten von Baumaltern

$$\theta_i^* = (58 \quad 67 \quad 73 \quad 69 \quad 85 \quad 32)^T \tag{4.14}$$

Nach der allgemeinen Darstellung des nichtlinearen, hierarchischen und gemischten Modells durch die Gleichungsfolge 3.9 im Abschnitt 3.2.2 ist der Parametervektor ϕ_i eine Linearkombination aus den fixen und zufälligen Parametern. Im endgültigen Modell (Gleichungsfolge 4.12) ist $\hat{k}_9^{(4)}$ jedoch ein Parameter einer nichtlinearen Funktion für den fixen Effekt auf Baum-Ebene. Deswegen ergibt sich der Parametervektor $\hat{\phi}_i$ nach einer Erweiterung durch:

$$\hat{\phi}_i = A_i^* \hat{\beta} + C_i r_i + B_i \hat{b}_i \tag{4.15}$$

Dabei ist $\hat{\beta}$ ein Vektor mit den Parameterwerten (außer $\hat{k}_9^{(4)}$) für die fixen Effekte (vgl. Tabelle 8.1 des Anhangs).

$$\hat{\beta} = \begin{pmatrix} \hat{\iota}^{(2)} \\ \hat{\iota}^{(3)} \\ \hat{\iota}^{(4)} \\ \hat{k}_1^{(2)} \\ \hat{k}_2^{(2)} \\ \hat{k}_3^{(2)} \\ \hat{k}_4^{(2)} \\ \hat{k}_5^{(2)} \\ \hat{k}_1^{(4)} \\ \hat{k}_2^{(4)} \\ \hat{k}_3^{(4)} \\ \hat{k}_4^{(4)} \\ \hat{k}_5^{(4)} \\ \hat{k}_6^{(4)} \\ \hat{k}_7^{(4)} \\ \hat{k}_8^{(4)} \end{pmatrix} = \begin{pmatrix} 0,526086 \\ 0,856212 \\ 23,483659 \\ -0,008690 \\ -0,013932 \\ -0,034741 \\ 0,003391 \\ -0,079738 \\ -2,838461 \\ -0,512227 \\ 1,673859 \\ -4,600783 \\ -3,938693 \\ 0,139219 \\ 0,103268 \\ 0,135276 \end{pmatrix} \tag{4.16}$$

Die Matrix A_i^* ist eine Design-Matrix für den Parametervektor $\hat{\beta}$, bestehend aus den Werten der Regressor-Kovariablen. Die Elemente in A_i^* mit den Werten gleich 1 gewährleisten, dass von den fixen Interzept-Parametern nur auf diejenigen zurückgegriffen wird, die auch benötigt werden, nämlich $\hat{\iota}^{(2)}, \hat{\iota}^{(3)}, \hat{\iota}^{(4)}, \hat{k}_4^{(2)}$ und $\hat{k}_4^{(4)}$. Die Elemente mit den Werten von 12 entsprechen dem fixen Effekt für das durchschnittliche Keimjahr an diesem Stichprobenpunkt. Dieser ist als Differenz zwischen dem Kalenderjahr, in dem die Bäume im Mittel das Alter von 50 Jahren erreicht haben oder erreichen werden, und dem Referenzjahr 1950

formuliert (vgl. Abschnitt 4.1.3.3). Die Bäume dieses Stichprobenpunktes erreichten also im Mittel im Jahr 1962 das Alter von 50 Jahren.

$$A_i^* = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 12 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 12 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 12 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 12 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 12 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 12 \end{pmatrix} \tag{4.17}$$

Der Vektor r_i beinhaltet den Effekt auf der Ebene des einzelnen Baumes auf den Höhenerwartungswert-Parameter, der mit einer vom Durchschnitt späteren Keimung verbunden ist.

$$r_i = \begin{pmatrix} r_{i1} \\ r_{i2} \\ r_{i3} \\ r_{i4} \\ r_{i5} \\ r_{i6} \end{pmatrix} = \begin{pmatrix} 1,2265533 \\ -1,059066 \\ -2,135967 \\ -0,29807 \\ -2,802932 \\ 2,8242977 \end{pmatrix} \tag{4.18}$$

Dabei ergibt sich r_{ij} in Abhängigkeit von der Keimjahresdifferenz $\delta(KJ)_{ij}$ aus der nichtlinearen Funktion (vgl. Gleichung 4.11 und Gleichungsfolge 4.12)

$$r_{ij} = -2.8926 + 5.7264/\left[1 + \exp\left(-\hat{k}_9^{(4)}\delta(KJ)_{ij}\right)\right] \tag{4.19}$$

Im Beispiel sind $\delta(KJ)_{i1} = 5$, $\delta(KJ)_{i2} = -4$, $\delta(KJ)_{i3} = -10$, $\delta(KJ)_{i4} = -1$, $\delta(KJ)_{i5} = -22$ und $\delta(KJ)_{i6} = 34$. Die Matrix C_i ist eine Design-Matrix für r_i

$$C_i = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.20}$$

Die Matrix B_i und der Vektor $\hat{b}_i$ werden weiter unten vorgestellt.

Mit den geschätzten Parameterwerten für die fixen Effekte $\hat{\beta}$, den fixen Effekten auf Baum-Ebene r_i und den Design-Matrizen A_i^* und C_i erhalten wir den Vektor mit Höhen-

Erwartungswerten zu den beobachteten Baumalterswerten nach Erweiterung von Gleichung 3.21:

$$f_i\left(\theta_i^*, A_i^*, \hat{\beta}, C_i, r_i, 0\right) = \begin{pmatrix} 24,791844 \\ 25,536488 \\ 26,276753 \\ 26,905399 \\ 28,839669 \\ 14,650997 \end{pmatrix} \tag{4.21}$$

Die Residuen zur Mittelwertkurve (mit Zufallsparameterwerten gleich 0) sind dann:

$$\hat{e}_i = y_i^* - f_i\left(\theta_i^*, A_i^*, \hat{\beta}, C_i, r_i, 0\right) = \begin{pmatrix} 2,1881554 \\ 1,8335116 \\ -1,206752 \\ 0,8146013 \\ 4,7603308 \\ -0,650995 \end{pmatrix} \tag{4.22}$$

.

Mit der a posteriori Kenntnis über die Kovarianzstruktur der Zufallsparameter ergibt sich $\hat{D}_i = diag\left(\hat{\Delta}_1, \hat{\Delta}_2, \hat{\Delta}_2, \hat{\Delta}_2, \hat{\Delta}_2, \hat{\Delta}_2, \hat{\Delta}_2\right)$ aus

$$\hat{\Delta}_1 = \begin{pmatrix} 0,0046129 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 9,8222569 \end{pmatrix} \tag{4.23}$$

und

$$\hat{\Delta}_2 = \begin{pmatrix} 0,003768 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2,799441 \end{pmatrix} \tag{4.24}$$

als ein Schätzer für die Kovarianzmatrix der Zufallsparameter am betreffenden Plot. Die partiellen Ableitungen in $\hat{F}_i$ ergeben $\hat{Z}_i$ über $\hat{Z}_i = \hat{F}_i^T B_i$. Dabei ist B_i die Design-Matrix für die Zufallsparameter.

$$\hat{F}_i = \begin{pmatrix} 5,7465566 & 0 & 0 & 0 & 0 & 0 \\ -12,12959 & 0 & 0 & 0 & 0 & 0 \\ 0,992321 & 0 & 0 & 0 & 0 & 0 \\ 0 & 11,397862 & 0 & 0 & 0 & 0 \\ 0 & -24,498 & 0 & 0 & 0 & 0 \\ 0 & 0,9946529 & 0 & 0 & 0 & 0 \\ 0 & 0 & 14,74725 & 0 & 0 & 0 \\ 0 & 0 & -32,03724 & 0 & 0 & 0 \\ 0 & 0 & 0,9965489 & 0 & 0 & 0 \\ 0 & 0 & 0 & 12,431086 & 0 & 0 \\ 0 & 0 & 0 & -26,81699 & 0 & 0 \\ 0 & 0 & 0 & 0,9801041 & 0 & 0 \\ 0 & 0 & 0 & 0 & 20,349834 & 0 \\ 0 & 0 & 0 & 0 & -45,04619 & 0 \\ 0 & 0 & 0 & 0 & 0,9759064 & 0 \\ 0 & 0 & 0 & 0 & 0 & -15,63217 \\ 0 & 0 & 0 & 0 & 0 & 30,552136 \\ 0 & 0 & 0 & 0 & 0 & 0,9384171 \end{pmatrix} \tag{4.25}$$

$$B_i = \begin{pmatrix}
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix} \tag{4.26}$$

Mit der Annahme normalverteilter und unabhängiger Restfehler und der a posteriori Kenntnis über die geschätzte Varianz des Restfehlers $\sigma^2 = 0.1331681$ erhalten wir für alle $n_i = 6$ Bäume an dem Plot dieses Beispiels die näherungsweise vorhergesagte Fehler-Matrix

$$\begin{aligned}
\hat{R}_i &= \sigma^2 I\,[n_i] \\
&= \begin{pmatrix}
0,1331681 & 0 & 0 & 0 & 0 & 0 \\
0 & 0,1331681 & 0 & 0 & 0 & 0 \\
0 & 0 & 0,1331681 & 0 & 0 & 0 \\
0 & 0 & 0 & 0,1331681 & 0 & 0 \\
0 & 0 & 0 & 0 & 0,1331681 & 0 \\
0 & 0 & 0 & 0 & 0 & 0,1331681
\end{pmatrix}
\end{aligned} \tag{4.27}$$

Mit den zuvor beschriebenen Matrizen können nun nach Gleichung 3.24 die Zufallsparameter näherungsweise vorhergesagt werden. Wir erhalten im vorliegenden Beispiel folgendes Vorhersageergebnis für die Zufallsparameterwerte:

$$\hat{b}_i \simeq \hat{D}_i \hat{Z}_i^T \left(\hat{Z}_i \hat{D}_i \hat{Z}_i^T + \hat{R}_i\right)^{-1} \hat{e}_i = \begin{pmatrix} 0,0415261 \\ 0 \\ 0,8732173 \\ 0,0077799 \\ 0 \\ 0,9981204 \\ 0,0062245 \\ 0 \\ 0,4035658 \\ -0,040034 \\ 0 \\ -2,009947 \\ -0,007669 \\ 0 \\ -0,449252 \\ 0,0538732 \\ 0 \\ 1,9194767 \\ 0,0137462 \\ 0 \\ -0,613089 \end{pmatrix} = \begin{pmatrix} \hat{z}_i^{(2)} \\ 0 \\ \hat{z}_i^{(4)} \\ \hat{z}_{i1}^{(2)} \\ 0 \\ \hat{z}_{i1}^{(4)} \\ \hat{z}_{i2}^{(2)} \\ 0 \\ \hat{z}_{i2}^{(4)} \\ \hat{z}_{i3}^{(2)} \\ 0 \\ \hat{z}_{i3}^{(4)} \\ \hat{z}_{i4}^{(2)} \\ 0 \\ \hat{z}_{i4}^{(4)} \\ \hat{z}_{i5}^{(2)} \\ 0 \\ \hat{z}_{i5}^{(4)} \\ \hat{z}_{i6}^{(2)} \\ 0 \\ \hat{z}_{i6}^{(4)} \end{pmatrix} \tag{4.28}$$

Darin sind die ersten drei Zeilen die Zufallsparameter $\left(\hat{z}_i^{(2)}, 0, \hat{z}_i^{(4)}\right)^T$ auf Plot-Ebene. Darunter schließen sich $n_i = 6$ Blocks aus den Zufallsparametern für die Baum-Ebene $\left(\hat{z}_{ij}^{(2)}, 0, \hat{z}_{ij}^{(4)}\right)^T$ mit $j = 1, \cdots, 6$ an. Nach der Konstruktionsweise des endgültigen Modells enthalten nur die Funktionsparameter $\psi_{ij}^{(2)}$ und $\psi_{ij}^{(4)}$ Zufallskomponenten, nicht der

Parameter $\psi_{ij}^{(3)}$ (vgl. Gleichungsfolge 4.12); daraus resultieren die 0-Zeilen in $\hat{b}_i$. Die vorhergesagten Funktions-Parameterwerte ergeben sich schließlich mit

$$\hat{\phi}_i = A_i^* \hat{\beta} + C_i r_i + B_i \hat{b}_i = \begin{pmatrix} \hat{\psi}_{i1}^{(2)} \\ \hat{\psi}_{i1}^{(3)} \\ \hat{\psi}_{i1}^{(4)} \\ \hat{\psi}_{i2}^{(2)} \\ \hat{\psi}_{i2}^{(3)} \\ \hat{\psi}_{i2}^{(4)} \\ \hat{\psi}_{i3}^{(2)} \\ \hat{\psi}_{i3}^{(3)} \\ \hat{\psi}_{i3}^{(4)} \\ \hat{\psi}_{i4}^{(2)} \\ \hat{\psi}_{i4}^{(3)} \\ \hat{\psi}_{i4}^{(4)} \\ \hat{\psi}_{i5}^{(2)} \\ \hat{\psi}_{i5}^{(3)} \\ \hat{\psi}_{i5}^{(4)} \\ \hat{\psi}_{i6}^{(2)} \\ \hat{\psi}_{i6}^{(3)} \\ \hat{\psi}_{i6}^{(4)} \end{pmatrix} = \begin{pmatrix} 0,5787831 \\ 0,856212 \\ 22,608394 \\ 0,5772276 \\ 0,856212 \\ 21,583629 \\ 0,5309687 \\ 0,856212 \\ 18,967416 \\ 0,5633337 \\ 0,856212 \\ 20,874049 \\ 0,6248763 \\ 0,856212 \\ 22,7713 \\ 0,5847494 \\ 0,856212 \\ 21,29792 \end{pmatrix} \tag{4.29}$$

Der Vektor mit Vorhersagen für die Baumhöhen zu beliebigen Alterswerten in θ_i resultiert aus:

$$\hat{y}_i = f_i \left(\theta_i, A_i^* \hat{\beta} + C_i r_i + B_i \hat{b}_i \right) \tag{4.30}$$

.

Die Vorhersage der Zufallsparamter für die Fortschreibungen im Rahmen der Modellevaluierung ist für das gleiche Beispiel zusätzlich in den Abbildungen 4.21 bis 4.24 bildlich dargestellt. In Abbildung 4.21 sind sämtliche Höhenmessungen aus den Trieblängenanalysen für die 6 Bäume des Stichprobenpunktes aufgezeigt. Die Bäume waren zum Zeitpunkt

der Fällung sehr verschieden alt. Das Alter der Bäume reichte bei der Aufnahme von 54 bis 110 Jahren. Für die Kalibrierung nach der ersten Variante wurden an allen Stichproben die jeweils ersten Beobachtungen aller Bäume als bekannt angenommen. Dieses ist durch die Punkt-Symbole in Abbildung 4.22 dargestellt. Die grüne Linie in dieser Abbildung ist der Höhenwachstumsverlauf, den man für einen Baum an dieser Stichprobe erwarten würde, wenn keine Höhenbeobachtungen für die Vorhersage der Zufallsparameter vorliegen würden. Wenn die Bäume an einem Stichprobenpunkt in unterschiedlichen Jahren gekeimt sind, kann zusätzlich für jeden Baum die Kenntnis über die Differenz seines Keimjahres zum durchschnittlichen Keimjahr am Stichprobenpunkt genutzt werden. Aufgrund der fixen Effekte auf Stichproben- und Baum-Ebene kann man den individuellen Höhenwachstumsverlauf für die einzelnen Bäume vermuten. Auf eine bildliche Darstellung wurde zur Wahrung der Übersichtlichkeit verzichtet. Das Ziel der Kalibrierung ist die Vorhersage der Zufallsparameter. Dazu werden die als bekannt angenommen Beobachtungen der einzelnen Bäume mit den Höhenwerten verglichen, die man zu den beobachteten Baumaltern aufgrund der fixen Effekte erwarten darf. Dadurch erhalten wir den Vektor $\hat{e}_i$. Dieser beschreibt die gesamte Abweichung, die nicht durch die bekannten Kovariablen erklärt werden kann. Diese restliche Abweichung wird zum einen durch die Zufallseffekte auf Ebene der Stichprobe und der Bäume und zum anderen durch den Restfehler verursacht. Anschaulich gesprochen gibt es Einflussgrößen, die uns nicht bekannt sind und die neben den bekannten Kovariablen zusätzlich Variation bewirken. Diese unbekannten Effekte können entweder dem Stichprobenort oder dem einzelnen Baum zugeordnet werden. Die Zufallseffekte auf Ebene der Stichprobenorte verursachen die Variation, die zwischen den verschiedenen Stichprobenpunkten herrscht; die Zufallseffekte auf Baum-Ebene verursachen die Variation innerhalb eines Stichprobenpunktes bzw. zwischen den Bäumen eines Punktes. Daneben haben wir bei der Modellbildung angenommen, dass es eine normalverteilte Restfehlergröße gibt. Diese bewirkt, dass die beobachteten Messwerte der Bäume um die geschätzte Höhenkurve herum oszillieren. Diese Oszillation kann zum einen von Messfehlern herrühren oder von uns unbekannten klimatischen, standörtlichen oder waldbaulichen Einflüssen. Bei der Modellanpassung haben wir die Varianzen der Zu-

fallsparameter und des Restfehlers geschätzt. Mit der Kenntnis über die Verteilungen der Zufallsparameter und des Restfehlers und dem abgeschätzten Ausmaß der unerklärlichen Variation in $\hat{e}_i$ können die Zufallsparameter selbst vorhergesagt werden. Diese werden über Gleichung 3.24 nacheinander für die Stichproben linear approximiert. Dieser Prozess ist äußerst effizient und beanspruchte für alle Bäume und Stichproben des Evaluierungsdatensatzes nur minimale Rechenzeiten. Mit den vorhergesagten Zufallsparametern und den fixen Effekten kann nun für jeden Baum eine Wachstumskurve zur Prognose geschätzt werden. Diese individuellen Schätzkurven sind in Abbildung 4.23 und 4.24 gezeichnet. In Abbildung 4.24 sind die als bekannt angenommenen Beobachtungen, anhand derer die Kalibrierung vorgenommen wurde, als Punkte dargestellt und die restlichen Beobachtungen als „+“-Symbole eingeblendet.

In Abbildung 4.25 sind die geschätzten Wachstumsverläufe für den gleichen Stichprobenpunkt dargestellt, wenn zusätzlich zu den ersten noch die elften Beobachtungen der Bäume als bekannte Vorinformationen vorliegen würden.

Die Güte des Modells soll anhand folgender Fragestellungen beurteilt werden:

1. Sind die geschätzten Wachstumsverläufe biologisch plausibel?
2. Sind die Prognosefehler akzeptabel klein und unverzerrt?
3. Ist die neue Vorhersagemethode besser als herkömmliche?

Zur Beantwortung der ersten Frage sollen die beobachteten und vorhergesagten Wachstumsverläufe verglichen werden. In Abbildung 4.26 sind die Beobachtungen und Vorhersagen der Baumhöhen dargestellt. Im unteren Teil der Abbildung kann man erkennen, dass die Höhenwachstumsverläufe aller Bäume biologisch plausibel geschätzt werden. Die Vorhersagen scheinen jedoch für einige wenige Bäume nicht die Wirklichkeit zu treffen. Es fallen drei Bäume ins Auge, die im relativ hohen Baumalter tatsächlich einen sehr steilen Wachstumsverlauf zeigen. Die geschätzten Verläufe sind bei beiden Fortschreibungsvarianten jedoch viel flacher. Für die Fortschreibungsvariante mit lediglich der ersten als bekannt angenommenen Beobachtung ist die Ursache leicht erklärbar. Bei diesen Bäumen ist der jeweils erste Höhenbeobachtungswert trotz des hohen Alters vergleichsweise gering.

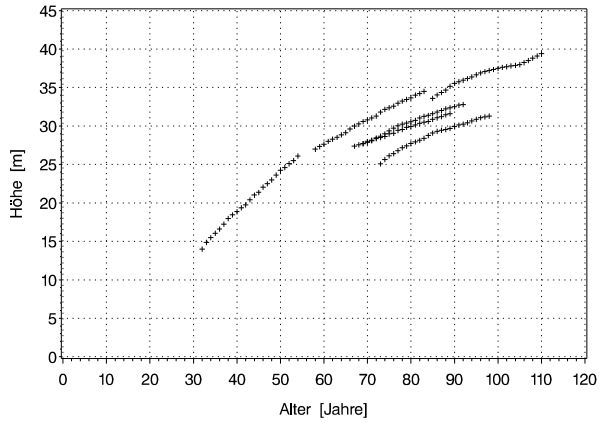


Abbildung 4.21: *Baumhöhenmessungen an einem Stichprobenpunkt.*

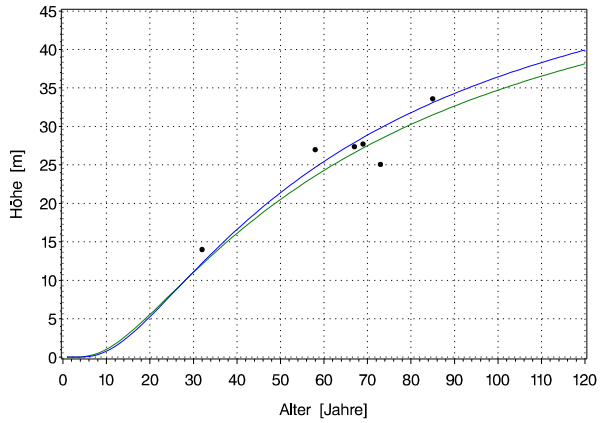


Abbildung 4.22: *Mittelwertkurven und erste Beobachtungen zur Kalibrierung.*
Grün: Stichproben-Mittelwertkurve nur aufgrund von fixen Effekten
Blau: Stichproben-Mittelwertkurve mit vorhergesagten Zufallsparametern für die Stichprobe.

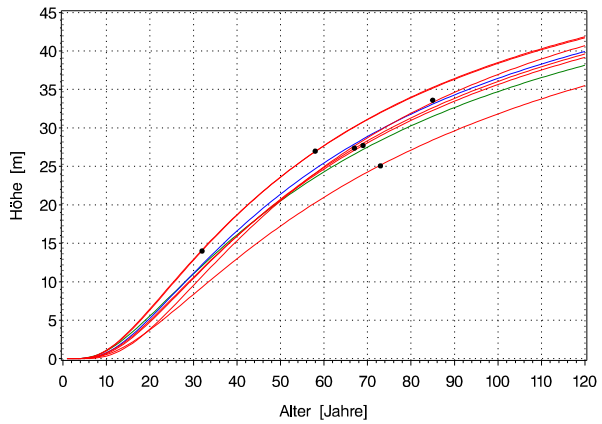


Abbildung 4.23: *Rot: Individuelle Kurven mit vorhergesagten Zufallsparametern für die Stichproben- und Baum-Ebene.*

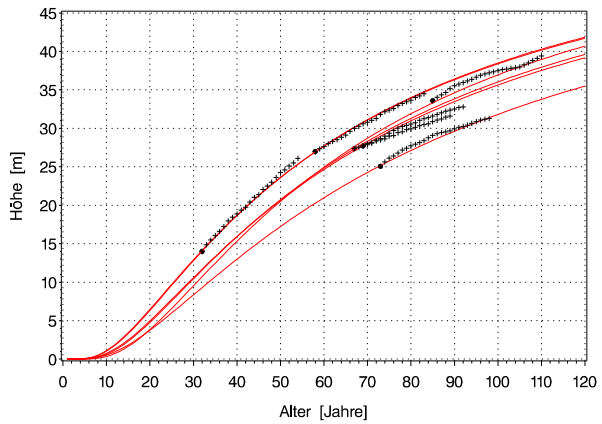


Abbildung 4.24: *Individuelle Kurven und alle Beobachtungen.*
Punkte: Beobachtungen zur Kalibrierung.

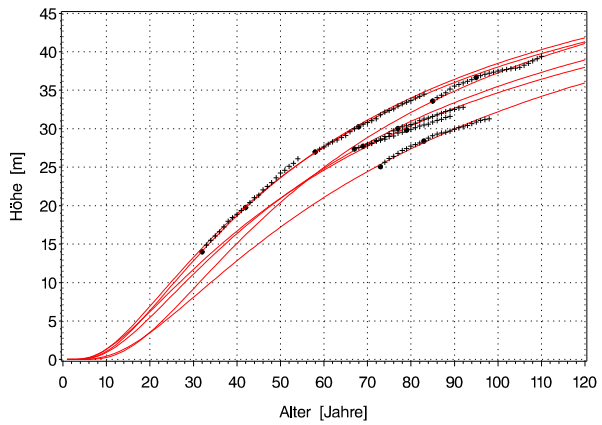


Abbildung 4.25: *Kalibrierung anhand von zwei Vorbeobachtungen.*

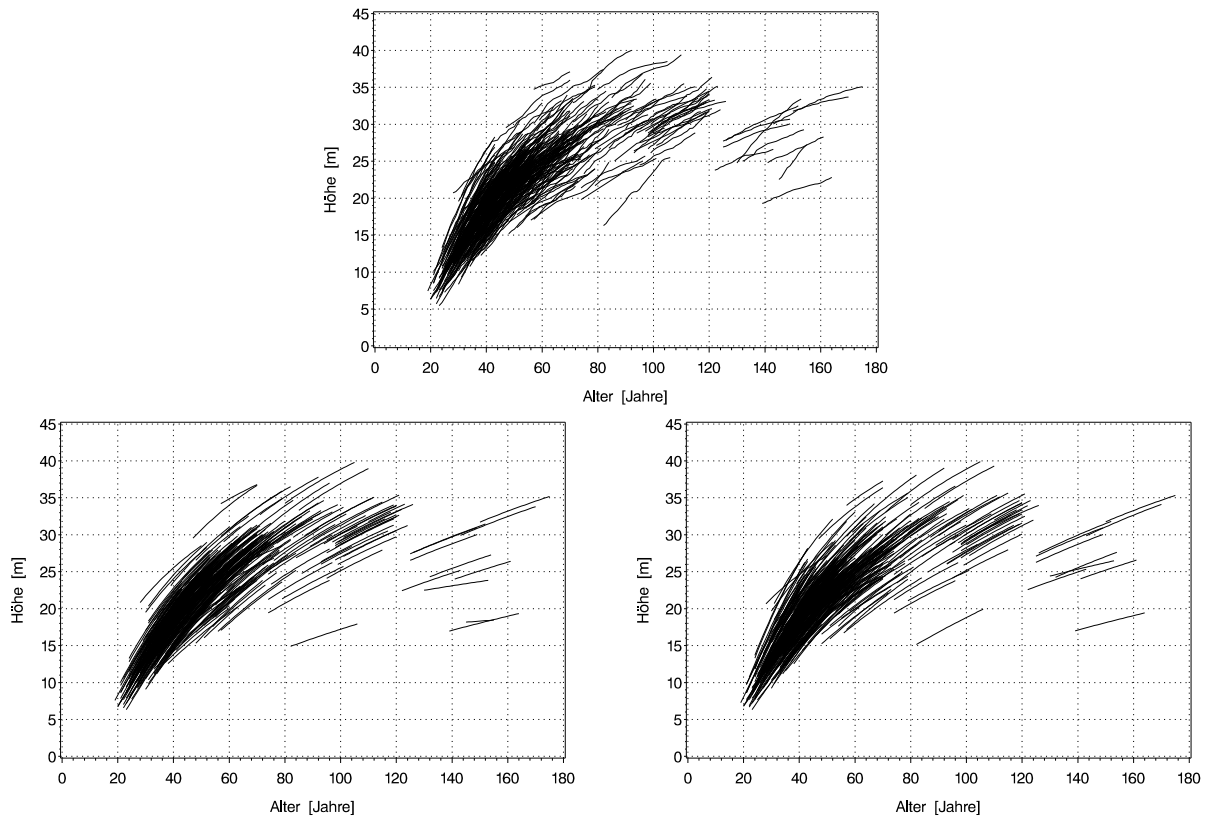


Abbildung 4.26: *Beobachtungen und Vorhersagen der Kalibrierungsdaten. Oben: Baumhöhenbeobachtungen der ZF-Daten. Unten: Vorhersagen für die Baumhöhen aufgrund der Kalibrierung mit BLUPs für die Zufallsparameter. Unten links: Kalibrierung anhand der jeweils ersten Beobachtung eines Baumes. Unten rechts: Kalibrierung anhand der ersten und elften Beobachtungen.*

Da nur der erste Beobachtungswert zur Kalibrierung herangezogen wurde, wird für diese Bäume ein Wachstumsverlauf prognostiziert, der sich weiterhin auf geringem Niveau befindet. Für die BLUPs gab es aufgrund der Kenntnis über den ersten niedrigen Höhenbeobachtungswert keinen Anlass, eine derartige Wachstumsbeschleunigung anzunehmen. Mit zwei Vorbeobachtungen zur Kalibrierung würde man eine wirklichkeitsgetreuere Vorhersage vermuten. Die Situation an diesem Stichprobenpunkt wird im linken Teil der Abbildung 4.27 dargestellt. Das Wachstumsverhalten des Baumes weicht extrem vom durchschnittlichen Verhalten am Stichprobenpunkt ab. Offensichtlich können die BLUPs der Zufallsparameter im Rahmen der zuvor ermittelten Varianzen auf Baum-Ebene diesen Ausreißer nicht abbilden. Man könnte darüber hinaus vermuten, dass bei diesem Baum die Alterswerte fehlerhaft erfasst wurden. Mit einem negativen Versatz von 80 Jahren würden sich die Beobachtungen gut in die der anderen Bäume einfügen. Ähnlich problematisch scheint die Prognose für einen Baum zu sein, dessen erster Höhenbeobachtungswert im jungen Alter von 28 Jahren mit 20,7 m Höhe sehr hoch ist. Nach Vorhersage der Zufallsparameter lediglich anhand dieser einen Beobachtung wird für diesen Baum der weitere Wachstumsverlauf viel höher als in Wirklichkeit geschätzt. Wenn die Kalibrierung zusätzlich anhand der elften Beobachtung vorgenommen wird, können die BLUPs der Zufallsparameter dieses stagnierende Wachstum beschreiben (s. rechten Teil der Abbildung 4.27).

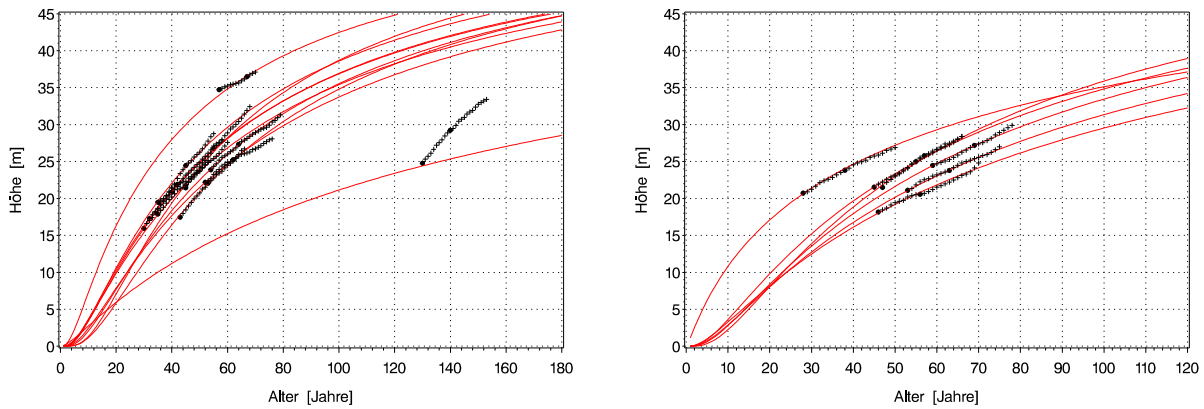


Abbildung 4.27: *Betrachtung einzelner Stichprobenpunkte hinsichtlich plausibler Vorhersagen. Links: Ungenaue Vorhersagen aus BLUPs für einen Baum, der stark vom Durchschnitt am Stichprobenpunkt abweicht. Rechts: Flexible und wirklichkeitsgetreue Vorhersagen für alle Bäume.*

Im folgen sollen die beiden weiteren Fragen geklärt werden; ob der Prognose-Fehler in einem akzeptablen Rahmen liegt und ob das auf Zufallsparametern basierende Modell einen methodischen Gewinn darstellt. Zur Untersuchung dieser Sachverhalte wird ein Modell als Referenz für einen Vergleich benötigt. Es liegt nahe für ein Vergleichsmodell auch die Trendfunktion nach Sloboda (1971) zu verwenden. Üblicherweise werden Wachstumsfächer zur Beschreibung der Höhenentwicklung von Oberhöhenbäumen verwendet. Bei variierenden Werten für den freien Bonitätsparameter können sich die resultierenden Kurven nicht schneiden. Vor dem Hintergrund, dass an späterer Stelle das auf Zufallsparametern basierende HNLMM dieser Studie mit den herkömmlichen Bonitätsfächern verglichen wird, sollen jedoch hier ausgehend von bestimmten Alters- und Höhenwerten die weiteren Höhen von einzelnen Bäumen mit einem Bonitätsfächer prognostiziert werden. Die Parameter sollen über ein empirisches Richtungsfeld von durchschnittlichen Höhenzuwächsen in Alters- und Höhenklassen nach der Methode von Wolf (1957) geschätzt werden. Das Vergleichsmodell wird dafür einmal an alle aktuelle IWE-Daten angepasst, die dieser Studie zugrunde lagen und einmal nur an die IWE-Daten aus dem Südschwarzwald. Wir wollen also das Vergleichsmodell mit zwei unterschiedlichen Modelldaten als Referenz heranziehen. Mit dem Bonitätsfächer der Trendfunktion können nicht wie bei der Kalibrierung mit BLUPs beliebig viele Vorbeobachtungen ausgenutzt werden. Von der Trendfunktion kann jeweils nur eine Beobachtung als Startinformation zur Fortschreibung verwendet werden. Zur Vergleichbarkeit soll deshalb in Anlehnung an die Auswahl der Vorinformationen bei der BULP-Kalibrierung in einer Fortschreibungsvariante nur die jeweils erste Beobachtung zugrunde gelegt werden und in einer weiteren nur die elfte.

Zunächst wurden alle Alterswerte t in 10-jährige Altersklassen eingeteilt und alle Höhenwerte H in Höhenklassen mit einer Stufenbreite von 2 m:

$$\begin{aligned}
&H_g = integer\left[\frac{H}{2}\right]\cdot 2 + integer\left[H - integer\left[\frac{H}{2}\right]\cdot 2\right]\cdot 2, \forall\ H >= 3\\
&H_g = 2, \forall\ H < 3\\
&t_g = integer\left[\frac{t}{10}\right]\cdot 10 + integer\left[\frac{t - integer\left[\frac{t}{10}\right]\cdot 10}{5}\right]\cdot 10, \forall\ t >= 15\\
&t_g = 10, \forall\ t < 15
\end{aligned} \tag{4.31}$$

Danach wurde für jeden Baum j aller n_{t_g,H_g} Bäume für sämtliche Alters-Höhen-Klassen der durchschnittliche jährliche Höhenzuwachs iH_{j,t_g,H_g} ermittelt über:

$$\begin{aligned}
&\delta H_{j,t_g,H_g} = max\left[H_{j,t_g,H_g}\right] - min\left[H_{j,t_g,H_g}\right]\\
&\delta t_{j,t_g,H_g} = max\left[t_{j,t_g,H_g}\right] - min\left[t_{j,t_g,H_g}\right]\\
&iH_{j,t_g,H_g} = \frac{\delta H_{j,t_g,H_g}}{\delta t_{j,t_g,H_g}}\\
&j = 1\ldots, n_{t_g,H_g}
\end{aligned}$$

Mit den Mittelwerten

$$iH_{t_g,H_g} = \frac{1}{n_{t_g,H_g}}\sum_{j=1}^{n_{t_g,H_g}} iH_{j,t_g,H_g} \tag{4.32}$$

dieser durchschnittlichen jährlichen Höhenzuwächse in den Alters-Höhen-Klassen wurden dann die empirischen Richtungsfelder aufgespannt. Über OLS mit Hilfe der `SAS 9.1`-Prozedur `NLIN` wurde dann die Differentialform der Trendfunktion (Gleichung 4.33) an das Richtungsfeld angepasst.

$$iH_{t_g,H_g} = \psi_2 \cdot H_g/(t_g^{\psi_3}) \cdot log(65^{\psi_1}/H_g) \tag{4.33}$$

Mit

$$\begin{aligned}
&\frac{\partial iH_{t_g,H_g}}{\partial \psi_3} = -\log(t_g)\cdot\psi_2\cdot H_g/t_g^{\psi_3}\cdot\log(65^{\psi_1}/H_g)\\
&\frac{\partial iH_{t_g,H_g}}{\partial \psi_2} = H_g/t_g^{\psi_3}\cdot\log(65^{\psi_1}/t_g)\\
&\frac{\partial iH_{t_g,H_g}}{\partial \psi_1} = \psi_2\cdot H_g/t_g^{\psi_3}\cdot\log(65)
\end{aligned}$$

wurden die partiellen Ableitungen nach den Parametern vorgegeben.

Da manche Alters-Höhen-Klassen schwach besetzt sind, soll bei einem Anpassungsvorgang das Richtungsfeld nur aus Klassen mit mehr als fünf Beobachtungen ausgeglichen werden. Da auch untersucht werden soll, wie sich der daraus resultierende Ausschluss von Beobachtungen auswirkt, werden in einem weiteren Anpassungsvorgang alle Felder mit mindestens einer Beobachtung verwendet.

Genau genommen erhalten wir mit vier verschiedenen Bonitätsfächern als Vergleichsmodelle und zwei Varianten von Vorinformationen acht verschiedene Fortschreibungsergebnisse, je nach dem

1. für welche Datengrundlage der Bonitätsfächer angepasst wird:
   - für alle IWE-Daten
   - nur für die IWE-Daten aus dem Südschwarzwald
2. welche Felder beim Anpassungsvorgang herangezogen werden:
   - Alters- und Höhenklassen mit mindestens einer Beobachtung
   - Alters- und Höhenklassen mit mindestens fünf Beobachtungen
3. welche Vorinformationen zur Fortschreibung gewählt werden:
   - nur die jeweils erste Beobachtung eines Baumes
   - nur die jeweils elfte Beobachtung eines Baumes

Abbildung 4.28 zeigt die vier angepassten Bonitätsfächer. Außerdem zeigen die Grafiken die zugrunde liegenden Beobachtungen und die empirischen Richtungsfelder. Die erzielten Parameterschätzwerte sind in Tabelle 4.4 aufgeführt.

In Abbildung 4.29 und 4.30 sind die Vorhersagewerte für die Baumhöhenprognosen dargestellt, die aus der Fortschreibung der ersten bzw. elften Beobachtung mit den verschiedenen Bonitätsfächern resultieren. Bei der Fortschreibungsvariante mit der elften

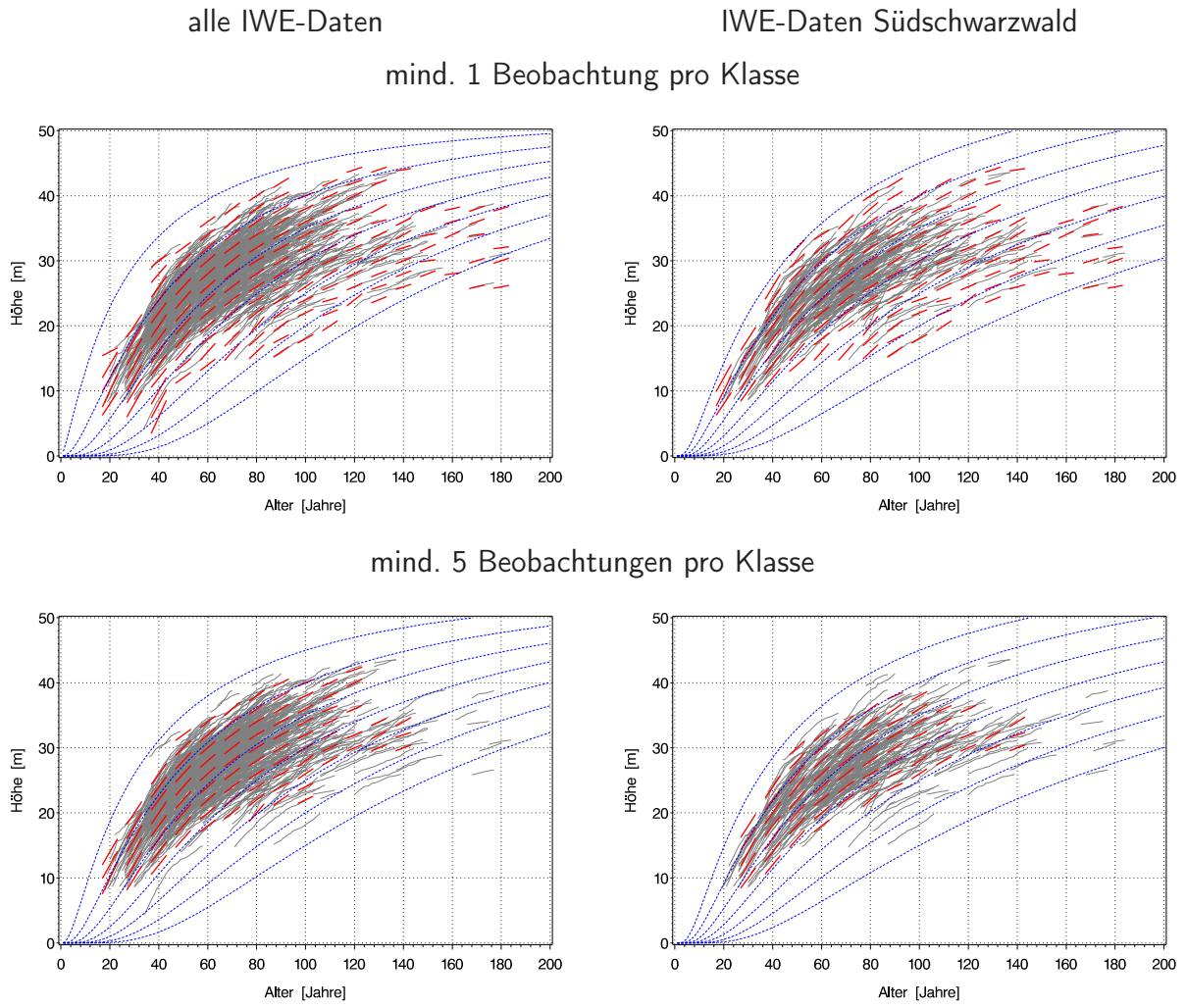


Abbildung 4.28: *Ausgleich empirischer Richtungsfelder. Links: Anpassung für alle IWE-Daten. Rechts: Anpassung für die IWE-Daten aus dem Südschwarzwald. Oben: Verwendung aller Alters-Höhenklassen mit mindestens einer Beobachtung. Unten: Verwendung von Alters-Höhenklassen mit fünf oder mehr Beobachtungen. Graue Linien im Hintergrund: Baumhöhenbeobachtungen. Rote Linien: durchschnittliche jährliche Höhenzuwächse in den Alters-Höhen-Klassen (empirische Richtungsfelder). Blaue unterbrochene Linien: angepasste Bonitätsfächer.*

	alle IWE-Daten		IWE-Daten aus Südschwarzwald	
	mind. 1 Beob.	mind. 5 Beob.	mind. 1 Beob.	mind. 5 Beob.
ψ_1	0.94786	0.96507	1.01094	0.99886
ψ_2	0.33968	0.45256	0.59629	0.70545
ψ_3	0.70298	0.79368	0.91463	0.94414

Tabelle 4.4: *Parameterschätzwerte aus der Anpassung des empirischen Richtungsfelds.*

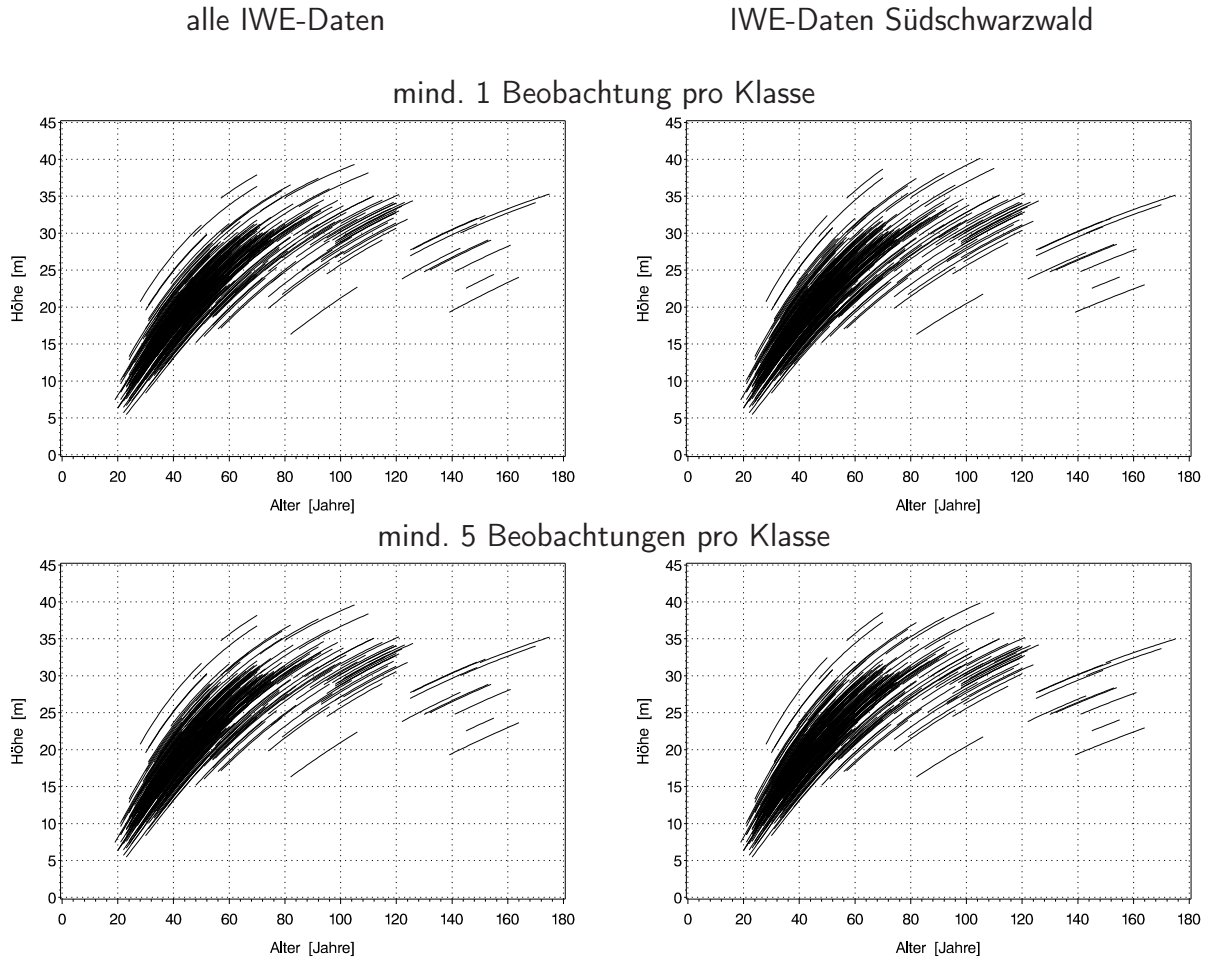


Abbildung 4.29: *Vorhersagen für die Daten des Projektes „Zukunftsorientierte Forstwirtschaft“ mit Hilfe der angepassten Bonitätsfächer mit der ersten Beobachtung als Startinformation. Links: Schätzungen mit dem Fächer, der an die IWE-Daten aus allen Regionen angepasst wurde. Rechts: Schätzungen mit dem Fächer für die IWE-Daten aus der Region Südschwarzwald / Baar-Wutach. Oben: Der Fächer wurde an alle Alters-Höhen-Klassen mit mindestens einer Beobachtung angepasst. Unten: Der zugrunde liegende Fächer wurde nur an Klassen mit mindestens fünf Beobachtungen angepasst.*

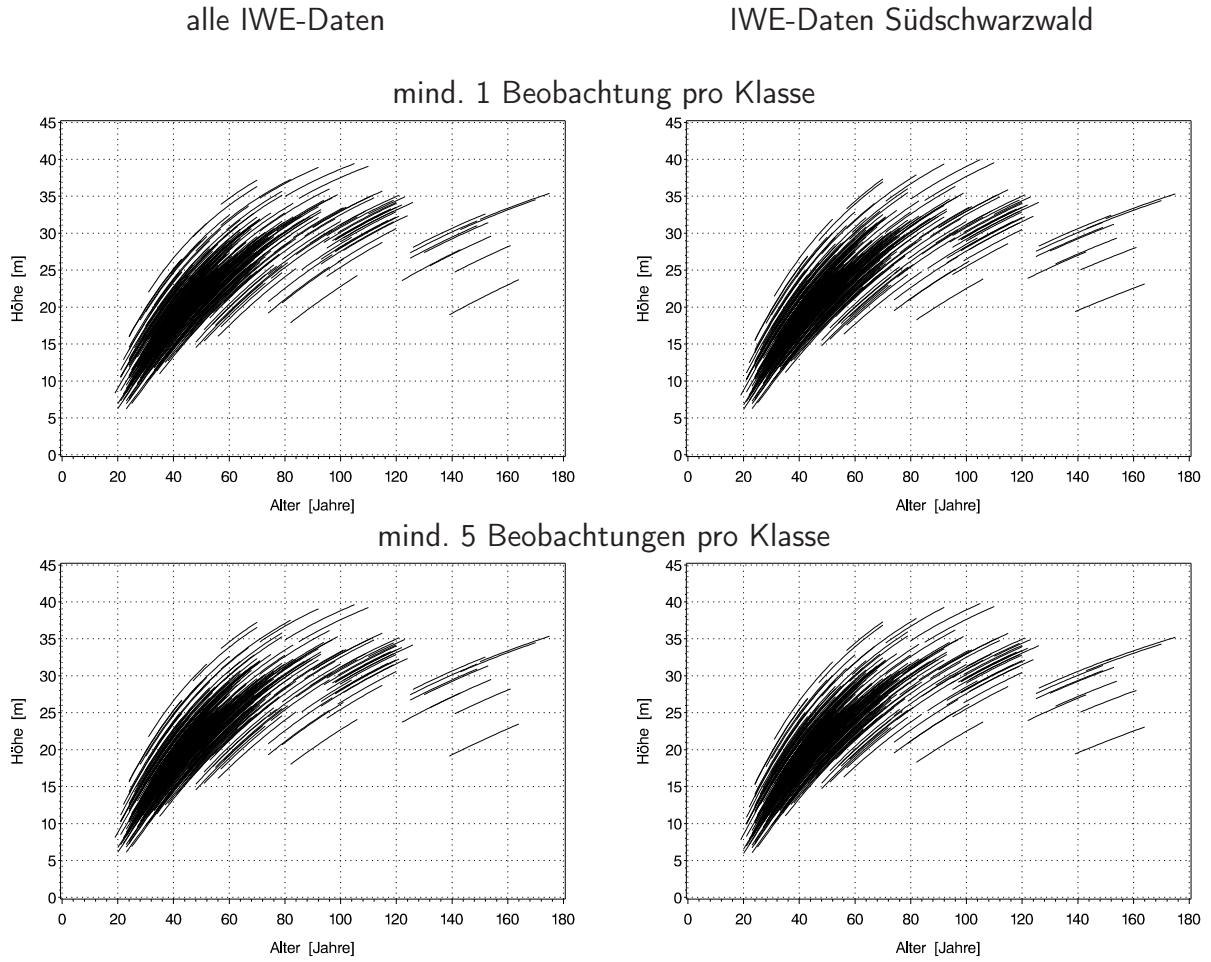


Abbildung 4.30: *Vorhersagen für die Daten des Projektes „Zukunftsorientierte Forstwirtschaft“ mit Hilfe der angepassten Bonitätsfächer mit der elften Beobachtung als Startinformation. Links: Schätzungen mit dem Fächer, der an die IWE-Daten aus allen Regionen angepasst wurde. Rechts: Schätzungen mit dem Fächer für die IWE-Daten aus der Region Südschwarzwald / Baar-Wutach. Oben: Der Fächer wurde an alle Alters-Höhen-Klassen mit mindestens einer Beobachtung angepasst. Unten: Der zugrunde liegende Fächer wurde nur an Klassen mit mindestens fünf Beobachtungen angepasst.*

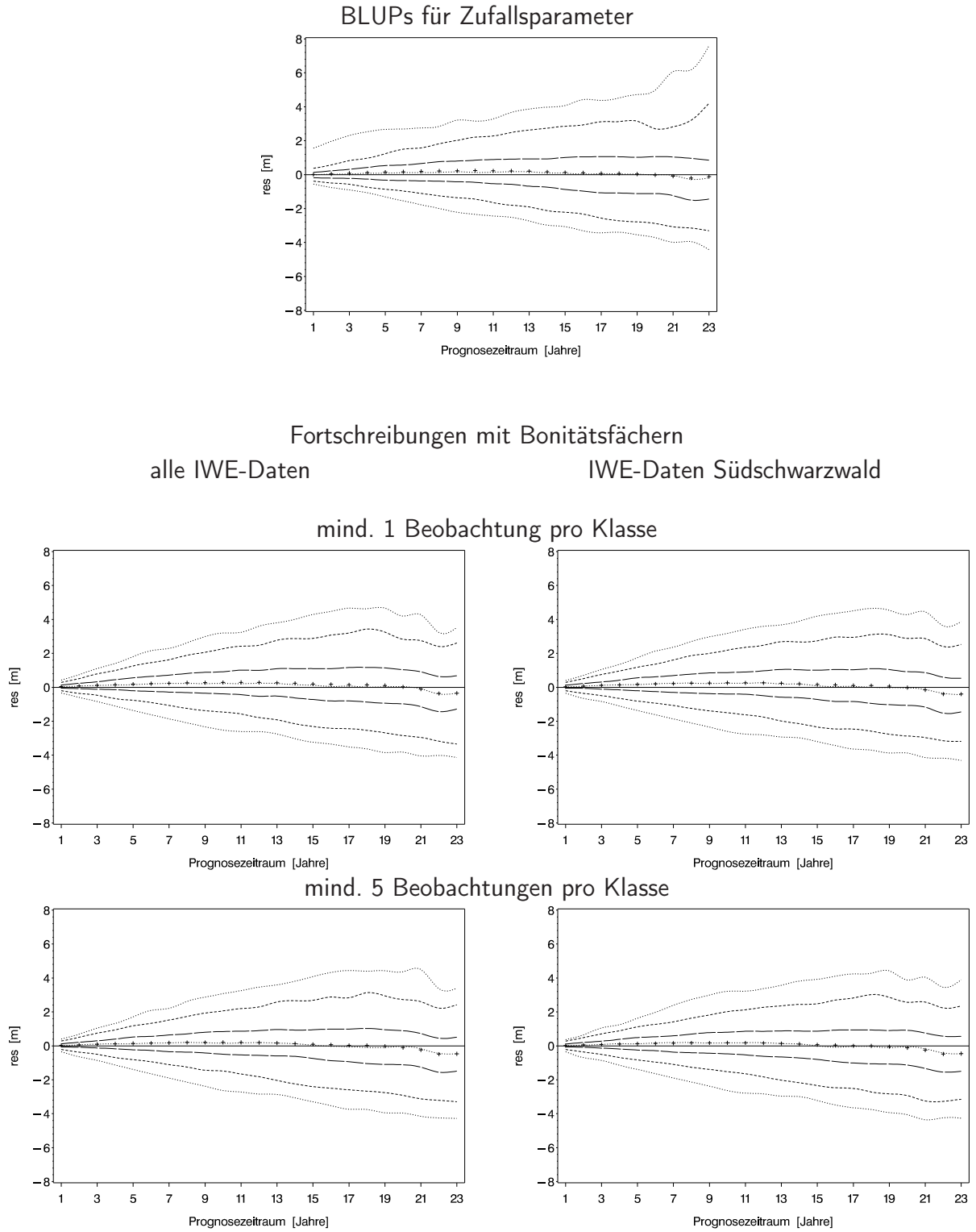


Abbildung 4.31: *Residuen der Baumhöhen-Prognosewerte mit der ersten Beobachtung als Startinformation. In den Grafiken sind die Perzentile der Residuen aus den Fortschreibungen für 1-jährige Klassen von Prognosezeiträumen dargestellt. Durch Linien verbunden sind in der Reihenfolge von oben nach unten: der P99, P95, Q3, Median, Q1, P5 und P1. „+"-Symbole: Mittelwerte der Residuen in den Klassen.*

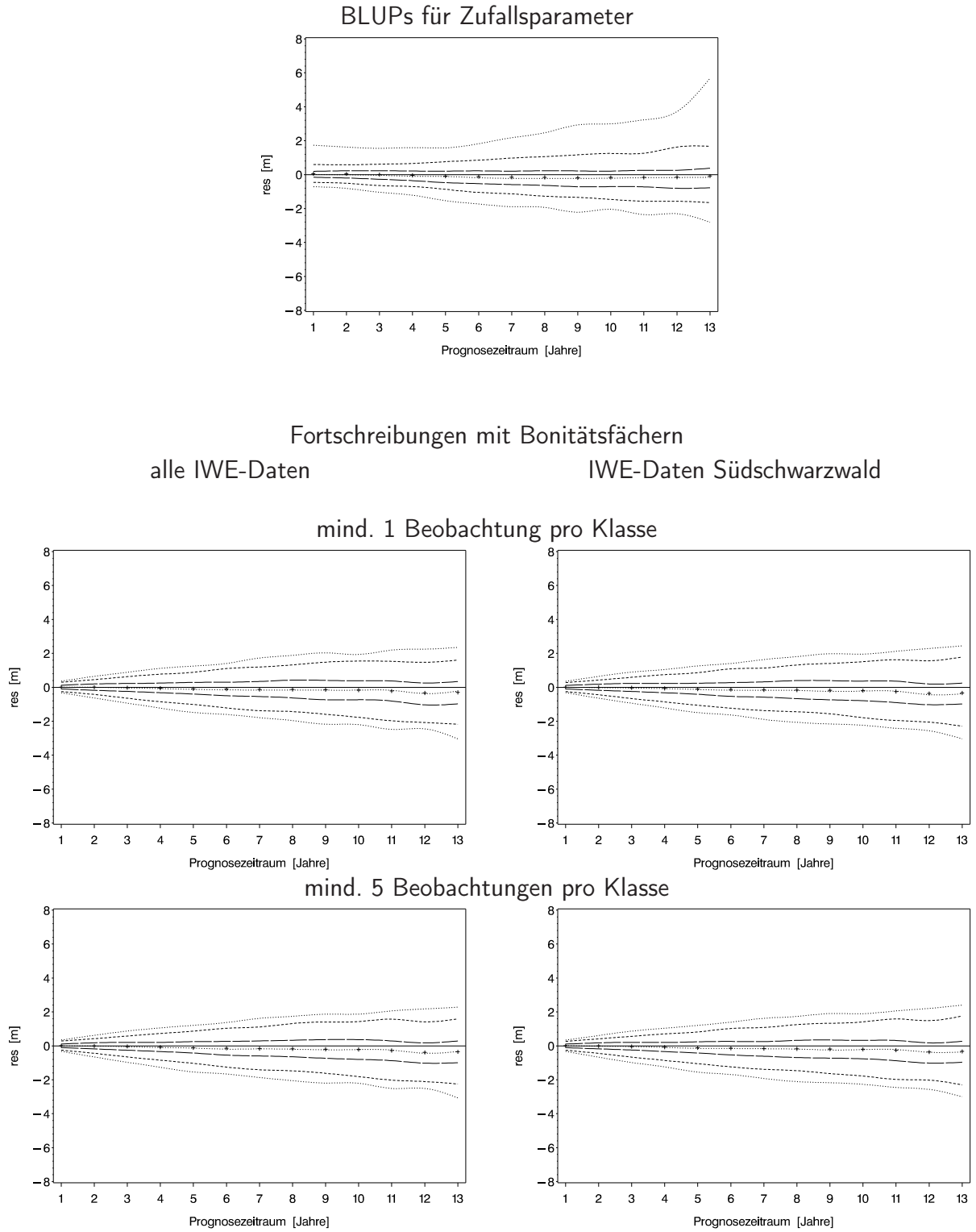


Abbildung 4.32: *Residuen der Baumhöhen-Prognosewerte mit der elften Beobachtung als Startinformation. In den Grafiken sind die Perzentile der Residuen aus den Fortschreibungen für 1-jährige Klassen von Prognosezeiträumen dargestellt. Durch Linien verbunden sind in der Reihenfolge von oben nach unten: der P99, P95, Q3, Median, Q1, P5 und P1. „+"-Symbole: Mittelwerte der Residuen in den Klassen.*

Beobachtung als Startinformation sind auch die Vorhersagewerte für zurückliegende Baumalter dargestellt. Die prognostizierten Höhenwachstumsverläufe erscheinen alle biologisch plausibel. Dieses ist nicht selbstverständlich, da aus Höhenfortschreibungen zur Bonitierung von Inventurdaten Fälle bekannt sind, bei denen sehr junge Bäume mit außergewöhnlich hohen Baumhöhen in den Extrapolationsbereich der Fächer gelangen können. Dort können die Kurvenverläufe kippen und negatives Wachstum beschreiben. Im okularen Vergleich scheinen sich die geschätzten Kurvenverläufe der verschiedenen Anpassungsvarianten nicht wesentlich zu unterscheiden. Aus dem Vergleich der Abbildung 4.26 mit den Abbildungen 4.29 und 4.30 wird ein Vorteil des auf Zufallsparametern basierenden gemischten Modell deutlich. Die Wirklichkeit kann flexibler abgebildet werden. Während der reine Bonitätsfächer ein relativ starres Wachstumsverhalten vorgibt, kann das gemischte Modell sogar sich überschneidende Wachstumsverläufe verschiedener Bäume beschreiben (siehe dazu auch die Abbildungen 4.24 und 4.27).

Zur Beurteilung der Treffgenauigkeit sollen die Verteilungen der Residuen der Schätzwerte zu den Beobachtungen über die Prognosezeiträume herangezogen werden. Die Verteilungen der Residuen sind den Abbildungen 4.31 und 4.32, sowie den Tabellen 8.5, 8.6 und 8.7 des Anhangs zu entnehmen.

Zunächst werden die Residuen der verschiedenen Modelle verglichen, die aus der Fortschreibung mit den ersten Beobachtungen als bekannte Startinformationen resultieren. Beim Vergleich der Grafiken in Abbildung 4.31 scheinen die Residuen aus der Fortschreibung mittels BLUPs für die Zufallsparameter über die Prognosezeiträume im Mittel genauso hoch zu sein wie bei der herkömmlichen Fortschreibung mit Hilfe der Bonitätsfächer. Auch die Spanne zwischen dem dritten und ersten Quartil (Q3-Q1) und zwischen dem 90er- und 10er-Perzentil scheint in ähnlicher Höhe zu liegen. Die hohe Treffgenauigkeit der Fortschreibungen mit dem Bonitätsfächer überrascht, da doch bei der Anpassung der Trendfunktion nicht die gruppierte Datenstruktur berücksichtigt werden konnte. Bemerkenswert ist, dass die Perzentile der Residuen aus der BLUP-Kalibrierung bei geringen Prognosezeiträumen breiter um die 0-Referenzlinie streuen, als bei den Vorhersagen mittels der Bonitätsfächer. Man könnte meinen, dass die Fortschreibungsergebnisse aus den

Bonitätsfächern anfangs noch besser auf den tatsächlichen Wachstumskurven verlaufen. Dieser Sachverhalt bedarf jedoch ausführlicherer Erläuterung. In Abbildung 4.24 kann man erkennen, dass die prognostizierten Kurven nicht genau durch die Beobachtungen verlaufen, anhand derer sie kalibriert wurden. Dieses scheint auf dem ersten Blick paradox zu sein. An dieser Stelle muss daran erinnert werden, dass der gemischte Modellansatz eine stochastische Restfehlerkomponente beinhaltet. Diese beschreibt die mögliche Variation innerhalb eines Baumes, z.B. aufgrund von Messfehlern oder Oszillationen. Die Varianz des Restfehlers wurde als Parameter durch die näherungsweise Likelihood-Maximierung im Rahmen der Modellanpassung geschätzt. Der Restfehler ist nun dafür verantwortlich, dass die prognostizierten Kurvenverläufe nicht direkt durch die Beobachtungen zur Kalibrierung verlaufen. Die Abschätzung des Restfehlers ist eine Leistung, die nur das gemischte Modell aber nicht das Modell des Bonitätsfächers vollbringt. Der Bonitätsfächer begreift die gemessenen Beobachtungen als „wahre“ und ohne Fehler gemessene Werte und zwingt die Wachstumskurve dazu, durch die Startbeobachtung zu verlaufen. Der Bonitätsfächer hat jedoch nicht den Anspruch stochastischer Natur zu sein, er beschreibt die Wachstumsverläufe vielmehr als Gesetzmäßigkeit mechanistischen Ursprungs.

Zum anschaulicheren Vergleich der Verteilungsmaße der Residuen dient die Abbildung 4.33. Der mittlere Restfehler der Fortschreibungen mittels BLUPs liegt bei geringen Prognosezeiträumen nahe bei 0 und wächst dann für einen Prognosezeitraum von 10 Jahren bis auf 24 cm an. Danach sinkt er wieder gegen 0. Ähnlich verhalten sich die Residuen aus den Fortschreibungen mittels der Bonitätsfächer. Offensichtlich wichen die Bäume in der Form ihres Wachstums zwischen den Modell-Daten der IWE und den Kalibrierungsdaten des ZF-Projekts auseinander. Die Residuen aus den Fortschreibungen beider Methoden (BLUPs und Bonitätsfächer) zeigen den gleichen Trend. Die Prognosefehler aus der BLUP-Fortschreibung liegen bis zur 20. Beobachtung im Mittelfeld aller Modelle. Für weitere Prognosezeiträume scheint das gemischte Modell mit Zufallsparametern aber treffgenauer zu sein. Während die Bonitätsfächer die Baumhöhen hier im Mittel um rund einen halben Meter überschätzen, überschätzt das gemischte Modell nur um 10 bis 20 cm. Der Tabelle 8.6 des Anhangs ist zu entnehmen, dass bei diesen langen Prognosezeiträumen noch

zahlreiche Beobachtungen verfügbar waren (21 Jahre (n=277), 22(n=210), 23(n=157)). Zugleich streuen die Residuen der BLUPs, gemessen an den Reichweiten zwischen Q3 und Q1 und zwischen P90 und P10, für längere Prognosezeiträume stärker als bei den Bonitätsfächern. Wenn jeweils nur eine Beobachtung als bekannte Vorinformation vorliegt, können die BLUPs zwar das Ausmaß der Variation zwischen und innerhalb eines Stichprobenpunktes beschreiben und den Restfehler abschätzen, sie liefern in diesem Fall jedoch erst für Prognosezeiträume ab 20 Jahren treffgenauere Fortschreibungsergebnisse als die herkömmlichen Bonitätsfächer.

Es soll nun genauer untersucht werden, ob die BLUPs zu treffgenaueren Höhen-Vorhersagen führen können, wenn nicht nur eine, sondern zwei Beobachtungen pro Individuum zur Kalibrierung vorliegen. Der Bonitätsfächer kann im Gegensatz dazu lediglich eine Beobachtung als Startinformation verwerten. Aus dem rechten Teil der Abbildung 4.33 wird ersichtlich, dass die Höhen-Vorhersagen mit BLUPs ab einem Prognosezeitraum von 12 Jahren die Baumhöhe nur um 4 bis 18 cm überschätzen. Die Bonitätsfächer unterschätzen für diesen Zeitraum um 30 bis 55 cm. Der Abbildung 4.33 ist weiterhin zu entnehmen, dass die Spannen zwischen Q3 und Q1 und zwischen P90 und P10 ab dem sechsten bzw. ab dem 4 Prognosejahr merklich kleiner sind als bei den Vorhersagen mit Hilfe der Bonitätsfächer.

Im folgenden soll ein wesentlicher, methodischer Vorteil des auf Zufallsparametern basierenden Modellansatzes vorgestellt werden. Im Abschnitt 3.2.4 wurde beschrieben, dass die Fehler für die Vorhersagewerte der Zufallsparameter und der Baumhöhen approximiert werden können. Es können demnach Vertrauensintervalle für die Baumhöhenfortschreibungen konstruiert werden. Hierbei dient Gleichung 3.25 als Schätzer für die Varianz des Vorhersagefehlers der Zufallsparameter. Die mit Zufallsparametern bedachten Funktionsparameter der zu kalibrierenden Mittelwertkurven sind abhängig von den fixen Parametern $\iota^{(2)}$, $\iota^{(4)}$, $k_2^{(4)}$, $k_4^{(4)}$, $k_8^{(4)}$ und $k_9^{(4)}$. Eine Besonderheit stellt der Parameter $k_9^{(4)}$ dar; er hat einen nichtlinearen Einfluss auf den Funktionsparameter $\psi_{ij}^{(4)}$. Da dieses in den Linearisierungen zur Fehler-Approximation nur schwer zu berücksichtigen ist, soll der Schätzfehler dieses Parameters einfachheitshalber vernachlässigt werden. Die Matrix

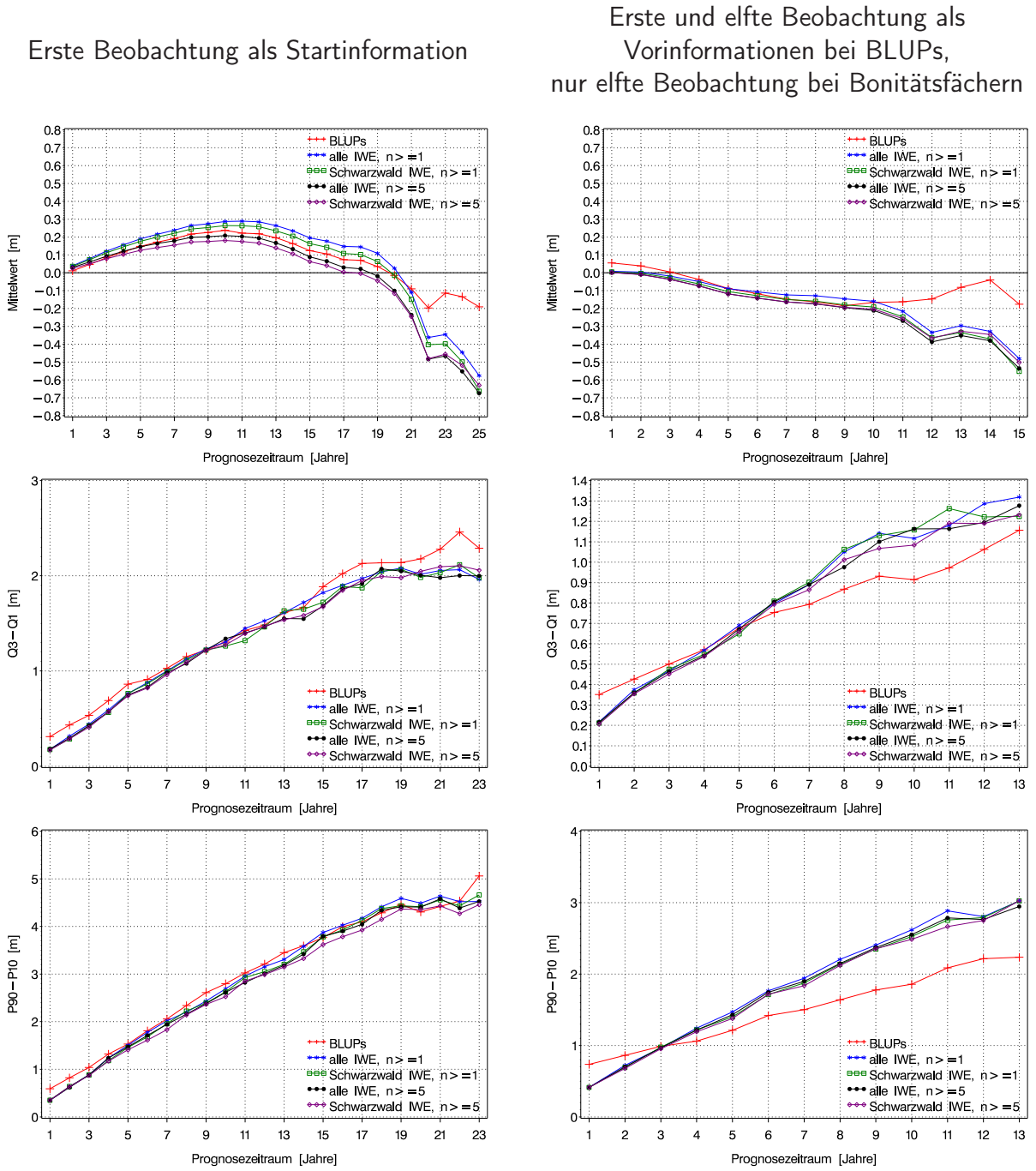


Abbildung 4.33: *Vergleiche der Verteilungsmaße für die Residuen der Vorhersagen. Durch rote Linien verbundene Plus-Symbole: Residuen der Vorhersagen mittels BLUPs für die Zufallsparameter. Übrige Linien: Residuen aus den Vorhersagen mit den angepassten Bonitätsfächern.*

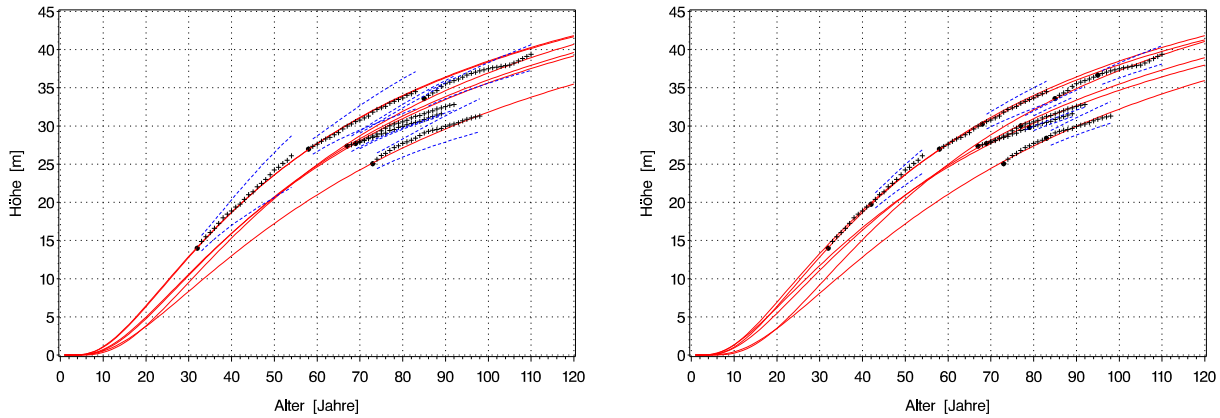


Abbildung 4.34: *Genauigkeit von Baumhöhenvorhersagen. Blau gestrichelte Linien: Vertrauensbereich auf dem 95%-Niveau. Links: Eine Beobachtung zur Kalibrierung. Rechts: Zwei Beobachtungen zur Kalbrierung.*

Ω aus Gleichungsfolge 3.26 enthält die geschätzten Kovarianzen für die fixen Parameter außer von $k_9^{(4)}$. Die geschätzten Kovarianzen aller fixen Parameter des endgültigen Modells sind in Tabelle 8.2 des Anhangs aufgeführt.

Als Varianzschätzer für die Fehler der Baumhöhenvorhersagewerte dient Gleichung 3.28. Um die Vorhersagewerte der Baumhöhen herum kann ein 95%-Vertrauensintervall geschätzt werden über:

$$\hat{y}_{ijk} \pm 1,96\sqrt{\hat{Var}\left(\hat{y}_{ijk} - y_{ijk}\right)} \tag{4.34}$$

Der t-Wert in Höhe von 1,96 rührt von einer annähernd unendlichen Anzahl von Freiheitsgraden her, die während der Schätzung der fixen Parameterwerte im Zuge der Modellanpassung zur Verfügung standen.

In der Abbildung 4.34 sind die Konfidenzintervalle für die Baumhöhen-Prognosewerte des Stichprobenpunktes aus den Abbildungen 4.21 - 4.25 exemplarisch dargestellt. Man kann daraus erkennen, dass sich der Vertrauensbereich der Vorhersagewerte für einen Vorhersagezeitraum gegebener Länge verkleinert, wenn zwei Beobachtungen zur Kalibrierung verwendet werden.

In Tabelle 4.5 sind die mittleren geschätzten Vorhersagefehler nach den Prognosezeiträumen auf dem 95%-Niveau dargestellt. Wenn nur eine Beobachtung zur Kalibrierung verwendet wird, dann können die Baumhöhen für einen Zeitraum von 25 Jahren mit einer

Genauigkeit (95%-Niveau) von rund ±2,8 m vorhergesagt werden und für einen Zeitraum von 15 Jahren mit einer Genauigkeit von rund ±2,3 m. Dieses entspricht einem relativen Fehler von rund 10%. Wenn dagegen zwei Beobachtungen zur Kalibrierung genutzt werden können, vermindert sich der geschätzte absolute Fehler bei einem Prognosezeitraum von 15 Jahren auf rund 1,5 m bzw. auf rund 5%.

	Eine Beob. zur Kalibrierung						Zwei Beob. zur Kalibrierung					
t	N	Mtw	StdAbw	m%	N_{vb}	$A\%_{vb}$	N	Mtw	StdAbw	m%	N_{vb}	$A\%_{vb}$
1	368	1,02	0,10	6,28 %	364	98,91 %	366	1,01	0,20	4,67 %	360	98,36 %
2	368	1,05	0,09	6,26 %	364	98,91 %	366	1,04	0,19	4,74 %	358	97,81 %
3	368	1,10	0,09	6,37 %	362	98,37 %	366	1,08	0,19	4,81 %	355	96,99 %
4	368	1,16	0,09	6,57 %	354	96,20 %	364	1,12	0,19	4,91 %	349	95,88 %
5	368	1,24	0,10	6,84 %	348	94,57 %	364	1,16	0,18	5,00 %	343	94,23 %
6	368	1,33	0,13	7,14 %	343	93,21 %	364	1,20	0,18	5,09 %	340	93,41 %
7	368	1,42	0,16	7,47 %	333	90,49 %	361	1,24	0,18	5,19 %	337	93,35 %
8	368	1,52	0,19	7,81 %	329	89,40 %	358	1,29	0,19	5,29 %	331	92,46 %
9	367	1,62	0,23	8,15 %	328	89,37 %	351	1,33	0,19	5,41 %	320	91,17 %
10	367	1,72	0,27	8,48 %	323	88,01 %	323	1,37	0,20	5,51 %	291	90,09 %
11	366	1,83	0,31	8,82 %	323	88,25 %	277	1,41	0,16	5,59 %	246	88,81 %
12	366	1,93	0,35	9,13 %	323	88,25 %	210	1,45	0,17	5,67 %	187	89,05 %
13	366	2,04	0,38	9,44 %	320	87,43 %	157	1,49	0,19	5,66 %	141	89,81 %
14	364	2,15	0,42	9,76 %	319	87,64 %	101	1,51	0,22	5,56 %	92	91,09 %
15	364	2,25	0,45	10,05 %	322	88,46 %	50	1,50	0,22	5,27 %	47	94,00 %
16	364	2,35	0,49	10,31 %	319	87,64 %	13	1,46	0,29	4,96 %	13	100 %
17	361	2,46	0,52	10,58 %	315	87,26 %	2	1,35	0,49	5,02 %	2	100 %
18	358	2,56	0,55	10,85 %	313	87,43 %	1	1,75		7,22 %	1	100 %
19	351	2,66	0,58	11,12 %	307	87,46 %						
20	323	2,74	0,62	11,26 %	284	87,93 %						
21	277	2,83	0,64	11,43 %	243	87,73 %						
22	210	2,90	0,66	11,42 %	183	87,14 %						
23	157	2,94	0,70	11,29 %	137	87,26 %						
24	101	2,89	0,75	10,76 %	88	87,13 %						
25	50	2,79	0,78	9,96 %	43	86,00 %						
26	13	2,62	0,92	8,97 %	12	92,31 %						
27	2	2,09	1,29	7,75 %	2	100 %						
28	1	3,08		12,21 %	1	100 %						
Gesamt	8072	1,91	0,73	8,86 %	7302	90,46 %	4394	1,23	0,24	5,15 %	4113	93,60 %

Tabelle 4.5: *Geschätzte Vorhersageungenauigkeit. t: Vorhersagezeitraum, N: Anzahl Beobachtungen, Mtw: mittlerer absoluter geschätzter Vorhersagefehler auf dem 95%-Niveau, StdAbw: Standardabweichung des absoluten Fehlers, m%: mittlerer prozentualer geschätzter Vorhersagefehler,* N_{vb}*: Anzahl von Beobachtungen im Vertrauensbereich der Vorhersage,* $A\%_{vb} = N_{vb}/N$.

Zur Beurteilung der Güte der Fehlerabschätzung können die geschätzen Vorhersagefehler mit den empirischen Residuen verglichen werden. Dazu wird geprüft, zu welchem Anteil die tatsächlichen Baumhöhenbeobachtungen im Vertrauensbereich ihrer Vorhersagen liegen. Man sollte erwarten, dass die Beobachtungen in 95% der Fälle im Vertrauensbereich liegen. Der Tabelle 4.5 ist zu entnehmen, dass die Vertrauensbereiche bei nur

einer Beobachtung zur Kalibrierung für einen Vorhersagezeitraum von bis zu vier Jahren noch zwischen 99% und 96% der Beobachtungen einschließen. Ab dem achten Jahr werden lediglich zwischen 87% und 89% der Beobachtungen eingeschlossen. Wenn zwei Beobachtungen zur Kalibrierung dienten, dann liegen bis zum achten Jahr noch zwischen 98% und 92% in den Vertrauensbereichen. Im ungünstigsten Fall werden die Fehler so stark unterschätzt, dass die Vertrauensbereiche nur rund 89% Beobachtungen abdecken (Vorhersagezeitraum von elf Jahren). Insgesamt liegen bei Verwendung einer Vorinformation rund 90,5% aller Realisationen in den Vertrauensbereichen und bei Verwendung von zwei Vorinformation rund 93,6%. Die Anteilswerte sind vor allem bei längeren Vorhersagezeiträumen gegenüber 95% etwas zu klein. Dieses kann mehrere Ursachen haben. Obwohl die Matrizen $\hat{F}_{ij}$ und $\hat{F}_{ijk}$ bzw. $\hat{Z}_{ij}$ und $\hat{Z}_{ijk}$ aus den Gleichungen 3.25 und 3.28 Schätzungen sind, werden sie bei der Approximation als fix betrachtet. Dadurch wird die Varianz eher zu niedrig eingeschätzt. Das kann schließlich dazu führen, dass weniger Realisationen in den Vertrauensbereichen liegen als erwartet. Außerdem wurden in der Modellkonstruktion die Kovarianzen zwischen den Zufallsparametern einer Ebene mit $Cov\left[z_i^{(2)}, z_i^{(4)}\right] = Cov\left[z_{ij}^{(2)}, z_{ij}^{(4)}\right] = 0$ unterdrückt, um numerische Schwierigkeiten bei der Optimierung im Zuge der Parameterschätzung zu vermeiden (vgl. Abschnitt 4.1.3.1). Weitere Ursachen für die eher optimistischen Fehlerschätzergebnisse können die Annahmen über die Unabhängigkeit und Normalverteilung der Residuen sein. In weiterführenden Studien sollte dieser Sachverhalt genauer beleuchtet werden und ein geeignetes Fehlermodell entwickelt werden. Zuletzt könnte der Linearisierungsansatz der Fehler-Approximation Variation unterdrückt haben.

Zusammenfassend kann festgestellt werden, dass die Varianz des Fehlers der Baumhöhenvorhersagen nach Gleichung 3.28 ein wenig zu optimistisch, aber dennoch in einem glaubwürdigen Rahmen approximiert werden kann.

In diesem Abschnitt sollte das nichtlineare, hierarchische und gemischte Modell evaluiert werden. Dabei wurde das gemischte Modell an den nicht zur Modellanpassung verwendeten Daten angewendet, die im Rahmen der Wuchsleistungs-Studie innerhalb des Projekts „Zukunftsorientierte Forstwirtschaft“ im Südschwarzwald erhoben wurden. Die

ersten Baumhöhenbeobachtungen wurden als bekannt angenommen. Über deren Abweichungen zu individuellen Mittelwertkurven mit Parametern aus lediglich fixen Effekten wurden die Zufallsparameter mittels BLUPs vorhergesagt. Mit den BLUPs der Zufallsparameter und den fix-Effekt-Parametern wurden individuelle Höhenkurven kalibriert, die für Fortschreibungen verwendet wurden. Die erzielten Schätzungen wurden mit den beobachteten Werten verglichen. Als Referenzmodelle wurden Bonitätsfächer der Trendfunktion an die IWE-Daten angepasst. Zusammenfassend ist festzustellen, dass das gemischte Modell mit BLUPs für die Zufallsparameter anhand nur einer Vorbeobachtung zur Kalibrierung den herkömmlichen Bonitätsfächern nur bei langen Prognosezeiträumen überlegen ist. Wenn zwei Vorbeobachtungen zur Kalibrierung herangezogen werden, sind die mittleren Fehler schon ab einem Fortschreibungszeitraum von 11 Jahren geringer als bei den Bonitätsfächern. Hinzu kommt, dass sich die Streubreite der Fehler im Vergleich zu den Bonitätsfächerschätzungen merklich reduziert.

Im Gegensatz zu den Fortschreibungsergebnissen aus den Bonitätsfachern konnten für die Vorhersagen mit dem gemischten Modell Vertrauensbereiche geschätzt werden. Die Vertrauensbereiche schlossen nicht wie erwartet 95%, sondern nur 90,5% (bei einer Beobachtung zur Kalibrierung) bzw. 93,6% (bei zwei Beobachtungen zur Kalibrierung) der tatsächlichen Beobachtungen ein. Obwohl die Fehlerabschätzung tatsächlich zu eher optimistischen Ergebnissen führte, kann der Schätzansatz vor dem Hintergrund des approximativen Charakters als geeignet angesehen werden.

Kapitel 5

Diskussion

5.1 Methodischer Erkenntnisgewinn

5.1.1 Hierarchische, gemischte Modelle vs. Traditionelle Site-Index Modelle

Seit rund 200 Jahren hat die forstliche Ertragskunde die Idee verfolgt, anhand der mittleren Entwicklung von Bestandes-Kennwerten über dem Alter, modellhafte Arbeitsgrundlagen zur Planung, Kontrolle und Beurteilung forstwirtschaftlicher Maßnahmen zu schaffen (Pretzsch, 2001). Die Ergebnisse dieser Bestandesmodelle mündeten in Ertragstafeln. Den Bestandesmodellen der Ertragstafeln liegen stets bestimmte Alters-Höhen-Beziehungen und Gesamtwuchsleistungskurven zugrunde. Die Bonität wurde als Maß zum Vergleich und zur Bewertung der Holzertragsfähigkeit von Waldbeständen begriffen (Wenk et al., 1990). Im Jahre 1880 beschloss der Verein deutscher forstlicher Versuchsanstalten einen einheitlichen Werterahmen für die Bonitäten der wichtigsten Hauptbaumarten. Die zu erwartende Gesamtwuchsleistung je ha im Alter von 100 Jahren wurde in fünf kategorische Klassen unterschiedlicher Breite eingeteilt. Diese Güte-Klassen wurden in ordinaler Weise mit römischen Ziffern aufsteigend von I bis V für fallende Bonitäten bezeichnet. Die Grenzen der Klassen waren jedoch willkürlich gezogen. Weil der Rahmen nicht das

gesamte mögliche Spektrum der Ertragsfähigkeit abdeckte, war man stets zu unzulässigen Extrapolationen gezwungen.

Aus diesen Gründen ging man dazu über, Ertragstafeln zu formulieren, die durch das Site-index-System charakterisiert sind. Grundlage für die Bonitierung ist die sogenannte Höhenbonitierung. Die Bestandeshöhen haben nach Wenk et al. (1990) mehrere Vorteile. Sie sind leicht messbar und korrelieren streng mit dem Holzertrag. Wegen der „rechnerischen Verschiebung“ von Bestandesmittelhöhen nach forstlichen Eingriffen, insbesondere bei Niederdurchforstungen ging man dazu über, Oberhöhen als Bonitätsmaßstab zu verwenden. Das Wachstum der Oberhöhenstämme ist stark vom Standort abhängig, auf dem der Waldbestand stockt. Zusätzlich kann aber auch die Genetik für Variation sorgen. Es ist jedoch schwierig, die zwei Faktoren Standort und Genetik zu trennen und ihren Einfluss separat zu beschreiben (Wenk et al., 1990); meist ignorieren Ertragskundler den Einfluss der genetischen Komponente auf das Baumwachstum. Monserud und Rehfeld (1990) unternehmen einen Versuch, den Site-index der Douglasie in Nord-Amerika in Abhängigkeit von genetischen Unterschieden zu beschreiben. Sie bedienen sich dabei des „genetischen Index“ nach Rehfeldt (1989). Nach Monserud und Rehfeld (1990) ist der Site-index als Bonität die durchschnittliche Höhe dominanter Bäume einer bestimmten Baumart in einem Referenzalter. Die zugrunde liegenden dominanten Bäume („site-trees“) weisen das beste Wachstum aller Bäume eines Bestandes auf. Sie haben gesunde Kronen, konsistentes Durchmesserwachstum und keine Anzeichen von Schädigungen oder Wuchsdepressionen in der Vergangenheit.

Site-index-Modelle können sich nach Rennolls (1995) darin unterscheiden, ob jede Wachstumstrajektorie separat modelliert wird oder ob der gesamte Datenbestand als Realisation eines einzigen Modells angesehen wird. Wenn jede einzelne Trajektorie modelliert wird, kann das Modell zusammengeführt werden, indem die angepassten Parameterwerte durch Regressionen auf Standortsvariablen zurückgeführt werden. Für diese Fälle besteht das Site-index-Modell grundsätzlich aus zwei Kompartimenten. Zum einen liegt ein Höhenwachstumsmodell für die dominanten „site-trees“ zugrunde und zum anderen Gleichungen für den Site-index. Wenn dagegen die Parameter des Site-index-Modells

gleichzeitig für den gesamten Datenbestand geschätzt werden, dann wird das Wachstum durch eine Schar von sich nicht schneidenden Kurven beschrieben. Sloboda (1971) formulierte Bedingungen für die Wachstumskurven eines Bonitätssystems. Demnach sollen die einzelnen Kurven einen gemeinsamen Knotenpunkt im Ursprung des kartesischen Koordinatensystem haben, und danach sollen die Kurven fächerförmig auseinander laufen und sich nicht überschneiden. Durch einen beliebigen Punkt soll nur eine Kurve verlaufen. Die Kurven sollen genau einen Wendepunkt und eine Asymptote aufweisen. Geleistet wird dieses von Wachstumsfunktionen, die einen freien Parameter haben, dieser ist die Bonität (Site-index) (Garcia, 1983; Wenk et al., 1990).

Cieszewski et al. (2006) bezeichnen die Gleichungs-Form zur separaten Parameterschätzung für die Trajektorien als *einfach* und die Form zur gemeinsamen Parameterschätzung als *dynamisch*. Zur Darstellung dieses Sachverhalts wird das Beispiel von Nord-Larsen (2006) benutzt. Wir nehmen an, das Wachstum wird beschrieben durch eine Richards-Funktion mit drei Parametern:

$$H(t) = \beta_1 \left(1 - \exp\left(-\beta_2 t\right)\right)^{\beta_3} \tag{5.1}$$

Wenn der asymptotische Parameter β_1 als standorts-spezifischer Parameter gewählt wird, erhalten wir zur Beschreibung des Wachstums für jede Trajektorie die einfache Form mit:

$$H(t, X) = X \left(1 - \exp\left(-\beta_2 t\right)\right)^{\beta_3} \tag{5.2}$$

Dabei sei X eine Kombination beliebiger standortsabhängiger Parameter. Auflösen nach X ergibt:

$$X = H(t, X) \left(1 - \exp\left(-\beta_2 t\right)\right)^{-\beta_3} \tag{5.3}$$

Da Standortsvariablen selbst nur schwer messbar oder zu interpretieren sind, bedient man sich der Bonität (Site-index) als Referenzgröße für die Baumhöhe, die in einem bestimmten Alter zu erwarten ist. Man erhält dann:

$$X = H_0 \left(1 - \exp\left(-\beta_2 t_0\right)\right)^{-\beta_3} \tag{5.4}$$

Durch Substitution von X in Gleichung 5.2 aus Gleichung 5.4 erhält man die dynamische Form:

$$H(t, t_0, H_0) = H_0 \left(\frac{1 - \exp(-\beta_2 t)}{1 - \exp(-\beta_2 t_0)} \right)^{\beta_3} \tag{5.5}$$

Jede Höhe H zum Alter t kann über die Kenntnis der Höhe H_0 zu einem Referenzalter t_0 geschätzt werden. H_0 wird dabei als Bonität (Site-index) begriffen. Die Variablen für die Güte der Standorte sind in dieser dynamischen Form durch die Bonität (den Site-index) substituiert. Mit veränderbaren Werten der Bonität wird für den gesamten Datenbestand ein gemeinsamer Fächer von Wachstums-Kurven aufgespannt.

Die resultierenden Kurvenfächer können wiederum von unterschiedlicher Gestalt sein. Cieszewski et al. (2006) unterscheiden anamorphe, polymorphe und fortgeschritten polymorphe Modelle. Wenn wie in Gleichung 5.5 aus dem obigen Beispiel der asymptotische Parameter durch die Bonität ersetzt wird, dann spricht man von einer anamorphen Gestalt. Wenn der asymptotische Parameter dagegen als fixe Größe angenommen wird und ein formbestimmender Parameter durch die Bonität substituiert wird, erhält der Wachstums-Fächer eine polymorphe Gestalt. Cieszewski et al. (2006) setzten sich zum Ziel, das Wachstum der Kiefer funktional zu beschreiben, indem für die Ertragstafelwerte von Schwappach ein geeignetes Site-index-Modell formuliert werden sollte. Die anamorphen und polymorphen Modelle reichten in ihrer Flexibilität nicht aus. Deshalb formulierten sie neben einem Formparameter noch den asymptotischen Parameter abhängig von der Bonität und nannten diese Modell-Gestalt fortgeschritten polymorph. In der Studie von Cieszewski et al. (2006) wurde der Fehler nicht modelliert. Man könnte kritisch anmerken, dass die Ertragstafeln Schwappachs aus Mittelwerten für willkürlich festgelegte relative Bonitätsklassen bestehen. Die Höhenerwartungswert-Bonitäten aus dem fortgeschritten polymorphen Modell dürfen deshalb streng genommen nicht interpoliert werden.

Die Konstruktion der Site-index-Modelle kann sich neben den bereits skizzierten Kriterien auch in der Behandlung des Referenzalters des Bonitätsparameters unterscheiden. Es gibt Ansätze, die entweder als Referenzalter-spezifisch oder als Referenzalter-invariant bezeichnet werden. Beim Referenzalter-spezifischen Ansatz werden die Höhen-Werte der

einzelnen Individuen oder die der Mittelhöhe an einem Stichprobenpunkt im Referenzalter als feste Werte für die Bonität angenommen. Für diese Werte selbst wird kein Fehler angenommen. Die Wachstumskurven werden durch die beobachtete Bonität hindurch gezwungen. Deswegen haben die empirischen Residuen aus der Modellanpassung biologisch unplausible Verläufe (Nord-Larsen, 2006). Für den Referenzalter-invarianten Ansatz dagegen wird die Bonität als zu schätzender Parameter angenommen. Der Referenzalter-spezifische Ansatz hat außer den unplausiblen Residuenverläufen noch weitere bedeutende Nachteile. Zum einen können grundsätzlich nur Zeitreihen-Beobachtungen von Individuen zur Modellanpassung verwendet werden, bei denen zum Referenzalter eine Höhenbeobachtung vorliegt; dieses kann zum Ausschluss umfangreicher und wertvoller Datenmengen führen. Teilweise wird der Auschluss von Daten umgangen, indem für fehlende Beobachtungen die Höhe im Referenzalter über Inter- oder Extrapolation hergeleitet wird; dieses birgt das Risiko verzerrter Schätzungen. Zum anderen führt dieser Ansatz, wie Strub und Cieszewski (2006) sehr anschaulich darstellen, zu unterschiedlichen Parameterschätzwerten und Wachstumsverläufen, je nach dem welches Referenzalter gerade ausgewählt wird. In jener Studie wurden für beide Ansätze über die Minimierung der Residuenquadratsumme sowohl Anpassungen der nichtlinearen Gleichungsform als auch der linearisierten Form vorgenommen. Nur beim Referenzalter-invarianten Ansatz wurden durch den nichtlinearen Fit und die Anpassung der linearisierten Gleichung jeweils die gleichen Parameter-Werte geschätzt. Zudem ist in dem Beispiel von Strub und Cieszewski (2006) die Summe der quadrierten Residuen für beide Anpassungsmethoden (linearisiert und nichtlinear) aus dem Referenzalter-spezifischen Ansatz höher, als wenn die Bonität als variabler Parameter begriffen wird.

Bei der Konstruktion des HNLMM in meiner Studie wurde die Idee des Referenzalter-invarianten Ansatzes aufgegriffen. Der Höhenerwartungswert-Parameter für das Alter von 50 Jahren wurde für jedes Individuum als zu schätzender Parameter begriffen.

Da die Wachstumsfunktionen der Site-index-Modelle normalerweise sehr unflexibel sind, stellt Garcia (1983) einen stochastischen Ansatz vor. Anhand von Wiederholungsmessungen aus Stammanalysen von Stichprobenpunkten wurde eine Bertalanffy-Richards-

Funktion angepasst. Basierend auf der Maximum-Likelihood-Technik entwickelte er eine Software-Anwendung, mit der einige Parameter als „globale Parameter“ - für die gesamte Population als konstant angenommen - geschätzt wurden und andere Parameter als „lokale Parameter“ spezifisch für die einzelnen Stichprobenpunkte. Ein freier Parameter bestimmte die Bonität. Zusätzlich wurden in dem Wachstumsmodell jener Studie noch zwei stochastische Komponenten angenommen. Eine Komponente sollte schwankenden Umwelteinflüssen Rechnung tragen, die das Wachstum beeinflussen können. Garcia (1983) wählte dafür eine Brownsche Bewegung, auch Wiener-Prozess genannt. Der Wiener-Prozess besteht aus stationären Zuwächsen, die normalverteilt und unabhängig sind und den gleichen Erwartungswert und die gleiche Varianz zu allen Zeitpunkten aufweisen. Die zweite Komponente stellt den unabhängigen Messfehler als normalverteilte Größe mit Mittelwert gleich 0 und bestimmter Varianz dar. Sein Algorithmus zur Likelihood-Maximierung schien jedoch in einer Evaluierung Schwierigkeiten zu haben, den Restfehler und die Schwankung durch den Wiener-Prozess zu separieren. Die Arbeit von Garcia (1983) kann jedoch in mancher Hinsicht als Meilenstein angesehen werden. Das Baumhöhenwachstum wird nicht mehr durch starre Bonitätsfächer mit dem Charakter von „Wuchsgesetzen“ abgebildet. Um die Mittelwertkurve der Population herum wird die Variation an den einzelnen Stichprobenpunkten durch die lokalen Parameter beschrieben. Außerdem werden durch den Wiener-Prozess Oszillationen um die Kurven ermöglicht. Der modellierte Restfehler steigert die „Ehrlichkeit“ des Modells in stochastischer Hinsicht. In diesen Punkten ist das Modell von Garcia (1983) dem vorgestellten HNLMM sehr ähnlich. Auch durch den hierarchischen und gemischten Modellansatz dieser Arbeit erhielten wir ein Mittelwertmodell für die Population. Die Variation, die nicht durch Kovariablen zu erklären war, wurde mit Hilfe normalverteilter Zufallskomponenten mit Mittelwerten gleich 0 beschrieben.

Eine Konstruktion eines traditionellen Site-index-Modells findet sich bei Monserud (1984). Er stellte heraus, dass nicht der einzelne Baum, sondern vielmehr der Standort der Gegenstand des Interesses ist. Der Baum sei lediglich ein „Phytometer“. Mit drei verschiedenen Ansätzen wurden von ihm Referenzalter-spezifische Site-index-Modelle angepasst.

In der Studie sollte geprüft werden, ob sich das Wachstum der Douglasie in fünf Wuchsräumen Nord-Amerikas unterscheidet. Ein Modell mit einem Ansatz aus zwei simultanen Richards-Funktionen, eine für die Bonität und eine für das Höhenwachstum, scheiterte aufgrund der schwachen numerischen Eigenschaften der Richards-Funktion. Da in meiner Studie ähnlich schlechte Erfahrungen mit der Richards-Funktion gemacht wurden, kam schließlich die Trendfunktion nach Sloboda (1971) zur Anwendung. Mit der zweiten Methode stellte Monserud (1984) in einem Gleichungssystem zunächst Höhe und Bonität als lineare Funktion von einander dar. Dann wurden die mittlere Höhe und die mittlere Bonität sowie die jeweiligen Steigungskoeffizienten in Abhängigkeit vom Alter funktional angepasst. Diese Funktionen wurden danach wieder in das simultane Gleichungs-System eingesetzt; dadurch erfuhren die Mittelwerte für Bonität und Höhe in Abhängigkeit vom Alter quasi eine Korrektur. Während die Site-index-Funktion aus dem angepassten Gleichungssystem noch einen plausiblen Verlauf zeigte, waren die Schätzungen aus der Höhenwachstumsfunktion besonders in niedrigen und hohen Altern biologisch unplausibel. Dieser Schwachpunkt resultiert aus dem polynomischen Ausgleich. Das HNLMM, das in dieser Arbeit entwickelt wurde, zeigt dagegen wegen der zugrunde liegenden nichtlinearen Wachstumsfunktion im gesamten Bereich möglicher Abszissenwerte biologisch plausible Höhen-Wachstumsverläufe.

Abhängig von Dummy-kategorial kodierten Gruppen von Wuchsräumen wurden von Monserud (1984) Steigungs-Koeffizienten für die Prädiktoren des simultanen Gleichungssystems geschätzt. Es konnten zwar unterschiedliche Kurvenverläufe für die Site-index-Funktion grafisch dargestellt werden, die Parameterwerte konnten wegen der polynomischen Funktion jedoch nicht interpretiert werden. Im hier entwickelten HNLMM konnte der Einfluss der fixen Effekte auf den Höhenerwartungswert-Parameter ausführlich gedeutet werden (siehe Abschnitt 4.1). Die fixen Effekte für die Merkmalskombinationen aus Datenherkunft und Region, sowie das Keimjahr am Stichprobenpunkt verursachen ein unterschiedliches Niveau für das Baumwachstum. Die Niveau-Unterschiede können direkt in Form der resultierenden Höhenerwartungswerte im Referenzalter von 50 Jahren beziffert werden. Wegen der Unterschiede zwischen den erfassten Kollektiven konnten

die einzelnen Wachstumsniveaus, verursacht durch fixe Parameter-Effekte, nicht zwischen den beiden Inventuren verglichen werden. Innerhalb der beiden Inventuren sind jedoch Vergleiche zwischen den Regionen statthaft. Außerdem konnten die Koeffizienten-Werte für das Keimjahr als fixer Effekt gedeutet werden, der den zeitlichen Trend der Wachstumsveränderung beschreibt.

Rennolls (1995) legte grundsätzliche Aspekte der Höhenwachstumsmodellierung dar und diskutierte zwei verschiedene Modell-Ansätze. Auf einen Gesichtspunkt wurde weiter oben schon ausführlich eingegangen. Zunächst sollte man entscheiden, ob jede einzelne Trajektorie separat angepasst werden soll oder für die gesamte Population eine Schar sich nicht schneidender Kurven. Eine weitere Frage sei, ob und wie die Variation um die Trajektorien beschrieben werden soll. Außerdem sollte man entscheiden, ob neben dem Wachstum auch die Mortalität berücksichtigt werden soll. Schließlich sei zu klären, ob der Restfehler in das Modell eingehen soll.

In den traditionellen Site-index-Modellen wird häufig ein simultanes System aus zwei Gleichungen aufgestellt, eine zur Schätzung des Höhenwachstums der dominanten Bäume und eine für die Schätzung des Site-index. Würden die Parameterschätzungen der beiden abhängigen Gleichungen anhand des selben Datensatzes mittels OLS-Methode vorgenommen werden, dann wären die Schätzergebnisse verzerrt, da die Fehler der beobachteten Baumhöhen und der Site-indices miteinander korrelieren (Reversibilitäts-Problem). Beaumont et al. (1999) untersuchten vor diesem Hintergrund verschiedene Methoden der Parameterschätzung. Sollten zusätzlich zum Reversibilitäts-Problem noch Autokorrelation und Heteroskedastizität der Residuen hinzutreten, sind nach seiner Meinung Hypothesen-Tests für Parameterwerte aus zwei- und dreistufigen Kleinstquadratschätzern (2SLS, 3SLS) nicht aussagekräftig. Beaumont et al. (1999) stellten deshalb die Methode GMM (generalized method of moments) von Hansen (1982) zur Parameterschätzung vor. Anhand von Stammanalysen von *Pinus mariana (Mill.) Britt.* in Quebec wurde ein Site-index-Modell auf Basis einer Weibull-Funktion angepasst. Zwei verschiedene Regionen und drei Wasserversorgungs-Stufen wurden als Dummy kodiert. Die Koeffizienten wurden mit einer modifizierten Form von GMM geschätzt. Außerdem fanden Autokorrelation und Hetero-

skedastizität Berücksichtigung. Ein unterschiedlicher Einfluss der beiden Regionen auf das Baumwachstum konnte durch signifikant unterschiedliche Koeffizienten auf einen Funktionsparameter nachgewiesen werden. Für einen anderen Funktionsparameter unterschieden sich die Koeffizienten für die Wasserhaushalts-Stufen signifikant. In einer Monte-Carlo-Simulation sollte die angewendete Methode mit anderen Parameterschätzmethoden verglichen werden. Dazu simulierten Beaumont et al. (1999) mit der Weibull-Funktion 20 Wertepaare von Alter und Höhe für 100 Bäume an 10 Stichproben. Die Parameter der Wachstumsfunktion wurden heuristisch festgelegt, wobei regionale Effekte nicht modelliert wurden. Die generierten Daten wurden willkürlich in zwei Regionen geteilt. Dann wurde an diese Daten das Site-index-Modell mit verschiedenen Parameterschätzmethoden angepasst. Hierbei wurden nun auf zwei Parametern wieder Koeffizienten für die Dummy-kodierten Regionen angenommen. Der gesamte Prozess wurde 1000 mal durchgeführt. Da bei der Generierung der Daten kein regionaler Effekt angenommen wurde, musste erwartet werden, dass für die regionalen Koeffizienten der Wert 0 bei der Anpassung geschätzt wird (0-Hypothese). Die 0-Hypothese wurde bei jedem Simulationslauf mit dem Wald-Test auf dem 0,05-Niveau geprüft. Beaumont et al. (1999) erwarteten schließlich, dass nach den 1000 Simulationen nur in 5% der Fälle die 0-Hypothese verworfen und damit als Fehler 1. Ordnung fälschlicherweise ein regionaler Einfluss auf die Parameterwerte angenommen worden wäre. Tatsächlich wurde die Erwartung nur bei seiner modifizierten GMM-Parameterschätzmethode erfüllt. Bei allen anderen Methoden, darunter auch OLS und 3SLS wurde die 0-Hypothese bei rund 96% bzw 52% der Simulationsdurchläufe fälschlicherweise verworfen. Bedauerlicherweise brachte GMM jedoch eine Schwierigkeit mit sich; nur in 85% der Fälle trat Konvergenz der Optimierungsalgorithmen zur Parameterschätzung ein. Beaumont et al. (1999) konnten somit zwar eine Methode präsentieren, die unverzerrte Parameterschätzungen lieferte, jedoch in 15% der Fälle nicht konvergierte und dort gar keine Parameter schätzen konnte. Das Site-index-Modell von Beaumont et al. (1999) beschreibt das Wachstum für Oberhöhenbäume durch eine Kurvenschar. Mit Hilfe der Koeffizienten für die zwei Regionen und die drei Wasserhaushaltsstufen wurde dort der Versuch unternommen, das Wachstum variabler abzubilden. Außerdem wurde

ein stochastischer Restfehler angenommen. Trotz dieser Bemühungen kann ein HNLMM dagegen viel flexibler reagieren. Neben regionalen fixen Effekten und dem Keimjahr bewirken vor allem die Zufallsparameter auf zwei verschiedenen Funktionsparametern die hohe Anpassungsfähigkeit. Diese wurde vom HNLMM auch im vorhinein gefordert, da das Modell nicht nur das Wachstum von Oberhöhenbäumen beschreiben soll. Zur Vereinfachung nahmen Beaumont et al. (1999) an, dass die Bonität an allen Stichprobenpunkten über die Zeit konstant ist. Im Verzicht auf diese Annahme sieht er weitere Forschungsaufgaben. In dem HNLMM der vorliegenden Studie wird diese Annahme schon nicht mehr getroffen. Das durchschnittliche Keimjahr dient für den aktuellen Datenbestandteil aus den Regionen Südschwarzwald / Baar-Wutach und Odenwald und für die historischen Daten der Moränen-Standorte als Substitut für den Einfluss sich ändernder Umweltbedingungen auf das Baumwachstum. An dieser Stelle sei angemerkt, dass zur Datengewinnung für die Studie von Beaumont et al. (1999) Sektionsschnitte in 1,3 m Höhe und jeden Meter darüber durchgeführt wurden. Die Höhenwerte wurden für fünfjährige Intervalle von Alterswerten über lineare Interpolation zwischen den Alters- und Höhenmessungen an den Sektionen ermittelt. Neben Messfehlern an den Beobachtungsstellen können deshalb auch „falsche“ Interpolationen eine Fehlerquelle darstellen.

Da der österreichische Waldwachstumssimulator PROGNAUS (Sterba et al., 1995) bei längeren Fortschreibungszeiträumen den Höhenzuwachs systematisch überschätzte, nahm Hasenauer (1999) dieses zum Anlass, ein neues Höhenzuwachsmodell für Österreich aufzustellen. Dazu wurde in logarithmisch transformierter Form der Höhenzuwachs des einzelnen Baumes mittels linear kombinierter Variablen des Baumes und des Standortes beschrieben. Die beschreibenden Baumvariablen waren Höhe, Bhd, jährlicher Durchmesserzuwachs und die Kronenlänge; die Standortsgrößen waren die Seehöhe und der Hangrichtungsazimut. Hasenauer (1999) versuchte, mit Hilfe von lokalen und individuellen Variablen auf regressionsanalytischem Weg die Wirklichkeit zu beschreiben. Bei der Kalibrierung des HNLMM durch Vorhersagen der Zufallsparameter können die verfügbaren Informationen jedoch noch besser ausgenutzt werden. Wie in Abschnitt 4.2 dargestellt ist, wird der Informationsgehalt einer Beobachtung optimal ausgenutzt. Mit der Abweichung

der Beobachtung zu einem Wert, der aufgrund der fixen Effekte durch die Kovariablen zu erwarten gewesen wäre, können die Zufallsparameter, die für die lokale und individuelle Variation verantwortlich sind, verzerrungsfrei geschätzt werden. Mit den Zufallsparametern ergeben sich individuelle Höhenkurven, die dann für die weitere Fortschreibung verwendet werden. Hasenauer (1999) schätzte in einer Evaluierung an unabhängigen Daten die 5-jährigen Höhenzuwächse für drei Perioden von 1983 bis 1997. Der mittlere geschätzte Höhenzuwachs war für die erste Periode um 0,03 m größer als der tatsächliche mittlere Zuwachs, für die zweite bzw. dritte jedoch um 0,28 m bzw. 0,1 m kleiner. Mit dem Modell wurden jedoch weder die gruppierte Datenstruktur von Bäumen an Stichproben noch ein stochastischer Restfehler berücksichtigt. Die entlogarithmierte Linearkombination stellt darüber hinaus kein biologisch begründbares Wachstumsmodell dar.

Im Gegensatz zu den klassischen Site-index-Modellen war es Ziel dieser Studie, das Wachstum für einzelne Bäume zu beschreiben. Mit der hierarchischen Modellstruktur konnte die unbalancierte Datenstruktur (ungleiche Anzahl von Bäumen an den Stichproben) berücksichtigt werden. Außerdem konnten die Varianzen der Zufallsparameter für die Ebenen der Stichproben und der Bäume geschätzt werden und so geklärt werden, wie viel Variation jeweils auf den Ebenen verursacht wird. Obwohl dem hier vorgestellten Modell die Trendfunktion nach Sloboda (1971) zugrunde liegt und ein Parameter als Höhenerwartungswert in einem Referenzalter formuliert wurde, ist das gemischte Modell kein Site-index-Modell. Es wird im Gegensatz zu den traditionellen Site-index-Modellen nicht nur das Wachstum von Oberhöhenbäumen beschrieben. Das Wachstums-Modell muss viel flexibler sein und sogar überschneidende Kurvenverläufe von einzelnen Bäumen erlauben. Deshalb ist neben dem Bonitätsparameter zusätzlich noch ein weiterer Parameter frei und variabel gestaltet worden. Beide freien Parameter haben Zufallskomponenten auf Stichproben- und Baum-Ebene. Das hier entwickelte Modell ist in der Lage, auch in ungleichaltrigen Beständen das Höhenwachstum einzelner Bäume zu beschreiben; dieses ist exemplarisch in den Abbildungen 4.8, 4.24, 4.25 und 4.27 (rechter Teil) dargestellt. Neben den Zufallsparametern bewirken auch die fixen Effekte Variation; dieses sind die Region/Datenherkunft und das durchschnittliche Keimjahr des Bestandes auf

Stichproben-Ebene und die Abweichung von diesem Keimjahreswert für den einzelnen Baum. Das Modell ließe sich dahingehend weiterentwickeln, dass künftig auch explizit das Höhenwachstum in Mischbeständen beschrieben werden kann. Maßzahlen für die interspezifische Konkurrenzwirkung könnten dabei als fixe Parameter-Effekte dienen. Die Daten, die dieser Studie zugrunde lagen, enthielten leider keine Informationen zu Nachbarschaftsbeziehungen oder Stammzahldichten. Das entwickelte Modell kann jedoch schon zum momentanen Zeitpunkt, sofern Vorbeobachtungen zur Kalibrierung vorliegen, das spezifische Höhenwachstumsverhalten in Mischbeständen im Rahmen der Zufallsparameter-Variation beschreiben. Die Variation, die durch die Mischbestandssituation hervorgerufen wird, ist dann in den vorhergesagten Werten der Zufallsparameter enthalten. Der Anteil der Parameter-Variation, der auf die interspezifische Konkurrenz entfällt, kann jedoch nicht geschätzt werden. Der normalverteilte Restfehler des gemischten Modells soll neben dem Messfehler auch die Oszillation der Beobachtungen um die individuellen Kurvenverläufe abbilden. Im Gegensatz zu den traditionellen Site-index-Modellen erlaubt es das gemischte Modell, Vertrauensbereiche für die Baumhöhenvorhersagen abzuschätzen.

5.1.2 Würdigung des methodischen Erkenntnisgewinns

In diesem Abschnitt sollen die Vorteile eines nichtlinearen, hierarchischen und gemischten Modells zusammenfassend gewürdigt werden. Dabei wird zum Teil nochmals auf einige Gesichtspunkte im Vergleich zu den traditionellen Site-index-Modellen eingegangen.

Lappi und Bailey (1988) diskutierten in ihrer bedeutenden Pionier-Arbeit sehr klar und umfassend die Vorzüge ihres auf Zufallsparametern basierenden Modells. Sie führen aus, dass gemischte Modelle den Umständen Rechnung tragen, dass in einem bestimmten Bestand die Höhen von Bäumen miteinander korrelieren und, dass die Baumhöhen und die mittlere Höhe des Bestandes im zeitlichen Verlauf miteinander korrelieren. Zufallsparameter auf verschiedenen Hierarchieebenen klären das Ausmaß der Variation zwischen und innerhalb von Gruppierungsmerkmalen auf. Außerdem führt die Methode auch bei gruppierter Datenstruktur mit ungleichmäßig vielen Beobachtungen pro Gruppe zu unverzerrten Parameter-Schätzungen. Im Gegensatz dazu wird in traditionellen Site-index-Modellen

die mögliche Verteilung der Höhenkurven gar nicht geschätzt. Site-index-Modelle können nur das Wachstum von Oberhöhenbäumen mit relativ unflexiblen Kurvengleichungen beschreiben.

Bei Site-index-Modellen können die vorliegenden Informationen von Baumhöhenbeobachtungen nicht vollständig zur Fortschreibung genutzt werden. Bei diesen Modellen kann nur eine einzige, meist die letzte Beobachtung genutzt werden. Das HNLMM kann beliebig viele Vorbeobachtungen zur Kalibrierung der Wachstumskurven nutzen. Wenn pro Individuum mehrere aktuelle Messungen vorliegen, können sich die Kurven an das aktuelle Zuwachsgeschehen kalibrieren und dieses fortschreiben. Mit dem HNLMM können somit auch kurzfristige Wachstumstrends beschrieben werden, ohne das Modell von Grund auf neu zu konstruieren. Die Bonitätsfächer aus Site-index-Modellen haben dagegen einen vergleichsweise statischen Charakter, der auch bei den herkömmlichen Ertragstafeln stets kritisiert wird. Liegt nur eine Vorbeobachtung pro Individuum vor, dann können mit Hilfe der Kalibrierung des HNLMM die Parameter von nichtlinearen Wachstumsfunktionen geschätzt werden. Es können also grundsätzlich die Werte für eine höhere Anzahl von Parametern geschätzt werden, als Beobachtungen vorliegen.

Die hohe Treffgenauigkeit der Fortschreibungen mit dem herkömmlichen Bonitätsfächer überraschte. Obwohl beim Bonitätsfächer mit dem Höhenerwartungswert zum Referenzalter nur ein freier Parameter existiert und nur eine Beobachtung als Vorinformation dient, waren die Prognosefehler für Fortschreibungsperioden bis zur Länge von zehn Jahren im Mittel genauso groß wie bei dem gemischten Modell, wenn die Zufallsparameter anhand von zwei Beobachtungen vorhergesagt wurden. Erst ab dem elften Jahr waren die Residuen beim gemischten Modell im Mittel betragsmäßig kleiner. Die Streubreiten der Fehler, gemessen an ihren Perzentilabständen, waren beim gemischten Modell aber schon bei kürzeren Prognosezeiträumen kleiner.

Der Bonitätsfächer liefert im Gegensatz zum gemischten Modell keine Informationen über die Verteilung der zu erwartenden Wachstumsverläufe. Wir erhalten aus dem Bonitätsfächer keine Informationen über die mittlere Wachstumskurve und über das mögliche Ausmaß der Variation um sie herum.

Die in der Modellanpassung geschätzten Varianzen der Zufallsparameter klären dagegen das Ausmaß der Variation auf verschiedenen Hierarchieebenen zwischen und innerhalb von Gruppierungsmerkmalen auf. Durch die Einführung einer Hierarchie von Gruppen in ein gemischtes Modell wird sparsam mit Freiheitsgraden umgegangen. Bei traditionellen Site-index-Modellen müsste jede Gruppe mit Indikatorvariablen (Dummy-Variablen) kodiert werden; dieses wäre mit einem drastischen Verlust an Freiheitsgraden verbunden.

Außerdem gelangt die Methode auch bei einer gruppierten Datenstruktur mit ungleichmäßig vielen Beobachtungen pro Gruppe zu unverzerrten Parameter-Schätzungen. Die Gruppierung von Bäumen an Stichprobenpunkten ist ein einfacher Ansatz zur Berücksichtigung räumlicher Korrelation (Lappi und Bailey, 1988). In weiteren Studien könnte das Verhalten der Zufallsparameter auf Stichproben-Ebene im Raum untersucht werden. Denkbar wäre, dass die Zufallsparameter auf Plot-Ebene, abhängig von der Distanz zwischen den Plots, unterschiedlich stark miteinander korrelieren. Für Orte ohne Beobachtungen könnten die Zufallsparameter dann mittels räumlicher Statistik vorhergesagt werden. Für diese Art der Kalibrierung könnten die Werte von Zufallsparametern aus Variogrammen vorhergesagt werden. In der hier vorgestellten Methode der Kalibrierung kombinieren die BLUPs verfügbare Beobachtungen und zuvor geschätzte Varianzen der Zufallsparameter in optimaler Weise.

Unter den aktuell praktizierten waldbaulichen Behandlungsrichtlinien erhöht sich die Variation innerhalb der Waldbestände. Einzelne Maßzahlen wie die Höhenbonität verlieren an Aussagekraft und können die Entwicklungen der Wirklichkeit nur unzureichend beschreiben. Gegenwärtig gewinnen deshalb Einzelbaumwachstumsmodelle, wie sie in den Simulatoren BWinPro und SILVA implementiert sind, immer mehr an Bedeutung. Die zugrunde liegenden Wachstumsmodelle sind jedoch, wie weiter unten ausgeführt wird, noch stark an das potentielle Höhenwachstum von Oberhöhenbäumen angelehnt. Hierarchische, nichtlineare und gemischte Modelle würden einen weiteren Fortschritt bei der Konstruktion solcher Simulatoren darstellen. Durch Implementierung von gemischten Modellen in Wachstumssimulatoren könnte mit Vertrauensbereichen für die Vorhersagen Auskunft über deren Glaubwürdigkeit gegeben werden. Die geschätzten Werte für den

Höhenerwartungswert-Parameter des HNLMM können aber auch dazu verwendet werden, um Aussagen über die Wuchskraft der Bestände und Standorte zu machen, der Parameter selbst hat hier jedoch im Gegensatz zu den Site-index-Modellen keine Bedeutung für die Methode von Vorhersagen (Lappi und Bailey, 1988).

5.1.3 Einsatzgebiet des hierarchischen, nichtlinearen und gemischten Modells

Im folgenden Abschnitt soll erläutert werden, wie das erstellte Modell in den Informationsfluss der Forstwirtschaft integriert werden kann und welchen praktischen Nutzen diese Arbeit hat.

In Baden-Württemberg sind die Richtlinien zur waldbaulichen Behandlung von Waldbeständen grundlegend verändert worden (Landesforstverwaltung Baden-Württemberg, 1999). Statt gleichaltriger Reinbestände, sollen mit dem Konzept der naturnahen Waldwirtschaft vielerorts gestufte und stabile Mischbestände erzielt werden. Die Bewertung der Bestände und ihrer Standorte gemessen an der Holzertragsfähigkeit hat jedoch weiterhin Bedeutung. Doch sind viele Annahmen, die den traditionellen Site-index-Modellen zugrunde liegen, nicht haltbar. Nach der Bonitierung wird angenommen, dass alle anderen Größen, die das Waldwachstum bestimmen, ebenfalls bekannt seien (Wenk et al., 1990, S. 241). Darüber hinaus wird unterstellt, dass das weitere Wachstum getreu dem Bonitätssystem verläuft.

In nahezu allen staatlichen Forstbetrieben Baden-Württembergs sind mindestens zu einem Termin Betriebsinventuren mit systematischen Stichproben auf Basis von konzentrischen Probekreisaufnahmen vorgenommen worden. Mit Fortschreibungen dieser Inventurdaten zu Prognosezwecken wäre eine verlässliche Herleitung von Hiebssätzen im Rahmen der Forsteinrichtung möglich. Für die Prognose des Wald-Wachstums werden Einzelbaum-Wachstumsmodelle notwendig. Traditionelle Site-index-Modelle können die Wirklichkeit nicht mehr in ausreichendem Maß beschreiben und nutzen die umfassenden Informationen aus Waldinventuren nicht genügend aus.

Pretzsch (2001, S. 264 ff.) stellt den Einsatz seines Wachstumssimulators SILVA 2.2 für eine dynamische Planung auf Betriebsebene im Stadtwald Traunstein vor. Das Höhenwachstumsmodell des SILVA-Wachstumssimulators basiert auf einem deterministischen Ansatz. In Abhängigkeit von neun Standortsvariablen werden die Parameter einer potentiellen Höhenwachstumskurve festgelegt (Pretzsch, 2001, S. 199-204). Als Grundmodell dient eine Richards-Funktion mit drei Parametern. Die Werte der Standortsvariablen werden über unimodale Dosis-Wirkungs-Funktionen auf das Intervall $[0; 1]$ transformiert. Die Wirkungswerte von jeweils drei Standortsvariablen werden zu drei ökologischen Faktoren aggregiert. Die drei Faktoren werden dann mit Hilfe von Regressionsparametern kombiniert, die aus langfristigen Versuchsflächen hergeleitet wurden. Dadurch gelangt man zum Schätzwert für den asymptotischen Parameter. Über Hilfsbeziehungen, die ebenfalls aus langfristigen Versuchsflächen hergeleitet wurden, können dann der Form- und der Steigungsparameter hergeleitet werden. Damit sind die Parameter des potentiellen Höhenwachstums bestimmt. Im SILVA-Wachstumssimulator wird der Höhenzuwachs des einzelnen Baumes bestimmt, indem aus der Umkehrfunktion der Höhenwachstumsgleichung für die beobachtete Höhe eines Baumes das sogenannte theoretische Alter kalkuliert wird (Pretzsch, 2001, S. 222-223). Der theoretische, potentielle Höhenzuwachs ergibt sich aus der Differenz der berechneten Baumhöhe im theoretischen Alter zu der berechneten Baumhöhe am Ende des Prognoseintervalls. Der geschätzte, tatsächliche Höhenzuwachs wird schließlich ermittelt, indem der theoretische, potentielle Höhenzuwachs durch Maßzahlen der Kronenmantelfläche und der Konkurrenzwirkung Korrektur erfährt. Im weit akzeptierten Ansatz des SILVA-Simulators werden alle Einflussgrößen als bekannt oder ohne Fehler schätzbar angenommen. Es werden keine stochastischen Komponenten zugelassen; die unbekannte Reststreuung wird vernachlässigt. Dem Höhenzuwachs-Modell liegt keine simultan angepasste Kurvenschar zugrunde, sondern separate Trajektorien für Kombinationen von Standortsvariablen. Die Standorts-Regressoren verschieben die Trajektorien zunächst flexibel; der Höhenzuwachs für einzelne Bäume wird jedoch nicht durch weitere Verschiebung auf der Ordinate, sondern durch Veränderung der Alterswerte auf der Abszisse ermittelt.

Etwas einfacher ist das Modell für das Höhenwachstum im Waldwachstumssimulator BWinPro 6.1 (Nagel, 1999; Nagel et al., 2002). Im ersten Schritt wird mit der Bestimmung der Oberhöhe im Alter von 100 Jahren die Bonität bestimmt. Die zugrunde liegenden Oberhöhen-Wachstumsverläufe verschiedener Baumarten stammen aus Ertragstafel-Modellen verschiedener Autoren. Für die Baumarten Eiche, Roteiche und J.-Lärche ist das Oberhöhen-Wachstum über dem Alter mit einer Richards-Funktion beschrieben worden; für Kirsche, Winterlinde und Bergahorn ist die Trendfunktion nach Sloboda (1971) anhand von empirischen Richtungsfeldern angepasst worden; für die Oberhöhen-Wachstumsverläufe von Buche, Esche, Fichte, Sitkafichte, Douglasie, Kiefer und E.-Lärche passte Nagel Regressionsmodelle an die Ertragstafelwerte an. Aus diesen Wachstumsfächern wird für die entsprechende Bonität der potentielle Zuwachs für die Oberhöhe iH_{pot} entnommen. Bezogen auf die Bonität ergibt sich dann der relative potentielle Zuwachs mit $iH_{relpot} = \frac{iH_{pot}}{H_{100}}$. Der geschätzte relative Höhenzuwachs kann dann ermittelt werden mit $iH_{rel} = iH_{relpot} + p_1(\frac{H_{100}}{h})^{p_2} + \epsilon$. Beim BWinPro-Wachstumssimulator werden also Bonitätsfächer eingesetzt, die für die gesamte Population angepasst wurden. Das Wachstum des einzelnen Baumes wird über die Relation seines Höhenwertes zum Wert der Oberhöhe hergeleitet. Die Größe ϵ ist eine um den Mittelwert 0 normalverteilte Zufallsvariable. Diese bewirkt eine größere Flexibilität für das Modell. Theoretisch ist ein „Umsetzen“ der Bäume möglich, d.h. die Wachstumskurven können sich überschneiden.

In weiterführenden Studien soll in Anlehnung an Fang et al. (2001) ein bivariates, simultanes, gemischtes Modell konstruiert werden, durch das neben dem Höhenwachstum der Bäume auch gleichzeitig deren Durchmesserwachstum beschrieben werden könnte. Es ist denkbar, dass das HNLMM in Wachstums-Simulatoren eingesetzt wird, um Bestandesdaten oder gar Inventurdaten ganzer Forstbetriebe fortzuschreiben. Bei der Anpassung dieses Modells wird die Variation innerhalb und zwischen Beständen oder Stichprobenpunkten geschätzt. Der Informationsgehalt vorliegender Beobachtungen kann optimal ausgenutzt werden. Im Rahmen zuvor geschätzter Varianzen können die Zufallsparameter vorhergesagt und die Wachstumskurven individuell und flexibel kalibriert werden. Die Variation bleibt auch nach Fortschreibungen erhalten. Während die Modellanpass-

ung komplexe Algorithmen zur numerischen Maximierung der Likelihood benötigt, erfolgt die Kalibrierung zur Vorhersage der Zufallsparameter rasch und effizient auf Basis einfacher Matrix-Algebra. Die zur Modellanpassung verwendete `nlme`-Bibliothek der `R`-Statistiksoftware ist in ihrer Bedienung einfach zu erlernen und trug im wesentlichen zum Gelingen dieser Arbeit bei. Aus diesen Gründen ist der Komplexitätsgrad dem Modellzweck unbedingt angemessen.

5.2 Waldwachstums- und ertragskundlicher Erkenntnisgewinn

Das ursprüngliche Ziel der umfangreichen Studie von Mitscherlich (1957, S. 171) war es, den Wachstumsgang der Fichte in Baden-Württemberg neu zu beschreiben. Ein Anlass dazu waren häufige Hinweise aus der Praxis, dass die damals gebräuchlichen Ertragstafeln von Zimmerle (1949) das Wachstum der Fichte nicht richtig beschreiben würden. Mitscherlich (1957) trug die Bonitäten der Versuchsflächen über ihrem Alter auf. Ursprünglich vermutete er, dass sich die Bonitäten in verschiedenen Landschaften Badens unterschiedlich über das Bestandesalter verändern würden (Mitscherlich, 1957, S. 177). Er zeigte sich überrascht, dass diese Veränderungen im Bonitätsverlauf sehr gering ausfielen und in manchen Landschaften gar zu nicht zu existieren schienen. So war z.B. beim Vergleich der Bonitätsverläufe des Hochschwarzwaldes mit denen des Bodenseegebiets kein Unterschied festzustellen.

Auf fast allen Standortsgruppen konnte Mitscherlich (1957, S. 183) jedoch einen Bonitätsanstieg in der Jugend beobachten. Er erklärte die höheren Bonitäten von jungen Fichtenbeständen im Badischen Landesteil gegenüber den Württembergischen Ertragstafeln von Zimmerle (1949) mit intensiverer Kulturpflege, weiteren Pflanzverbänden und frühzeitigeren Läuterungen (Mitscherlich, 1957, S. 224). Das Zurückbleiben der Badischen Fichtenbeständen gegenüber der Ertragstafel in höheren Altern führte er darauf zurück, dass sich Zimmerle (1949) bei der Aufstellung der Ertragstafeln vor allem auf die Beobachtungen in Versuchsflächen stützte, die in sehr wuchskräftigen oberschwäbischen

Beständen lagen. Trotz der Auswertung von mehr als 30000 Bohrspänen und den Höhentriebsmessungen an 3000 Fichten auf 309 Probeflächen gelang es Mitscherlich (1957) nicht (S. 231), „vom Wachstumsgang der Fichte in Baden ein anschauliches und klares Bild zu gewinnen“. Aus den Beobachtungen der recht umfangreichen Erhebungen konnte er für die Landschaften oder Standortsgruppen keine „bindenden Regeln oder strenge Gesetzmäßigkeiten“ ableiten. Besonders bei der Auswertung der Bohrspäne kam er zum Ergebnis, dass es *unbekannte Ursachen* für das Ausmaß der Streuung geben muss. Mitscherlich (1957) schlug deshalb vor, noch erheblich umfangreicheres Datenmaterial zu sammeln, um endgültige Klarheit über die Zusammenhänge des Wachstums zu gewinnen. Die Erkenntnisse von Mitscherlich (1957) zeigen, dass es außerordentlich schwierig ist, das Wachstum der Bäume anhand bekannter Gesetzmäßigkeiten allgemein gültig zu formulieren und in abstrakter Darstellungsweise in Ertragstafeln festzuschreiben. Nebenbei sei angemerkt, dass damit auch das hohe Ausmaß an unerklärbarer Variation verloren gehen würde. Noch umfangreichere Erhebungen als sie Mitscherlich (1957) ohnehin schon durchführte, hätten vermutlich auch nicht zur besseren Erklärung beigetragen.

Der in dieser Studie verwendete Modell-Ansatz mit Zufallsparametern kann vor diesem Hintergrund als sinnvolle und effiziente Lösung gewertet werden. Im ersten Augenblick enttäuschte es zwar, dass Standortsvariablen keine signifikante Regressoren für fixe Modell-Effekte darstellen konnten. Doch bleibt die beobachtete Variation durch die Zufallsparameter erhalten. Anhand von Beobachtungen an Stichproben oder auf Versuchsflächen können die Zufallsparameter vorhergesagt und das Modell effizient kalibriert werden. Wir sind zum gegenwärtigen Zeitpunkt folglich nicht in der Lage, einen Großteil der beobachteten Streuung mit effizienten Methoden der Datengewinnung und Modellkonstruktion zu erklären. Wir können aber das Ausmaß der Streuung und zu erwartende mittlere Kurvenverläufe für das Wachstum schätzen. Über die Abweichung der Beobachtungen von den mittleren Kurven können die individuellen Wachstumsverläufe ermittelt werden und die Beobachtungen in die Zukunft fortgeschrieben werden.

Seit längerem wird das weit verbreitete Phänomen beobachtet, dass junge Bestände im Durchschnitt offensichtlich bessere Bonitäten zu haben scheinen als ältere Bestände.

Zunächst gibt es dafür ungeachtet aller übrigen Erklärungsversuche nach Monserud (1984) eine sehr einfache Ursache: Die absinkenden Bonitäten mit steigendem Bestandesalter spiegeln die frühzeitigeren Endnutzungen auf besseren Standorten wider. Eine ähnliche Auffassung teilt auch Mitscherlich (1957); er beobachtete, dass Bestände besserer Bonität in geringeren Produktionszeiträumen bewirtschaftet werden.

Mit Hilfe des Studiums langfristiger Versuchsflächen und der Analyse von Longitudinaldaten konnten zahlreiche Arbeiten längerfristige Veränderungen von Wachstumsniveaus beschreiben. Mitscherlich (1957, S. 182) kam zu dem Schluss, dass Anstiege der Bonität in Jugendjahren auf eine veränderte waldbauliche Praxis zurückzuführen waren, wie die Hinkehr zu Bestandesbegründungen durch Pflanzung statt Saat und neuerdings unternommene frühzeitige Läuterungen. Wie schon kurz im Abschnitt 4.1.3.2 angeführt wurde, konnten Schöpfer et al. (1994) einen Bonitätsanstieg der Fichte in jüngerer Vergangenheit herleiten. Bei Schöpfer et al. (1994) findet sich auch eine äußerst interessante Untersuchung zur aktuellen Veränderung des eigentlichen Höhenwachstums. Sie teilten die Bäume der IWE aus dem Jahr 1983 aufgrund ihrer Alter und der Höhen im Jahr der Fällung in fünf relative Ertragsklassen ein, die jeweils 20% der statistischen Masse umfassten. Für diese Kollektive wurden mit deren durchschnittlichen jährlichen Höhenzuwächsen in Alters- und Höhenklassen empirische Richtungsfelder aufgestellt und mit Bonitätsfächern ausgeglichen. Die Kurven des angepassten Bonitätsfächers passten aber überhaupt nicht überein mit den mittleren Wachstumskurven für die Kollektive über deren gesamtes Lebensalter. Die Zuwächse aus den Beobachtungsperioden jüngerer Vergangenheit zeigten einen starken Anstieg. Dieses hohe Zuwachsniveau hatten die Bäume nicht ihr ganzes Leben lang, sondern nur in den letzten Jahrzehnten. Damit zeigten Schöpfer et al. (1994) nebenbei, dass herkömmliche Ertragstafeln die aktuelle Bestandesentwickklung nur unzulänglich wiedergeben können. Sie konnten herleiten, dass sich der von Mitscherlich (1957) beobachtete Bonitätsanstieg von alten zu jungen Beständen in jüngerer Vergangenheit noch verdoppelt hat. Schöpfer et al. (1994) führten dafür als Erklärungen an, dass die gestiegenen atmosphärischen Stickstoffeinträge der vergangenen Jahrzehnte auf ärmeren und gut wasserversorgten Standorten einen besonders hohen Dünge-Effekt bewirkt ha-

ben müssen. Diese Aussagen zu den Wachstumstrends decken sich mit den Erkenntnissen meiner Studie. Danach stieg der Höhenerwartungswert für eine 50-jährige Fichte in der näheren Vergangenheit im Schwarzwald um rund 13 cm pro Jahr (s. Abb. 4.18). Dieser Anstieg wird in dem hier entwickelten Wachstumsmodell durch den Koeffizient für das durchschnittliche Keimjahr an den Stichprobenpunkten als fixer Effekt ausgedrückt.

Untheim (1996) verglich die Baumhöhen von Fichten gleichen Alters zu unterschiedlichen Jahren der Keimung anhand von gefällten Probestämmen auf der Ostalb. Für einen etwa 60-jährigen Beobachtungszeitraum konnte er zeigen, dass die Höhen im Alter von 50 Jahren auf drei verschiedenen Standortseinheiten um 10,8 bis 12,4 m pro 100 Jahre anstiegen. Das zeigt, dass auch der Koeffizient des HNLMM für die IWE Daten aus der Region Odenwald mit einem rechnerischen Anstieg von 14 m pro 100 Jahre im plausiblen Wertebereich liegt. Für die IWE-Daten von den Moränen war der Koeffizient nicht signifikant. Für die historischen Mitscherlich-Daten von den Moränen bewirkte der Koeffizient für das Keimjahr einen Anstieg der Höhe um 10,3 m pro 100 Jahre. Vermutlich hat sich der Dünge-Effekt der atmosphärischen Einträge auf den Hochleistungsstandorten der Moräne in der jüngeren Zeit erschöpft.

Die Studie von Moosmayer et al. (2003) basiert auf den Höhenmessungen der IWE aus den Jahren 1983 und 1988. Sie untersuchten die Differenz von durchschnittlichen Höhen zu bestimmten Altern zwischen zwei benachbarten 20-jährigen Klassen von Keimjahren differenziert für verschiedene Humusformen und Wuchsregionen. Die Höhendifferenzen streuten stark; klare Abhängigkeiten waren nicht auszumachen. Die Baumhöhen waren in allen Altersklassen bei einer 20 Jahre späteren Keimung zwischen 3,5 m höher bis 0,5 m niedriger. Der Schwerpunkt der Streuwolken lag bei etwa 1,5 m, um die die Baumhöhen nach 20 Jahren späterer Keimung im gleichen Alter höher gewesen sind.

Wenk und Vogel (1996) analysierten das Höhenwachstum von 2277 Fichten aus 448 Beständen in niederen Bergland-Regionen Ostdeutschlands. Sie kamen zu dem Ergebnis, dass junge Bestände in heutiger Zeit schneller wachsen. Die Änderung des Wachstumsverhaltens setzte ungefähr zum Anfang des vorletzten Jahrhunderts ein und setzt sich in heutiger Zeit fort. Wenk und Vogel (1996) kamen zu der Erkenntnis, dass Ertragsta-

feln solche Veränderungen in den Wachstumsmustern nicht berücksichtigen können. Das Wachstum junger Bestände dürfe deshalb nicht mit Ertragstafeln für Perioden in weiter Zukunft fortgeschrieben werden. Künftig sollten passende Wachstumsmodelle für Zwecke der Fortschreibung entwickelt werden. Für dieses Problem bietet das HNLMM dieser Studie eine Lösung. Der Wachstumsverlauf der Bäume wird mit diesem Modell nicht als konstant für alle Zeit angesehen. Einerseits sorgen die Koeffizienten des Keimjahres als fixer Effekt für eine dynamische Beschreibung des Wachstums. Andererseits kann anhand von Beobachtungen vor Ort über Vorhersagen der Zufallsparameter das Modell effizient kalibriert werden. Dadurch können mittelfristige Wachstumsänderungen erkannt und fortgeschrieben werden. Genau diese dynamischen und selbstkalibrierenden Modell-Eigenschaften werden von Wenk und Vogel (1996) gefordert. Sie fordern auf, dass Modelle entwickelt werden, die es uns ermöglichen, Wachstumsveränderungen direkt aus dem Wachstumsverhalten selbst herzuleiten. Wenk und Vogel (1996) führen als Gründe für Veränderungen des Wachstums den Wechsel von Saat zu Pflanzung, weitere Pflanzverbände, intensivere Pflege, Stickstoffdeposition und den Anstieg von Temperatur und CO_2 an. Anhand von Wiederholungsmessungen aus der nationalen Forstinventur Österreichs entdeckte Schadauer (1996) einen Anstieg des Höhenzuwachses bei jungen Fichten bis ins Alter von 40 Jahren. Als Ursache vermutete er ein erhöhtes Stickstoffangebot durch Regeneration der Böden nach dem Verzicht auf Streunutzung und durch höhere atmosphärische Einträge. Nach Stratifizierungen fand Schadauer (1996) Zuwachsanstiege auf Standorten mit ursprünglich geringem Angebot an Stickstoff; in einigen Fällen seien die Ergebnisse jedoch widersprüchlich gewesen.

Zusammenfassend kann das hier entwickelte HNLMM zeitliche Veränderungen im Wachstum beschreiben, einerseits durch den fixen Effekt des Keimjahres und andererseits durch die Kalibrierung an Beobachtungen über Vorhersagen für die Zufallsparameter. Die modellierten Anstiege der Höhenzuwächse liegen in einem Rahmen, der von den Ergebnissen anderer Autoren bestätigt wird.

5.3 Würdigung der zugrunde liegenden Daten

In diesem Abschnitt sollen die Daten hinsichtlich ihrer Brauchbarkeit für den Zweck dieser Studie gewürdigt werden. Die vorliegenden Daten aus Mitscherlichs Erhebung und aus den Inventuren der IWE stellen wertvolle longitudinale Beobachtungsreihen dar. Es sei an dieser Stelle noch einmal darauf hingewiesen, dass für diese Arbeit keine zusätzlichen neuen Daten erhoben wurden, die Kosten verursacht hätten. Dem unermüdlichen Forscherdrang von Prof. Schöpfer und Prof. Hradetzky ist es zu verdanken, das die historischen Daten Mitscherlichs digital gesichert werden konnten. Prof. Schöpfer ist es daneben auch zu verdanken, dass die Daten aus der IWE vorliegen, auf die schon zahlreiche und gehaltvolle Studien aufbauten. Die ersten Inventuren der IWE in den Jahren 1983 und 1988 unterstanden seiner Verantwortung. Entgegen der Erhebung von Mitscherlich sind die Stichproben der IWE in einem systematischen Raster angelegt. Für höher aggregierte Aussagen mangelt es den Daten Mitscherlichs folglich an räumlicher Repräsentativität. Dieses Problem wurde dadurch umgangen, dass die Daten nur für gemeinsame Regionen zusammengeführt wurden.

Beide Inventuren sind geclusterte Stichproben mit mehreren Bäumen pro Stichprobenpunkt. Es war möglich, die Variation des Baumwachstums innerhalb und zwischen den Stichprobenpunkten zu beschreiben. Bei den periodisch wiederkehrenden Aufnahmen der IWE kann es durch Veränderungen der Waldbestände vorkommen, dass die zwei Probestämme vor Ort aus einem anderen Bestand entnommen worden sind als bei der Voraufnahme. Zu Beginn dieser Studie wurden die Höhenbeobachtungen für jede Stichprobe visuell geprüft; dabei zeigte sich, dass dieses nur selten der Fall war.

Die Höhentriebe an den gefällten Stämmen der IWE wurden jahresweise für einen Zeitraum von 10 bis 25 Jahren zurückverfolgt. Mitscherlich analysierte die Höhenzuwächse für Perioden von fünf Jahren rund 60 Jahre in die Vergangenheit zurück. Anhand der Mitscherlich-Daten konnte das Wachstum eines jeden Baumes für eine größere Spanne der Lebensalter untersucht werden. Da so weit zurück häufig keine Anzeichen von Astquirlen oder Knospenschuppenspuren zu erkennen waren, wurden Sektionsschnitte zur Synchro-

nisation durchgeführt. Bei der IWE ist dieses Vorgehen in der heutigen Zeit in diesem Umfang nicht mehr zu leisten; das Holz der gefällten Probebäume soll noch in sinnvoller Weise wirtschaftlich verwertbar sein.

Der Informationsgehalt der jährlichen Messungen bei der IWE wurde in dieser Studie nicht vollends ausgeschöpft. Die jährliche Oszillation wurde lediglich mit einer stochastischen, normalverteilten Störgröße modelliert. Nur für komplexe Modellierungen des Restfehlers wären diese jährliche Messungen von Nutzen. Es sollte deshalb diskutiert werden, das Raster der IWE weniger intensiv zu gestalten, dafür aber die Höhentriebe an den einzelnen Bäumen über Sektionsschnitte weiter zurückzuverfolgen.

Aus den Daten der Standortskartierung und den Raster-Geodaten für Klimagrößen ließen sich für diese Arbeit keine signifikanten Prädiktoren ableiten. Die zu beobachtende Variation im Baumhöhenwachstum der Fichte ließ sich mit Hilfe von Standortsvariablen nicht erklären. Die Modellierungsversuche auf Basis dieser Datengrundlagen gerieten wegen der geringen Aussagekraft schnell in Gefahr, ineffizient zu werden. Wie in dieser Arbeit gezeigt wird, eignen sich Beobachtungen aus Waldinventuren hervorragend dazu, das Modell an örtliche Verhältnisse zu kalibrieren. Vor diesem Hintergrund wird es als unverzichtbar angesehen, bereits etablierte Verfahren des forstlichen Monitorings fortzuführen und weiter zu stärken. Deshalb müssen die Organisationsstrukturen zur Durchführung und Administration von Waldinventuren auf Ebene der Forstbetriebe (Betriebsinventuren), des Landes (Waldwachstumskundliche Erhebung '99) und auf nationaler Ebene (Bundeswaldinventur) aufrechterhalten und weiterhin gefördert werden. Diese Waldinventuren liefern zugleich die Datenbasis, anhand derer das hier entwickelte HNLMM durch Vorhersagen der Zufallsparameter im Rahmen der Kalibrierung praktische Anwendung erfahren kann.

Kapitel 6

Schlussfolgerung

Zusammenfassend wird nun dargelegt, in wieweit die aufgestellten Hypothesen ihre Gültigkeit behalten. In dieser Arbeit wurde ein nichtlineares, hierarchisches und gemischtes Modell für das Baumhöhenwachstum der Fichte erstellt. Das Modell wurde an longitudinale Beobachtungsreihen von Höhentriebsanalysen angepasst. Die Schätzungen für die Parameter der fixen Effekte und für die Varianzen der Zufallsparameter, sowie die BLUP-Vorhersagen der Zufallsparameter selbst führten zu biologisch plausiblen Höhenwachstumskurven auf Basis der nichtlinearen Trendfunktion nach Sloboda (1971). Die Residuen waren unverzerrt. Durch Zufallsparameter auf zwei verschiedenen Hierarchieebenen wurde der gruppierten Datenstruktur Rechnung getragen. Die vorhergesagten empirischen Zufallsparameter verstießen nicht gegen die zuvor getroffenen Annahmen über deren Verteilung. Damit ist die erste Hypothese haltbar, die methodischen Charakter hatte. Das Baumhöhenwachstum der Fichte in Baden-Württemberg lässt sich mit einem nichtlinearen, hierarchischen Modell beschreiben.

Die Modellanpassung geschah gleichzeitig an Daten aus aktuellen Immissionsökologischen Waldschadensinventuren (IWE) und aus der Erhebung von Mitscherlich (1957), welche eine Epoche mit möglicherweise anderen Wachstumsbedingungen abdeckt und in dieser Arbeit daher auch als „historisch“ bezeichnet wird. Die Daten beider Aufnahmen wurden für drei Regionen (Odenwald, westlicher Teil von Alt- und Jungmoräne des südwestdeutschen Alpenvorlandes, Südschwarzwald und Baar-Wutach) zusammengeführt.

Jedoch unterschieden sich die Kollektive der Probebäume zwischen beiden Erhebungen. Während für die Studie Mitscherlichs ausschließlich vorherrschende Bäume gefällt wurden, enthalten die Daten der IWE auch Bäume niederer soziologischer Stellungen. Bei der IWE wurde zudem ein höherer Anteil an alten und mattwüchsigen Bäume beprobt. Gemeinsame Aussagen konnten für die Daten beider Inventuren deshalb nicht getroffen werden. Das Modell musste das Höhenwachstum beider Inventuren mit Indikatorvariablen getrennt beschreiben. Innerhalb der beiden Inventuren drücken Koeffizienten für Indikatorvariablen jeweils ein unterschiedliches Wachstumsniveau in den drei Regionen aus. Durch diese Interzepte für fixe Effekte konnte die räumliche Variation abgebildet werden. Das durchschnittliche Keimjahr am Stichprobenpunkt zeigte bei den aktuellen IWE-Daten aus dem Odenwald und dem Schwarzwald und bei den historischen Daten von den Moränen einen signifikanten Einfluss auf den Höhenerwartungswertparameter. Die individuelle Abweichung vom durchschnittlichen Keimjahr am Stichprobenpunkt konnte für den einzelnen Baum zusätzliche Variation erklären. Mit den Parametern für diese fixen Effekte gelang es, zeitliche Trends für das Baumhöhenwachstum modellhaft zu beschreiben und damit die dritte Hypothese zu verifizieren.

In einer Anwendung wurde das mögliche Einsatzgebiet des nichtlinearen, hierarchischen und gemischten Modells aufgezeigt. Das Modell wurde anhand von longitudinalen Daten evaluiert, die in einer aktuellen und unabhängigen Studie im Südschwarzwald erhoben wurden. Dabei wurden in einer Variante die jeweils ersten Beobachtungen und in einer zweiten Variante die ersten und elften Beobachtungen als bekannt angenommen. Mit diesen Informationen konnten die Zufallsparameter durch BLUPs vorhergesagt werden. In einem effizienten Verfahren konnten auf diese Weise auch für ungleichaltrige Bestandessituationen individuelle Höhenkurven für alle Bäume kalibriert werden. Mit den Schätzkurven wurden danach Fortschreibungen zur Prognose vorgenommen und die Vorhersagen mit den als unbekannt angenommen restlichen Beobachtungen verglichen. Die Prognosefehler waren nahezu unverzerrt. Für ein Vergleichsmodell wurden Bonitätsfächer an empirische Richtungsfelder für die aktuellen IWE-Daten angepasst. Mit den Bonitäts-

fächern wurden ebenfalls Fortschreibungen vorgenommen, wobei analog zur Kalibrierung die ersten bzw. elften Beobachtungen als Startinformationen dienten.

Mit nur einer Beobachtung als Startinformation hatte das gemischte Modell erst ab einem Prognosezeitraum von 21 Jahren im Mittel geringere Fehler als die Bonitätsfächer und stellte sich deshalb nicht wesentlich besser dar. Bei zwei Vorbeobachtungen zur Kalibrierung waren die Fehler des gemischten Modells ab dem elften Fortschreibungsjahr betragsmäßig geringer. Die Streubreite der Fehler war zudem merklich kleiner als bei den Fortschreibungen mittels der Bonitätsfächer. Insgesamt zeigte sich also das nichtlineare, hierarchische und gemischte Modell vorteilhafter. Damit kann schließlich auch die dritte und letzte Hypothese aufrecht erhalten werden.

Kapitel 7

Zusammenfassung

Die Abschätzung des Volumenzuwachses von Waldbeständen ist für betriebliche und wissenschaftliche Fragestellungen bedeutsam. Der Zuwachs für Stichprobenkollektive oder gesamte Forstbetriebe ergibt sich aus Waldinventuren über Hochrechnungen der Messungen an einzelnen Bäumen. Der Volumenzuwachs eines einzelnen Baumes für eine bestimmte Periode ist die Differenz seiner Volumina zu den Zeitpunkten am Periodenende und -anfang. Das Volumen zu einem bestimmten Zeitpunkt kann über Regressionen in Abhängigkeit von leicht messbaren Größen wie dem Brusthöhendurchmesser (Bhd) und der Höhe beschrieben werden (Kublin, 2003). Für nur einmalig erfasste Bäume, bei denen keine Messungen zu zwei Zeitpunkten vorliegen, kann der Zuwachs der abgelaufenen Periode nicht direkt über Volumenbilanzierung ermittelt werden. Zunächst müssen für den Bhd und die Höhe Schätzungen für den Zeitpunkt ohne Messwerte vorgenommen werden. Damit kann das Volumen hergeleitet werden, das schließlich in die Bilanzierung eingeht. Die Schätzungen für den Bhd und die Höhe stammen aus Wachstumsmodellen. Besondere Bedeutung haben die Wachstumsmodelle für Fortschreibungen zum Zweck von Prognosen.

In dieser Studie wurde ein Höhenwachstums-Modell am Beispiel der Fichte (*Picea abies (L.) Karst.*) in Baden-Württemberg erstellt. Die zugrunde liegenden Daten waren longitudinale Beobachtungen aus Trieblängenmessungen an gefällten Bäumen. Zum einen stammen diese Daten aus aktuellen Aufnahmen der Immissionsökologischen Waldschaden-

serhebung (IWE) und zum anderen aus historischen Aufnahmen von Mitscherlich (1957). In dieser Studie wurde ein nichtlineares, hierarchisches und gemischtes Modell konstruiert. Eine nichtlineare Wachstumsfunktion ist sparsam im Gebrauch von Freiheitsgraden und liefert auch im Extrapolationsbereich biologisch plausible Werte für die Baumhöhe als Response-Variable. Als nichtlineare Grundfunktion zeigte die Trendfunktion nach Sloboda (1971) die besten numerischen Eigenschaften.

Die zugrunde liegenden Modelldaten besaßen eine gruppierte Struktur auf verschiedenen hierarchischen Ebenen. An jedem Stichprobenpunkt wurde eine Gruppe von mehreren Bäumen zur Vermessung gefällt, und für jeden Baum wurde eine Gruppe von Höhenmessungen über die Zeit erfasst. Weil die Höhenmessreihen der Bäume an einem Stichprobenpunkt miteinander korreliert waren und an den Stichprobenpunkten jeweils eine unterschiedliche Anzahl von Bäumen erfasst wurden, musste die hierarchische Gruppierung in der Modellkonstruktion berücksichtigt werden.

Die Variation des Baumwachstums wurde durch variierende Parameterwerte beschrieben. Dabei wurde ein Teil der zu beobachtenden Variation durch Regressionsparameter für bekannte Kovariablen als fixe Effekte erklärt. Neben dem erklärbaren Anteil der Parametervariation gab es einen großen Teil, der nicht erklärbar war. Für diesen wurde der Zufall als Ursache angenommen. Die zufälligen Effekte wurden durch nomalverteilte Zufallsparameter modelliert. Zufallsparameter sowohl auf der Ebene der Bäume, als auch auf der Ebene der Stichproben trugen der hierarchisch gruppierten Datenstruktur Rechnung. Die nichterklärbare Variation zwischen den Bäumen innerhalb einer Stichprobe wurde durch Zufallsparameter auf der Baum-Ebene beschrieben; diejenige zwischen den Stichproben durch Zufallsparameter für die Stichproben-Ebene. Außerdem enthielt das Modell eine stochastische Restfehlerkomponente für den Messfehler und die Oszillationen aufgrund schwankender Umwelteinflüsse. Mit Zufallseffekten ausschließlich auf der Stichproben-Ebene konnten mittlere Wachstums-Kurven für jeden Stichprobenpunkt analysiert werden. In dem gemischten Modellansatz wurden die fixen Parametereinflüsse und die Zufallsparameter miteinander kombiniert.

Ein Ziel dieser Studie war es, durch das Modell die zeitliche und räumliche Variation des Fichtenwachstums in Baden-Württemberg zu beschreiben. Die Mitscherlich-Daten waren konzentriert in drei Gebieten des badischen Landesteils erhoben worden. Nach einer Zusammenführung der IWE-Daten mit den Mitscherlich-Daten sollte untersucht werden, ob sich das Wachstum der Fichte in den drei Gebieten im Laufe der Zeit verändert hat. Es stellte sich jedoch heraus, dass die Baummessreihen der beiden Inventuren nicht unmittelbar vergleichbar waren. Bei der IWE wurde ein höherer Anteil an alten und mattwüchsigen Bäumen erfasst. Und obwohl bei beiden Aufnahmen vorwiegend dominante Bäume ausgewählt werden sollten, wurden zumindest bei der IWE im hohen Ausmaß Bäume aus den Kraft'schen Klassen 2 und niedriger beprobt. Es war beabsichtigt, die zeitliche Variation in Abhängigkeit des Kalenderjahres zu modellieren, in dem die Bäume eines Stichprobenpunktes im Mittel gekeimt waren. Junge Probebäume aus Mitscherlichs Aufnahme und alte aus der IWE konnten jedoch im gleichen Jahr gekeimt sein. Über die beiden Aufnahmen hinweg hatte das Keimjahr deshalb keine Aussagekraft. Für die Daten der beiden Inventuren konnten somit keine gemeinsamen Aussagen getroffen werden. Aber innerhalb der beiden Inventuren lagen die Aufnahmezeitpunkte der Stichproben dicht beieinander, weshalb das Keimjahr hier als fixer Effekt in das Modell eingehen konnte. Die Kombinationen aus den Regionen und den Inventuren wurden mit Indikatorvariablen kodiert. Mit verschiedenen Parameter-Interzepten für die Merkmalskombinationen ergaben sich unterschiedliche mittlere Wachstumskurven für die jeweiligen Inventurdaten aus den drei Regionen.

Für die IWE-Daten aus dem Odenwald und der Region Südschwarzwald / Baar-Wutach und für die Mitscherlich-Daten von den Moränen-Standorten konnte über das durchschnittliche Keimjahr an den Stichproben ein Teil der Variation des Höhenerwartungswert-Parameters signifikant erklärt werden. Demnach wurden die Fichten während des Beobachtungszeitraums der Mitscherlich-Daten im Alter von 50 Jahren mit jedem Jahr, das sie später aufgewachsen bzw. gekeimt waren, um 0,1 m höher. Für die IWE-Daten aus dem Odenwald und dem Südschwarzwald war ein Anstieg um 0,14 m pro Jahr zu verzeichnen. Für die Schätzung der Parameterwerte der fixen Effekte und der Vari-

anzen der Zufallsparameter und des Restfehlers wurden Maximum-Likelihood-Techniken angewendet. Wegen der nichtlinearen Responsefunktion gibt es keine explizite Darstellung der Likelihood. Deswegen wurde das Integral der marginalen Verteilungsfunktion der Response-Realisationen über eine Linearisierungsmethode aus der `nlme`-Funktion der `R`-Statistiksoftware approximiert.

Es wurde versucht, Kovariablen von Standortsgrößen als fixe Parametereffekte zu gewinnen. Die Koeffizienten zeigten jedoch keine Signifikanz.

Nach der Konstruktion wurde das Modell an unabhängigen longitudinalen Daten evaluiert, die nicht zur Anpassung verwendet wurden. Anhand von Beobachtungen, die als bekannt angenommen wurden, konnten die Zufallsparameter mit der a posteriori Kenntnis aus der Modellanpassung über deren Varianzen und der des Restfehlers durch BLUPs (best linear unbiased predictor) mit Bayes'schem Charakter vorhergesagt werden (siehe Vonesh und Chinchilli (1996, S. 362), Davidian und Giltinan (1995, S. 78)). In einer Variante wurde die jeweils erste Beobachtung eines Baumes als bekannt angenommen und in einer weiteren die jeweils erste und elfte. Die Matrix-Algebra wurde in `IML` von `SAS 9.1` programmiert. Die vorhergesagten Zufallsparameter führten zu individuellen Wachstumskurven, mit denen die Vorbeobachtungen fortgeschrieben wurden. Diese Art der Kalibrierung stellte sich einfach und effizient dar. Die Fortschreibungsergebnisse wurden mit den tatsächlichen Beobachtungen verglichen. Zur Bewertung des Erkenntnisgewinns wurden Vergleichsmodelle herangezogen. Dazu wurde die Differentialform der Trendfunktion nach Sloboda (1971) an verschiedene Varianten von empirischen Richtungsfeldern angepasst. Die Richtungsfelder bestanden aus den durchschnittlichen Höhenzuwächsen in Alters-Höhen-Klassen. Als Datengrundlage wurden einmal alle Beobachtungen der IWE-Daten herangezogen, die auch zur Anpassung des nichtlinearen, hierarchischen und gemischten Modells dienten und einmal nur diejenigen aus der Region Südschwarzwald / Baar-Wutach. In einem Anpassungsprozess wurden nur Felder mit mindestens fünf Beobachtungen benutzt, in einem weiteren alle Felder mit mindestens einer Beobachtung. Mit diesen Bonitätsfächern wurden dann Fortschreibungen zur Prognose vorgenommen. Zum

Vergleich mit dem gemischten Modell wurde einmal jede erste Beobachtung als Startinformation verwendet und einmal jede elfte.

Die Fortschreibungen mittels der Bonitätsfächer, bei der die ersten Beobachtungen als Startinformationen dienten, wurden mit denjenigen Ergebnissen verglichen, die aus der BLUP-Kalibrierung anhand der ersten Beobachtung resultierten. Die Prognosefehler der Vergleichsmodelle waren für einen Fortschreibungszeitraum bis zu 20 Jahren betragsmäßig genauso groß wie die aus den BLUP-Fortschreibungen. Gleiches galt für die Streubreite der Prognosefehler gemessen an den Perzentils-Abständen. Erst ab einem Prognosezeitraum von 21 Jahren waren die mittleren Prognosefehler des gemischten Modells geringer als die der Vergleichsmodelle. Als mit der jeweils ersten und elften Beobachtung zwei Vorbeobachtungen zur Kalibrierung des gemischten Modells verwendet wurden, waren die Prognosefehler des gemischten Modells bereits ab dem elften Jahr betragsmäßig geringer. Die Abstände zwischen dem ersten und dritten Quartil der Residuen-Verteilungen waren für das gemischte Modell ab dem sechsten Fortschreibungsjahr enger, zwischen dem 90er- und 10er-Perzentil ab dem vierten Jahr. Bei der Verwendung von zwei Vorbeobachtungen zur Kalibrierung zeigte sich das gemischte Modell treffgenauer und flexibler als die herkömmlichen Bonitätsfächer. Auch für Stichprobenpunkte in sehr ungleichaltrigen Beständen konnte das Wachstum der einzelnen Bäume plausibel vorhergesagt werden.

Das gemischte Modell beschrieb zusätzlich die Variation des Höhenwachstums zwischen den Stichprobenpunkten und innerhalb dieser und lieferte geschätzte mittlere Kurven für die Population und die einzelnen Stichprobenpunkte. Aufgrund eines wesentlichen methodischen Vorteils konnten für die Baumhöhenvorhersagen Vertrauensbereiche angegeben werden. Wegen der approximativen Schätzmethode lagen bei der Kalibrierung anhand einer Beobachtung als Vorinformation insgesamt 5% weniger Beobachtungen als erwartet in den Vertrauensbereichen, bei zwei Vorinformationen rund 1% weniger. Angesichts der effizienten Methode zur Vorhersage der Zufallsparameter und der optimalen Verwertung auch mehrerer Vorinformationen pro Individuum ist das nichtlineare, hierarchische und gemischte Modell uneingeschränkt zur Implementierung in Waldwachstumssimulatoren geeignet.

Kapitel 8

Anhang

	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5	Modell 6	Modell 7
log(Likelihood)	-35169.35	-32288.86	-32447.45	-32382.23	-32338.28	-32286.63	-31848.75
$\iota^{(2)}$	0.4757	0.4952	0.4915	0.5051	0.5245	0.5248	0.5261
	*0.0041****	*0.0056****	*0.0056****	*0.0084****	*0.0110****	*0.0109****	*0.0110****
$\iota^{(3)}$	0.8515	0.8576	0.8555	0.8550	0.8561	0.8563	0.8562
	*0.0021****	*0.0021****	*0.0021****	*0.0021****	*0.0022****	*0.0022****	*0.0022****
$\iota^{(4)}$	21.6136	21.5561	21.5423	24.0671	23.5013	23.5026	23.4837
	*0.0768****	*0.1960****	*0.1972****	*0.3379****	*0.4762****	*0.4261****	*0.4278****
$k_1^{(2)}$				0.0062	-0.0085	-0.0079	-0.0087
				0.0104	*0.0151*	*0.0151*	*0.0151*
$k_2^{(2)}$				-0.0332	-0.0125	-0.0125	-0.0139
				*0.0091***	*0.0131*	*0.0131*	*0.0131*
$k_3^{(2)}$					-0.0334	-0.0334	-0.0347
					*0.0138**	*0.0138**	*0.0138**
$k_4^{(2)}$					0.0021	0.0033	0.0034
					0.0131	*0.0130*	*0.0131*
$k_5^{(2)}$					-0.0785	-0.0786	-0.0797
					*0.0117****	*0.0117****	*0.0117****
$k_1^{(4)}$				-1.2683	-0.9258	-2.8887	-2.8385
				*0.4790**	*0.7277*	*0.7492****	*0.7504***
$k_2^{(4)}$				-4.5229	-0.5673	-0.5694	-0.5122
				*0.4183****	*0.6416*	*0.5732*	*0.5765*
$k_3^{(4)}$					1.1213	1.6569	1.6739
					0.6744	*0.6157**	*0.6196**
$k_4^{(4)}$					-4.0301	-4.4081	-4.6008
					*0.6274****	*0.5635****	*0.5656****
$k_5^{(4)}$					-3.9506	-3.9521	-3.9387
					*0.5673****	*0.5071****	*0.5097****
$k_6^{(4)}$						0.1396	0.1392
						*0.0262****	*0.0262****
$k_7^{(4)}$						0.1026	0.1033
						*0.0245****	*0.0247****
$k_8^{(4)}$						0.1276	0.1353
						*0.0157****	*0.0158****
$k_9^{(4)}$							0.1882
							*0.0094****
$Var\left[z_i^{(2)}\right]^{0.5}$	0.0943	0.0759	0.0760	0.0732	0.0679	0.0678	0.0679
$Var\left[z_i^{(4)}\right]^{0.5}$	4.4441	3.9998	4.0286	3.4893	3.4810	3.0889	3.1340
$Cov\left[z_i^{(2)}, z_i^{(4)}\right]^{0.5}$	0.086	0.189					
$Var\left[z_{ij}^{(2)}\right]^{0.5}$		0.0630	0.0611	0.0610	0.0612	0.0612	0.0614
$Var\left[z_{ij}^{(4)}\right]^{0.5}$		2.0191	1.9678	1.9681	1.9680	1.9669	1.6732
$Cov\left[z_{ij}^{(2)}, z_{ij}^{(4)}\right]^{0.5}$		-0.32					
σ	0.3638	0.3638	0.3649	0.3649	0.3648	0.3648	0.3649

Tabelle 8.1: *Parameter-Schätzwerte und deren Standardabweichungen aus den Modellanpassungen. p-Werte: 1%→*; 0,1%→**; 0,01%→***.*

	$\iota^{(2)}$	$\iota^{(3)}$	$\iota^{(4)}$	$k_1^{(2)}$	$k_2^{(2)}$	$k_3^{(2)}$	$k_4^{(2)}$	$k_5^{(2)}$	$k_1^{(4)}$
$\iota^{(2)}$	0,00012								
$\iota^{(3)}$	0,00006	0,00017							
$\iota^{(4)}$	0,00000	0,00000	0,00014						
$k_1^{(2)}$	-0,00001	0,00000	0,00000	0,00000					
$k_2^{(2)}$	-0,00329	-0,00029	0,00007	0,00045	0,18301				
$k_3^{(2)}$	-0,00011	-0,00001	0,00000	0,00002	0,00351	0,00023			
$k_4^{(2)}$	-0,00010	0,00000	0,00000	0,00001	0,00321	0,00011	0,00017		
$k_5^{(2)}$	-0,00012	-0,00002	0,00000	0,00002	0,00378	0,00013	0,00012	0,00019	
$k_1^{(4)}$	0,00007	-0,00002	-0,00501	-0,00003	-0,00257	-0,00008	-0,00008	-0,00008	0,56314
$k_2^{(4)}$	0,00008	0,00002	-0,00501	-0,00001	-0,00247	-0,00009	-0,00008	-0,00010	0,18299
$k_3^{(4)}$	0,00007	-0,00002	-0,00501	-0,00001	-0,00239	-0,00007	-0,00007	-0,00009	0,18320
$k_4^{(4)}$	0,00006	-0,00003	-0,00501	-0,00001	-0,00242	-0,00008	-0,00017	-0,00009	0,18292
$k_5^{(4)}$	0,00008	0,00002	-0,00501	-0,00001	-0,00262	-0,00008	-0,00007	-0,00009	0,18283
$k_6^{(4)}$	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	-0,00968
$k_7^{(4)}$	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
$k_8^{(4)}$	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
$k_9^{(4)}$	0,00000	0,00000	0,00000	0,00000	-0,00002	0,00000	0,00000	0,00000	0,00004

	$k_2^{(4)}$	$k_3^{(4)}$	$k_4^{(4)}$	$k_5^{(4)}$	$k_6^{(4)}$	$k_7^{(4)}$	$k_8^{(4)}$	$k_9^{(4)}$
$k_2^{(4)}$	0,33231							
$k_3^{(4)}$	0,18287	0,38391						
$k_4^{(4)}$	0,18290	0,18292	0,31986					
$k_5^{(4)}$	0,18306	0,18286	0,18305	0,25980				
$k_6^{(4)}$	0,00000	0,00000	0,00000	0,00000	0,00069			
$k_7^{(4)}$	0,00000	0,00320	0,00000	0,00000	0,00000	0,00061		
$k_8^{(4)}$	0,00000	0,00000	-0,00073	0,00000	0,00000	0,00000	0,00025	
$k_9^{(4)}$	0,00002	0,00001	-0,00001	0,00001	0,00000	0,00000	0,00000	0,00009

Tabelle 8.2: *Geschätzte Kovarianzen der fixen Parameter des endgültigen Modells.*

	Odenwald		Moränen		Schwarzwald		ZF-
	IWE	Mitscherlich	IWE	Mitscherlich	IWE	Mitscherlich	Daten
Anzahl Stichproben-punkte	43	68	57	55	78	132	51
			Anzahl Bäume pro Stichprobenpunkt				
Mittelwert	5.2	9.0	5.9	9.5	5.9	9.4	7.2
StdAbw.	1.6	1.4	1.4	1.0	1.4	1.2	4.1
Anzahl Bäume	224	613	338	522	461	1239	368
			Alter zum Fällzeitpunkt				
Mittelwert	72.7	66.8	77.2	58.7	86.7	71.9	68.4
StdAbw.	18.1	19.4	17.6	17.5	25.1	22.2	25.2
			Höhe zum Fällzeitpunkt				
Mittelwert	29.5	27.6	31.6	26.8	29.6	25.1	26.6
StdAbw.	4.1	5.5	4.5	4.1	4.6	6.2	4.4
			Anzahl Beobachtungen pro Baum				
Mittelwert	18.5	10.3	16.8	9.0	16.5	9.8	22.9
StdAbw.	6.1	2.7	6.5	1.6	6.7	1.8	2.4
			Altersspanne der Beobachtungen pro Baum				
Mittelwert	17.5	46.7	15.8	40.2	15.5	44.3	21.9
StdAbw.	6.1	13.4	6.5	8.0	6.7	8.9	2.4

Tabelle 8.3: *Beschreibung der Datengrundlage. Kennzahlen für die Modelldaten aus den Inventuren der IWE und aus Mitscherlichs Erhebung und für die Evaluierungsdaten aus dem Projekt „Zukunftsorientierte Forstwirtschaft“.*

Alters-	Odenwald		Moränen		Schwarzwald		ZF
klasse	IWE	Mitscherlich	IWE	Mitscherlich	IWE	Mitscherlich-	Daten
			Anzahl Beobachtungen nach Altersklassen				
10		214		341		412	
20		804	43	779	21	1087	83
30	119	1109	167	915	173	1765	953
40	563	1176	441	922	540	2080	1924
50	886	1037	864	762	802	2032	1965
60	923	737	1020	423	1079	1632	1266
70	726	484	1152	254	1359	1182	756
80	415	383	1032	169	1258	671	354
90	275	248	491	81	830	455	266
100	127	111	224	34	523	339	274
110	21	23	162	22	409	232	226
120		4	68	11	221	149	116
130	28	2	11	2	166	54	49
140	37				110	19	71
150	22				45	9	82
160	3				10	3	38
170					33		16
180					23		1
			mittlere Höhen nach Altersklassen				
10		4.4		4.6		3.4	
20		8.6	12.7	10.1	10.5	7.4	9.0
30	14.9	13.6	17.2	15.8	14.1	11.8	13.8
40	20.4	18.5	21.4	20.5	19.0	16.0	18.5
50	24.0	22.4	24.6	24.0	22.7	19.5	22.2
60	26.5	25.6	27.6	26.4	24.9	22.5	24.7
70	28.6	28.0	30.1	28.8	27.0	25.2	26.9
80	30.4	30.4	32.1	30.3	28.6	27.6	28.9
90	31.1	32.4	33.6	31.2	29.8	28.8	30.2
100	33.3	32.8	34.2	32.4	30.8	30.2	30.2
110	33.9	31.1	36.4	34.6	31.5	31.8	31.5
120		35.1	39.9	35.3	32.1	33.6	32.5
130	29.4	35.8	41.7	35.7	32.1	34.0	27.4
140	32.0				32.0	35.7	27.7
150	34.2				31.9	38.1	28.4
160	38.8				33.6	41.2	29.1
170					32.9		34.2
180					31.4		35.1
			Standardabweichung der mittleren Höhen nach Altersklassen				
10		1.74		2.13		1.62	
20		2.88	2.18	2.74	1.41	2.67	1.79
30	2.89	3.48	2.36	3.13	2.60	3.48	2.76
40	3.62	3.85	2.69	3.06	3.22	3.97	2.92
50	3.50	3.99	3.03	3.15	3.57	4.12	2.61
60	3.06	3.67	2.90	3.24	3.47	4.32	3.11
70	3.12	3.59	3.16	3.40	3.28	4.40	3.32
80	3.62	3.60	3.49	2.94	3.77	4.63	3.93
90	3.93	3.77	3.81	2.49	4.22	4.26	4.25
100	3.40	4.17	4.50	2.23	4.32	4.21	3.40
110	2.03	3.93	4.29	1.47	3.90	4.47	2.25
120		0.60	2.94	1.31	3.90	4.71	1.96
130	3.30	0.57	0.37	0.99	4.29	5.00	1.95
140	3.66				3.71	4.86	3.02
150	4.16				2.26	4.20	3.85
160	0.08				2.53	1.45	4.73
170					4.35		0.58
180					3.35		

Tabelle 8.4: *Beschreibung der Datengrundlage. Kennzahlen für die Modelldaten aus den Inventuren der IWE und aus Mitscherlichs Erhebung und für die Evaluierungsdaten aus dem Projekt „Zukunftsorientierte Forstwirtschaft“ nach Altersklassen.*

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N	368	368	368	368	368	368	368	368	367	367	366	366	366	364	364
	BLUPs für Zufallsparameter														
P99	1.56	1.96	2.31	2.53	2.67	2.69	2.75	2.84	3.21	3.15	3.29	3.66	3.87	3.98	4.07
P95	0.38	0.57	0.82	0.97	1.23	1.49	1.58	1.82	2.02	2.21	2.28	2.48	2.62	2.73	2.85
P90	0.29	0.42	0.56	0.77	0.90	1.07	1.24	1.37	1.47	1.61	1.74	1.78	1.88	1.93	1.90
Q3	0.13	0.22	0.31	0.42	0.53	0.57	0.66	0.76	0.80	0.85	0.89	0.91	0.93	0.93	1.01
Med	-0.01	0.02	0.04	0.08	0.08	0.11	0.12	0.13	0.17	0.13	0.14	0.17	0.16	0.10	0.08
Mean	0.01	0.05	0.08	0.12	0.15	0.17	0.19	0.22	0.23	0.24	0.22	0.22	0.20	0.16	0.12
Q1	-0.18	-0.21	-0.22	-0.27	-0.33	-0.35	-0.37	-0.39	-0.42	-0.45	-0.54	-0.57	-0.68	-0.74	-0.87
P10	-0.30	-0.40	-0.47	-0.54	-0.62	-0.73	-0.81	-0.97	-1.14	-1.19	-1.28	-1.43	-1.57	-1.66	-1.86
P5	-0.38	-0.49	-0.57	-0.74	-0.86	-0.97	-1.11	-1.25	-1.36	-1.45	-1.65	-1.80	-1.91	-2.12	-2.22
P1	-0.56	-0.76	-0.90	-1.07	-1.31	-1.55	-1.78	-2.00	-2.22	-2.34	-2.44	-2.51	-2.74	-2.98	-3.07
	Bonitätsfächer für alle IWE-Daten und alle Alters-Höhen-Klassen mit n≥1														
P99	0.43	0.75	1.11	1.41	1.81	2.16	2.3	2.63	2.99	3.2	3.24	3.6	3.8	4.01	4.29
P95	0.29	0.51	0.79	0.99	1.26	1.45	1.62	1.89	2.06	2.28	2.42	2.48	2.77	2.87	2.88
P90	0.22	0.4	0.57	0.83	0.99	1.16	1.27	1.42	1.58	1.67	1.8	1.94	1.93	2.07	2.21
Q3	0.12	0.24	0.33	0.45	0.56	0.64	0.72	0.81	0.88	0.92	1.01	1	1.1	1.1	1.1
Med	0.04	0.07	0.11	0.13	0.16	0.2	0.23	0.23	0.22	0.23	0.22	0.24	0.23	0.15	0.12
Mean	0.04	0.08	0.12	0.16	0.19	0.22	0.24	0.26	0.27	0.29	0.29	0.29	0.26	0.23	0.2
Q1	-0.05	-0.08	-0.12	-0.14	-0.2	-0.24	-0.28	-0.31	-0.35	-0.39	-0.44	-0.53	-0.51	-0.62	-0.72
P10	-0.13	-0.23	-0.31	-0.4	-0.51	-0.62	-0.74	-0.79	-0.85	-1.02	-1.16	-1.21	-1.38	-1.51	-1.66
P5	-0.21	-0.36	-0.47	-0.67	-0.77	-0.9	-1.07	-1.25	-1.38	-1.46	-1.57	-1.78	-1.92	-2.15	-2.31
P1	-0.34	-0.59	-0.84	-1.11	-1.36	-1.62	-1.86	-2.1	-2.33	-2.52	-2.61	-2.61	-2.76	-3.03	-3.24
	Bonitätsfächer für IWE-Daten aus Südschwarzwald und Alters-Höhen-Klassen mit n≥1														
P99	0.4	0.7	1.05	1.35	1.75	2.13	2.37	2.7	2.98	3.18	3.42	3.59	3.68	3.9	4.2
P95	0.29	0.51	0.75	0.96	1.19	1.37	1.61	1.87	2	2.19	2.35	2.48	2.69	2.67	2.74
P90	0.21	0.4	0.55	0.77	0.93	1.07	1.22	1.41	1.5	1.67	1.78	1.81	1.84	1.99	2.13
Q3	0.12	0.22	0.31	0.42	0.56	0.62	0.69	0.78	0.85	0.87	0.91	0.97	1.05	1.03	1.01
Med	0.03	0.06	0.12	0.15	0.16	0.21	0.24	0.25	0.22	0.24	0.24	0.26	0.2	0.19	0.09
Mean	0.04	0.07	0.11	0.14	0.18	0.2	0.22	0.25	0.25	0.26	0.26	0.26	0.23	0.21	0.16
Q1	-0.05	-0.06	-0.11	-0.15	-0.21	-0.25	-0.3	-0.34	-0.37	-0.39	-0.41	-0.5	-0.58	-0.62	-0.71
P10	-0.14	-0.24	-0.33	-0.41	-0.53	-0.61	-0.73	-0.82	-0.89	-0.96	-1.15	-1.24	-1.36	-1.48	-1.65
P5	-0.2	-0.35	-0.49	-0.66	-0.81	-0.94	-1.08	-1.25	-1.38	-1.5	-1.62	-1.75	-1.99	-2.15	-2.34
P1	-0.35	-0.66	-0.84	-1.11	-1.37	-1.62	-1.87	-2.11	-2.35	-2.6	-2.76	-2.8	-2.93	-2.98	-3.21
	Bonitätsfächer für alle IWE-Daten und Alters-Höhen-Klassen mit n≥5														
P99	0.4	0.69	1.05	1.32	1.71	2.09	2.21	2.62	2.87	3.08	3.26	3.46	3.59	3.81	4.08
P95	0.28	0.5	0.75	0.94	1.18	1.32	1.52	1.73	1.93	2.06	2.2	2.29	2.59	2.66	2.68
P90	0.21	0.38	0.54	0.79	0.91	1.06	1.17	1.34	1.46	1.59	1.63	1.73	1.74	1.85	2.02
Q3	0.12	0.21	0.3	0.4	0.52	0.56	0.64	0.71	0.8	0.84	0.86	0.9	0.96	0.93	0.95
Med	0.02	0.05	0.1	0.12	0.12	0.16	0.17	0.2	0.18	0.16	0.17	0.19	0.16	0.13	0.03
Mean	0.03	0.06	0.09	0.12	0.15	0.16	0.18	0.2	0.2	0.21	0.2	0.19	0.17	0.13	0.09
Q1	-0.06	-0.08	-0.13	-0.17	-0.23	-0.27	-0.34	-0.37	-0.42	-0.5	-0.54	-0.57	-0.59	-0.62	-0.74
P10	-0.14	-0.24	-0.34	-0.45	-0.56	-0.65	-0.77	-0.82	-0.92	-1.03	-1.19	-1.28	-1.44	-1.57	-1.77
P5	-0.22	-0.37	-0.5	-0.69	-0.81	-0.94	-1.11	-1.25	-1.43	-1.49	-1.66	-1.82	-2.04	-2.24	-2.41
P1	-0.35	-0.64	-0.86	-1.13	-1.39	-1.65	-1.9	-2.14	-2.39	-2.63	-2.72	-2.84	-2.86	-3.07	-3.3
	Bonitätsfächer für IWE-Daten aus Südschwarzwald und Alters-Höhen-Klassen mit n≥1														
P99	0.37	0.65	1.05	1.25	1.63	1.98	2.39	2.73	2.99	3.22	3.22	3.37	3.57	3.82	3.92
P95	0.28	0.48	0.71	0.89	1.12	1.27	1.42	1.63	1.82	2.02	2.13	2.27	2.36	2.44	2.48
P90	0.2	0.37	0.52	0.72	0.84	0.97	1.1	1.29	1.37	1.47	1.63	1.66	1.71	1.75	1.9
Q3	0.11	0.2	0.27	0.38	0.49	0.53	0.59	0.69	0.78	0.8	0.86	0.85	0.88	0.87	0.87
Med	0.02	0.04	0.08	0.13	0.12	0.17	0.19	0.19	0.16	0.18	0.18	0.18	0.13	0.11	0
Mean	0.02	0.05	0.08	0.1	0.13	0.14	0.15	0.17	0.17	0.18	0.18	0.17	0.14	0.11	0.06
Q1	-0.06	-0.09	-0.13	-0.19	-0.25	-0.29	-0.37	-0.4	-0.43	-0.47	-0.53	-0.62	-0.66	-0.71	-0.81
P10	-0.15	-0.27	-0.35	-0.45	-0.57	-0.64	-0.73	-0.85	-0.99	-1.05	-1.22	-1.32	-1.44	-1.58	-1.72
P5	-0.21	-0.36	-0.5	-0.69	-0.83	-0.96	-1.09	-1.26	-1.39	-1.53	-1.65	-1.83	-2.02	-2.17	-2.34
P1	-0.35	-0.69	-0.86	-1.15	-1.39	-1.64	-1.89	-2.13	-2.37	-2.64	-2.79	-2.84	-2.96	-2.99	-3.22

Tabelle 8.5: *Verteilungen der Residuen über die Prognosezeiträume bei Kenntnis der ersten Beobachtung - Teil 1.*

t	16	17	18	19	20	21	22	23	24	25	26	27
N	364	361	358	351	323	277	210	157	101	50	13	2
	BLUPs für Zufallsparameter											
P99	4.42	4.38	4.52	4.72	4.98	6.08	6.17	7.61	6.55	4.73	3.27	-0.11
P95	2.93	3.11	3.13	3.15	2.71	2.82	3.20	4.21	4.06	3.56	3.27	-0.11
P90	2.04	2.11	2.19	2.27	1.90	1.85	1.79	2.34	2.56	3.09	1.54	-0.11
Q3	1.05	1.06	1.05	1.02	1.06	1.04	0.96	0.85	0.75	0.90	0.57	-0.11
Med	0.03	0.06	0.03	0.03	-0.01	-0.04	-0.28	-0.19	-0.19	-0.37	0.40	-0.40
Mean	0.11	0.07	0.07	0.03	-0.01	-0.09	-0.20	-0.11	-0.13	-0.19	0.08	-0.40
Q1	-0.97	-1.07	-1.08	-1.12	-1.12	-1.24	-1.50	-1.44	-1.37	-1.30	-0.75	-0.68
P10	-1.95	-1.98	-2.10	-2.19	-2.41	-2.57	-2.74	-2.72	-2.71	-2.84	-1.07	-0.68
P5	-2.33	-2.56	-2.72	-2.79	-2.88	-3.07	-3.15	-3.31	-3.14	-3.26	-3.67	-0.68
P1	-3.31	-3.43	-3.39	-3.54	-3.71	-3.99	-3.95	-4.42	-3.74	-3.57	-3.67	-0.68
	Bonitätsfächer für alle IWE-Daten und alle Alters-Höhen-Klassen mit n≥1											
P99	4.47	4.66	4.63	4.67	4.19	4.28	3.23	3.52	3.06	2.39	1.19	-0.44
P95	3.08	3.2	3.43	3.25	2.83	2.75	2.41	2.62	1.92	1.53	1.19	-0.44
P90	2.25	2.32	2.41	2.45	2.35	2.19	1.87	1.86	1.13	1.14	0.95	-0.44
Q3	1.1	1.17	1.17	1.14	1.04	0.91	0.63	0.68	0.57	0.46	0.5	-0.44
Med	0.05	0.07	0.02	0.04	0.04	-0.09	-0.4	-0.34	-0.26	-0.33	-0.25	-1.04
Mean	0.18	0.15	0.14	0.11	0.02	-0.11	-0.36	-0.35	-0.45	-0.58	-0.42	-1.04
Q1	-0.8	-0.8	-0.87	-0.94	-0.98	-1.14	-1.43	-1.28	-1.28	-1.3	-1	-1.64
P10	-1.77	-1.86	-2.01	-2.14	-2.14	-2.45	-2.65	-2.67	-2.78	-2.9	-1.95	-1.64
P5	-2.42	-2.44	-2.54	-2.68	-2.85	-2.96	-3.17	-3.34	-3.29	-3.36	-4.04	-1.64
P1	-3.35	-3.51	-3.63	-3.84	-3.83	-4.05	-4.02	-4.14	-4.18	-3.93	-4.04	-1.64
	Bonitätsfächer für IWE-Daten aus Südschwarzwald und Alters-Höhen-Klassen mit n≥1											
P99	4.36	4.51	4.65	4.54	4.28	4.45	3.62	3.87	3.75	2.89	1.05	-0.16
P95	2.96	2.95	3.11	3.1	2.89	2.87	2.4	2.53	1.91	1.53	1.05	-0.16
P90	2.19	2.23	2.32	2.38	2.16	2.1	1.76	1.83	1.03	1.08	0.54	-0.16
Q3	1.05	1.04	1.09	1.04	0.92	0.86	0.59	0.54	0.52	0.32	0.49	-0.16
Med	0.05	0.05	0	-0.02	-0.03	-0.14	-0.35	-0.44	-0.32	-0.49	-0.09	-0.59
Mean	0.14	0.11	0.1	0.06	-0.01	-0.15	-0.4	-0.4	-0.5	-0.66	-0.47	-0.59
Q1	-0.84	-0.84	-0.95	-1.03	-1.06	-1.17	-1.53	-1.45	-1.32	-1.5	-1.35	-1.03
P10	-1.74	-1.9	-2.06	-2.06	-2.25	-2.46	-2.67	-2.84	-2.79	-2.97	-1.67	-1.03
P5	-2.46	-2.46	-2.59	-2.77	-2.87	-2.96	-3.15	-3.18	-3.05	-3.11	-4.19	-1.03
P1	-3.43	-3.64	-3.7	-3.87	-3.87	-4.14	-4.18	-4.31	-4.51	-4.07	-4.19	-1.03
	Bonitätsfächer für alle IWE-Daten und Alters-Höhen-Klassen mit n≥5											
P99	4.33	4.44	4.4	4.43	4.37	4.5	3.37	3.4	3.12	2.42	1.1	-0.32
P95	2.88	2.83	3.14	2.94	2.73	2.6	2.22	2.42	1.72	1.33	1.1	-0.32
P90	2.08	2.13	2.24	2.19	2.04	2.02	1.64	1.67	1.1	1.05	0.75	-0.32
Q3	0.99	0.98	1.03	0.96	0.89	0.73	0.45	0.51	0.45	0.32	0.46	-0.32
Med	-0.02	-0.06	-0.09	-0.1	-0.03	-0.17	-0.45	-0.47	-0.37	-0.51	-0.24	-0.9
Mean	0.07	0.03	0.02	-0.02	-0.1	-0.24	-0.48	-0.47	-0.55	-0.67	-0.49	-0.9
Q1	-0.87	-0.94	-1.04	-1.09	-1.12	-1.25	-1.55	-1.49	-1.37	-1.29	-1.35	-1.49
P10	-1.83	-1.92	-2.1	-2.22	-2.37	-2.56	-2.74	-2.86	-2.79	-3.05	-1.8	-1.49
P5	-2.5	-2.59	-2.67	-2.76	-2.92	-3.12	-3.21	-3.3	-3.32	-3.39	-4.18	-1.49
P1	-3.52	-3.74	-3.75	-3.95	-3.96	-4.16	-4.25	-4.28	-4.36	-4.06	-4.18	-1.49
	Bonitätsfächer für IWE-Daten aus Südschwarzwald und Alters-Höhen-Klassen mit n≥1											
P99	4.11	4.24	4.28	4.42	3.88	4.04	3.45	3.9	3.57	2.72	1.15	0.02
P95	2.7	2.82	3.02	2.86	2.57	2.58	2.23	2.36	1.64	1.33	1.15	0.02
P90	1.97	1.97	2.06	2.17	1.99	1.87	1.53	1.64	1.04	1.14	0.66	0.02
Q3	0.93	0.94	0.94	0.91	0.91	0.76	0.56	0.56	0.54	0.49	0.64	0.02
Med	-0.08	-0.04	-0.04	-0.11	-0.05	-0.17	-0.42	-0.48	-0.26	-0.47	0.08	-0.5
Mean	0.04	0	-0	-0.04	-0.12	-0.25	-0.48	-0.46	-0.52	-0.63	-0.4	-0.5
Q1	-0.92	-1.01	-1.06	-1.07	-1.13	-1.34	-1.54	-1.5	-1.46	-1.51	-1.35	-1.02
P10	-1.82	-1.96	-2.09	-2.2	-2.37	-2.56	-2.74	-2.82	-2.85	-2.96	-1.67	-1.02
P5	-2.47	-2.51	-2.72	-2.83	-2.93	-3.25	-3.27	-3.15	-3.08	-3.16	-4.17	-1.02
P1	-3.49	-3.65	-3.75	-3.94	-4.06	-4.36	-4.24	-4.26	-4.45	-4.05	-4.17	-1.02

Tabelle 8.6: *Verteilungen der Residuen über die Prognosezeiträume bei Kenntnis der ersten Beobachtung - Teil 2.*

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
N	366	366	366	364	364	364	361	358	351	323	277	210	157	101	50	13
	BLUPs für Zufallsparameter															
P99	1.74	1.62	1.55	1.59	1.58	1.82	2.18	2.47	2.93	2.99	3.23	3.71	5.67	4.47	3.38	0.95
P95	0.60	0.58	0.61	0.66	0.76	0.85	0.97	1.06	1.16	1.25	1.26	1.63	1.67	1.76	0.99	0.95
P90	0.42	0.45	0.48	0.48	0.50	0.59	0.59	0.64	0.66	0.69	0.82	0.90	1.00	0.86	0.87	0.69
Q3	0.19	0.23	0.23	0.21	0.21	0.22	0.21	0.23	0.22	0.21	0.25	0.26	0.38	0.43	0.24	0.28
Med	0.01	-0.01	-0.04	-0.10	-0.11	-0.18	-0.22	-0.23	-0.22	-0.21	-0.18	-0.17	-0.16	-0.17	-0.31	-0.05
Mean	0.05	0.04	0.00	-0.04	-0.09	-0.12	-0.15	-0.16	-0.18	-0.17	-0.16	-0.15	-0.08	-0.04	-0.18	0.02
Q1	-0.16	-0.20	-0.27	-0.36	-0.47	-0.53	-0.59	-0.64	-0.71	-0.71	-0.72	-0.81	-0.78	-0.81	-0.78	-0.34
P10	-0.32	-0.41	-0.50	-0.58	-0.71	-0.83	-0.91	-1.00	-1.11	-1.16	-1.27	-1.31	-1.23	-1.28	-1.20	-0.61
P5	-0.46	-0.51	-0.64	-0.70	-0.87	-1.05	-1.14	-1.27	-1.34	-1.46	-1.57	-1.57	-1.65	-1.38	-1.39	-0.73
P1	-0.71	-0.81	-1.04	-1.21	-1.54	-1.73	-1.89	-1.93	-2.22	-2.04	-2.36	-2.30	-2.81	-2.11	-1.44	-0.73
	Bonitätsfächer für alle IWE-Daten und alle Alters-Höhen-Klassen mit n≥1															
P99	0.36	0.65	0.88	1.11	1.25	1.41	1.73	1.88	2.03	1.93	2.19	2.25	2.35	2.31	1.91	0.89
P95	0.30	0.45	0.63	0.77	0.88	1.10	1.20	1.32	1.48	1.55	1.53	1.47	1.60	1.47	1.16	0.89
P90	0.21	0.37	0.46	0.60	0.66	0.79	0.82	0.97	1.05	1.13	1.26	1.09	1.31	0.80	0.81	0.44
Q3	0.12	0.20	0.22	0.25	0.29	0.30	0.35	0.42	0.41	0.38	0.37	0.25	0.34	0.28	0.19	0.15
Med	0.00	0.00	-0.04	-0.08	-0.12	-0.15	-0.16	-0.16	-0.17	-0.17	-0.18	-0.37	-0.23	-0.27	-0.42	-0.23
Mean	0.01	0.00	-0.02	-0.05	-0.09	-0.11	-0.12	-0.13	-0.15	-0.16	-0.22	-0.33	-0.30	-0.33	-0.48	-0.34
Q1	-0.10	-0.18	-0.24	-0.32	-0.40	-0.50	-0.55	-0.63	-0.73	-0.73	-0.81	-1.03	-0.98	-0.91	-0.99	-0.77
P10	-0.20	-0.35	-0.50	-0.64	-0.81	-0.98	-1.12	-1.24	-1.35	-1.49	-1.62	-1.72	-1.72	-1.91	-2.01	-1.41
P5	-0.25	-0.42	-0.65	-0.85	-1.01	-1.21	-1.38	-1.43	-1.59	-1.77	-1.97	-2.08	-2.18	-2.04	-2.15	-1.87
P1	-0.32	-0.63	-0.93	-1.23	-1.48	-1.60	-1.79	-1.96	-2.18	-2.20	-2.48	-2.45	-3.04	-2.59	-2.54	-1.87
	Bonitätsfächer für IWE-Daten aus Südschwarzwald und Alters-Höhen-Klassen mit n≥1															
P99	0.35	0.64	0.89	1.05	1.26	1.41	1.64	1.81	1.98	1.96	2.12	2.28	2.44	2.25	2.16	0.74
P95	0.28	0.43	0.59	0.75	0.86	1.09	1.14	1.32	1.41	1.51	1.63	1.57	1.79	1.69	1.21	0.74
P90	0.21	0.36	0.43	0.57	0.62	0.75	0.80	0.92	0.99	1.05	1.15	1.03	1.25	0.85	0.51	0.42
Q3	0.11	0.18	0.23	0.22	0.24	0.28	0.32	0.39	0.39	0.37	0.36	0.20	0.25	0.23	0.00	0.07
Med	0.00	-0.02	-0.04	-0.08	-0.13	-0.17	-0.18	-0.19	-0.24	-0.22	-0.23	-0.43	-0.38	-0.43	-0.49	-0.32
Mean	0.00	0.00	-0.03	-0.06	-0.11	-0.13	-0.15	-0.16	-0.18	-0.19	-0.25	-0.36	-0.34	-0.37	-0.55	-0.40
Q1	-0.10	-0.18	-0.25	-0.32	-0.41	-0.53	-0.58	-0.67	-0.74	-0.79	-0.90	-1.03	-0.98	-0.85	-1.08	-0.65
P10	-0.21	-0.33	-0.54	-0.65	-0.78	-0.96	-1.07	-1.21	-1.36	-1.47	-1.60	-1.76	-1.78	-1.83	-1.90	-1.72
P5	-0.25	-0.44	-0.66	-0.86	-1.06	-1.23	-1.37	-1.44	-1.56	-1.79	-1.96	-2.06	-2.29	-2.08	-2.12	-1.81
P1	-0.31	-0.64	-0.94	-1.22	-1.49	-1.63	-1.89	-2.07	-2.16	-2.24	-2.42	-2.56	-3.04	-2.69	-2.59	-1.81
	Bonitätsfächer für alle IWE-Daten und Alters-Höhen-Klassen mit n≥5															
P99	0.34	0.62	0.87	1.06	1.21	1.37	1.62	1.74	1.86	1.87	2.06	2.18	2.28	2.22	1.94	0.82
P95	0.28	0.42	0.58	0.73	0.87	1.03	1.11	1.32	1.40	1.43	1.57	1.42	1.59	1.40	1.16	0.82
P90	0.21	0.35	0.43	0.55	0.61	0.76	0.76	0.88	0.98	1.06	1.14	1.00	1.23	0.74	0.70	0.44
Q3	0.11	0.18	0.21	0.21	0.25	0.26	0.29	0.33	0.37	0.37	0.29	0.18	0.29	0.21	0.12	0.20
Med	0.00	-0.01	-0.05	-0.10	-0.13	-0.19	-0.18	-0.20	-0.23	-0.22	-0.25	-0.43	-0.34	-0.39	-0.50	-0.26
Mean	0.00	-0.01	-0.04	-0.07	-0.12	-0.14	-0.16	-0.17	-0.20	-0.21	-0.27	-0.39	-0.35	-0.38	-0.54	-0.38
Q1	-0.11	-0.18	-0.26	-0.34	-0.42	-0.55	-0.60	-0.65	-0.73	-0.80	-0.88	-1.02	-0.99	-0.88	-1.07	-0.71
P10	-0.21	-0.35	-0.54	-0.66	-0.82	-0.99	-1.14	-1.27	-1.39	-1.49	-1.65	-1.77	-1.71	-1.91	-2.01	-1.56
P5	-0.26	-0.44	-0.65	-0.84	-1.03	-1.25	-1.41	-1.47	-1.62	-1.81	-2.02	-2.10	-2.25	-2.09	-2.20	-1.89
P1	-0.33	-0.66	-0.97	-1.26	-1.54	-1.66	-1.87	-2.05	-2.19	-2.21	-2.51	-2.51	-3.06	-2.66	-2.54	-1.89
	Bonitätsfächer für IWE-Daten aus Südschwarzwald und Alters-Höhen-Klassen mit n≥1															
P99	0.34	0.61	0.87	1.04	1.21	1.40	1.62	1.73	1.90	1.90	2.06	2.21	2.40	2.23	2.11	0.83
P95	0.26	0.41	0.56	0.70	0.82	1.02	1.08	1.25	1.33	1.41	1.58	1.49	1.75	1.73	1.16	0.83
P90	0.21	0.34	0.41	0.54	0.59	0.73	0.75	0.88	0.96	1.00	1.07	0.99	1.24	0.75	0.61	0.55
Q3	0.10	0.18	0.21	0.20	0.24	0.26	0.25	0.32	0.35	0.32	0.32	0.17	0.27	0.25	0.10	0.17
Med	0.00	-0.02	-0.04	-0.09	-0.15	-0.16	-0.17	-0.19	-0.24	-0.22	-0.22	-0.38	-0.37	-0.39	-0.44	-0.20
Mean	0.00	-0.01	-0.04	-0.07	-0.12	-0.14	-0.16	-0.17	-0.19	-0.20	-0.26	-0.37	-0.33	-0.35	-0.50	-0.33
Q1	-0.10	-0.18	-0.25	-0.34	-0.42	-0.53	-0.61	-0.69	-0.72	-0.76	-0.88	-1.02	-0.96	-0.85	-0.96	-0.60
P10	-0.21	-0.34	-0.54	-0.65	-0.79	-0.99	-1.09	-1.24	-1.39	-1.49	-1.59	-1.77	-1.79	-1.82	-1.87	-1.66
P5	-0.26	-0.44	-0.66	-0.86	-1.05	-1.24	-1.38	-1.46	-1.63	-1.78	-1.98	-2.02	-2.30	-2.10	-2.15	-1.79
P1	-0.33	-0.67	-0.98	-1.24	-1.55	-1.69	-1.92	-2.11	-2.17	-2.27	-2.45	-2.57	-3.00	-2.64	-2.54	-1.79

Tabelle 8.7: *Verteilungen der Residuen über die Prognosezeiträume bei Kenntnis der elften Beobachtung für die Fortschreibung mittels der Bonitätsfächer und der ersten und elften Beobachtungen für BLUPs der Zufallsparameter.*

Literaturverzeichnis

S. L. Beal und L. B. Sheiner. Estimating population kinetics. *CRC Crit. Rev. Biomed. Eng.*, 8:195–222, 1982. 30

S. L. Beal und L. B. Sheiner. Heteroskedastic nonlinear regression. *Technometrics*, 30: 327–338, 1988. 30

J. F. Beaumont, C. H. Ung, und M. Bernier-Cardou. Relating site index to ecological factors in black spruce stands: Tests of hypotheses. *Forest Science*, 45(4):484–491, 1999. 122, 123, 124

R. Calama und G. Montero. Interregional nonlinear height-diameter model with random coefficients for stone pine in spain. *Canadian Journal of Forest Research*, 34(1):150–163, 2004. 8

N. Calegario, R. F. Daniels, R. Maestri, und R. Neiva. Modeling dominant height growth based on nonlinear mixed-effects model: a clonal eucalyptus plantation case study. *Forest Ecology and Management*, 204(1):11–20, 2005. 9

C. J. Cieszewski, M. Zasada, und M. Strub. Analysis of different base models and methods of site model derivation for scots pine. *Forest Science*, 52(2):187–197, 2006. 117, 118

M. Davidian. *ST 762 - Nonlinear Models for Univariate and Multivariate Response*. Department of Statistics North Carolina State University, 2001. Lecture Notes. 21, 26, 27, 29

M. Davidian. *ST 732 - Applied Longitudinal Data Analysis*. Department of Statistics North Carolina State University, 2005. Lecture Notes. 5

M. Davidian und D. M. Giltinan. *Nonlinear Models for Repeated Measurement Data*. Chapman and Hall, London, 1995. 21, 26, 27, 28, 29, 36, 146

M. Davidian und D. M. Giltinan. Nonlinear models for repeated measurement data: An overview and update. *Journal of Agriculture, Biological and Environmental Statistics*, 8(4):387–419, 2003. 21

Z. Fang, R. L. Bailey, und B. D. Shiver. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. *Forest Science*, 47(4):550–562, 2001. 8, 131

Z. X. Fang und R. L. Bailey. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. *Forest Science*, 47(3):287–300, 2001. 8, 18, 37

O. Garcia. A stochastic differential equation model for the height growth of forest stands. *Biometrics*, 39:1059–1072, 1983. 117, 119, 120

T. G. Gregoire und O. Schabenberger. A non-linear mixed-effects model to predict cumulative bole volume of standing trees. *Journal of Applied Statistics*, 23(2-3):257–272, 1996a. 7

T. G. Gregoire und O. Schabenberger. Non-linear mixed-effects modeling of cumulative bole volume with spatial correlated within-tree data. *Journal of Agricultural, Biological, and Environmental Statistics*, 1:107–119, 1996b. 7

D. B. Hall und R. L. Bailey. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *Forest Science*, 47(3):311–321, 2001. 7, 60

D. B. Hall und M. Clutter. Multivariate multilevel nonlinear mixed effects models for timber yield predictions. *Biometrics*, 60(1):16–24, 2004. 8, 59

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982. 122

W. Härdle und L. Simar. *Applied Multivariate Statistical Analysis*. Springer-Verlag, New York, 2003. 31

H. Hasenauer. Höhenzuwachsmodelle für die wichtigsten Baumarten Österreichs. *Forstwissenschaftliches Centralblatt*, 118:14–23, 1999. 124, 125

J. Hradetzky. Zuwachsermittlung bei Waldinventuren. -, -(-):–, 2002. unveröffentlicht. 3

G. Kändler und U. Riemer. *Der Zusammenhang zwischen Wuchsleistung und Standort bei Fichte, Tanne und Buche auf den wichtigsten Standortseinheiten des Südschwarzwaldes*. Springer-Verlag, Berlin, 2005. 59-84; in: Waldumbau für eine zukunftsorientierte Waldwirtschaft. K. v. Teuffel, M. Baumgarten, M. Hanewinkel, W. Konold, U. H. Sauter, H. Spiecker und K. v. Wilpert. 15

E. Kublin. Einheitliche Beschreibung der Schaftform - Methoden und Programme - BDATPro. *Forstwissenschaftliches Centralblatt*, 122(3):183–200, 2003. 2, 143

Landesforstverwaltung Baden-Württemberg. *Verwaltungsvorschrift des Ministeriums für Ernährung und Ländlichen Raum über eine Dienstanweisung für die Forsteinrichtung im öffentlichen Wald Baden-Württembergs (FED 2000); 55-8632.00*. Ministerium Ländlicher Raum Baden-Württemberg, Stuttgart, 2002. 3

Landesforstverwaltung Baden-Württemberg. *Hilfstabellen für die Forsteinrichtung*. Ministerium für Ländlichen Raum, Ernährung, Landwirtschaft und Forsten Baden-Württemberg, Stuttgart, 1993. 71

Landesforstverwaltung Baden-Württemberg. *Richtlinie Landesweiter Waldentwicklungstypen*. Ministerium Ländlicher Raum Baden-Württemberg, Stuttgart, 1999. 129

J. Lappi. Calibration of height and volume equations with random parameters. *Forest Science*, 37(3):781–801, 1991. 7, 82

J. Lappi. A longitudinal analysis of height/diameter curves. *Forest Science*, 43(4):555–570, 1997. 7, 79

J. Lappi und R. L. Bailey. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. *Forest Science*, 34(4):907–927, 1988. 7, 35, 83, 126, 128, 129

J. Lappi und J. Malinen. Random-parameter height/age models when stand parameters and stand age are correlated. *Forest Science*, 40(4):715–731, 1994. 7

M. J. Lindstrom und D. M. Bates. Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46:673–687, 1990. 7, 21, 33

R. McGill, J. W. Tukey, und W. A. Larsen. Variations of box plots. *The American Statistician*, 32:12–16, 1978. 57

L. Mehtatalo. A longitudinal height-diameter model for norway spruce in finland. *Canadian Journal of Forest Research*, 34(1):131–140, 2004. 7

L. Mehtatalo. Height-diameter models for scots pine and birch in finland. *Silva Fennica*, 39(1):55–66, 2005. 7

Ministerkonferenz zum Schutz der Wälder in Europa. Resolution H1. *Zweite Ministerkonferenz zum Schutz der Wälder in Europa*, Helsinki/Finland, 1993. 1

G. Mitscherlich. Das Wachstum der Fichte in Baden. I. Das Höhenwachstum und die Bonitätsveränderung. II. Grundfläche, Stammzahl, Durchmesser und Massenzuwachs. III. Über den Aufbau reiner, gleichaltriger Fichtenbestände. *Allgemeine Forst- und Jagdzeitung*, 128(8/9; 10/11; 12):171–184; 219–232; 245–256, 1957. 9, 70, 71, 132, 133, 134, 139, 144

R. A. Monserud. Height growth and site index curves for inland douglas-fir based on stem analysis data and forest habitat type. *Biometrics*, 30(4):943–965, 1984. 120, 121, 134

R. A. Monserud und G. E. Rehfeld. Genetic and environmental components of variation of site index in inland douglas-fir. *Forest Science*, 36(1):1–9, 1990. 116

H. U. Moosmayer, W. Schöpfer, E. Kublin, und U. Tuch. *Standörtliche Einflüsse auf Wachstumsänderungen der Fichte in Baden Württemberg*. Verein für Forstliche Standortskunde und Forstpflanzenzüchtung e.V., 2003. 11, 135

J. Nagel. *Konzeptionelle Überlegungen zum schrittweisen Aufbau eines waldwachstumskundlichen Simulationssystems für Nordwestdeutschland*. Schriften aus der Forstlichen Fakultät der Universität Göttingen und der Nieders. Forstl. Versuchsanstalt, Band 128, 1999. 131

J. Nagel, M. Albert, und M. Schmidt. Das waldbauliche Prognose- und Entscheidungsmodell BWINPro. *Forst u. Holz*, 57(15/16):486–493, 2002. 3, 131

T. Nord-Larsen. Developing dynamic site index curves for european beech (Fagus sylvatica L.) in Denmark. *Forest Science*, 52(2):173–181, 2006. 117, 119

J. C. Pinheiro und D. M. Bates. *Mixed-effects models in S and S-plus*. Springer-Verlag, New York, 2002. 17, 21, 26, 30, 33, 34, 41, 43, 56, 59, 60

H. Pretzsch. *Modellierung des Waldwachstums*. Blackwell Wissenschafts-Verlag, Berlin, 2001. 1, 7, 115, 129, 130

H. Pretzsch, P. Biber, und J. Ďurský. The single tree-based stand simulator silva: construction, application and evaluation. *Forest Ecology and Management*, 162(1):3–21, 2002. 3

David A. Ratkowsky. *Handbook of nonlinear regression models*. Marcel Dekker, New York, 1989. 17, 18, 19, 78

G. E. Rehfeldt. Ecological adaptations in douglas-fir (pseudotsuga meziesii var. glauca): A synthesis. *Forest Ecology and Management*, 28:203–215, 1989. 116

K. Rennolls. Forest height growth modelling. *Forest Ecology and Management*, 71:217–225, 1995. 116, 122

SAS Institute Inc. *SAS OnlineDoc© 9.1*. SAS Institute Inc., Cary, NC, 2003. 32

O. Schabenberger und F. J. Pierce. *Contemporary Statistical Models for the Plant and Soil Sciences*. CRC Press, New York, 2002. 4, 6, 17, 20, 21, 26, 28, 31, 33, 34, 43

K. Schadauer. *Growth Trends in Austria*. Springer-Verlag, New York, 1996. 275-289; in: Growth trends in European forests. Spiecker, H. and Mielikäinen, K. and Köhl, M. and Skovsgaard, J.P. 136

W. Schöpfer und J. Hradetzky. Zielsetzungen, Methoden und Probleme der terrestrischen Waldschadensinventur Baden-Württemberg 1983. *Mitteilungen der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg*, 107, 1983. 13, 70

W. Schöpfer, J. Hradetzky, und E. Kublin. Wachstumsänderungen der Fichte in Baden Württemberg. *Forst und Holz*, 49(21):633–644, 1994. 70, 71, 134

W. Schöpfer. Immissionsökologische Waldzustandserhebungen. *Forst und Holz*, 49(21): 618, 1994. 13

B. Sloboda. Zur Darstellung von Wachstumsprozessen mit Hilfe von Differentialgleichungen erster Ordnung. *Mitteilungen der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg*, 32, 1971. 19, 98, 117, 121, 125, 131, 139, 144, 146

D. Sorensen. *Likelihood, Bayesian and MCMC methods in quantitative genetics*. Springer-Verlag, New York, 2002. 18, 27

H. Spiecker, K. Mielikäinen, M. Köhl, und J.P. Skovsgaard. *Growth trends in European forests*. Springer-Verlag, New York, 1996. 62

H. Sterba, M. Moder, H. Hasenauer, und R. A. Monserud. *PROGNAUS - Ein abstandsunabhängiger Wachstumssimulator für ungleichaltrige Mischbestände*. Deutscher Verband Forstl. Forschungsanstalten - Sektion Ertragskunde. Eberswalde/Berlin, 1995. 124

M. Strub und C.J. Cieszewski. Base-age invariance properties of two techniques for estimating the parameters of site index models. *Forest Science*, 52(2):182–186, 2006. 119

H. Untheim. Zur Veränderung der Produktivität von Waldstandorten. Untersuchungen zum Höhen und Volumenwachstum von Fichte (Picea abies (L.) Karst.) und Buche (Fagus sylvatica L.) auf Standorteinheiten der Ostalb und des Flächenschwarzwaldes. *Mitteilungen der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg*, 198, 1996. 11, 63, 135

E. F. Vonesh und V. M. Chinchilli. *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. Marcel Dekker, New York, 1996. 36, 37, 146

G. Wenk und M. Vogel. *Height Growth Investigations of Norway Spruce (Picea abies, [L.] Karst.) in the Eastern Part of Germany During the Last Century*. Springer-Verlag, New York, 1996. 99-106; in: Growth trends in European forests, Spiecker, H. and Mielikäinen, K. and Köhl, M. and Skovsgaard, J.P. 135, 136

G. Wenk, V. Antanaitis, und Š. Šmelko. *Waldertragslehre*. Dt. Landwirtschaftsverl., Berlin, 1990. 115, 116, 117, 129

J. Wolf. Zeitberücksichtigendes und zeitloses Richtungsfeld der Entwicklung. *Acta Uni. Agric. et Silvic., Brno Čisto 1*, 32:12–16, 1957. 19, 98

D. H. Zhao, M. Wilson, und B. E. Borders. Modeling response curves and testing treatment effects in repeated measures experiments: a multilevel nonlinear mixed-effects model approach. *Canadian Journal of Forest Research*, 35(1):122–132, 2005. 9

H. Zimmerle. Beiträge zur Biologie der Fichte in Württemberg. *Mitteilungen der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg*, 8, 1949. 132

Lebenslauf

Arne Nothdurft

Im Etter 6 79117 Freiburg Tel.: 0761-6008436	Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg Wonnhaldestr. 4, 79100 Freiburg Tel.: 0761-4018-195 email: arne.nothdurft@forst.bwl.de

Angaben zur Person

Geburtsdatum und -ort	15.02.1974 in Nürnberg.
Nationalität	Deutsch.
Familienstand	Ledig.

Berufliche Tätigkeiten

Seit April 2002	Wissenschaftlicher Referent der Abteilung für Biometrie und Informatik, FVA Freiburg.
Juli 2001 - April 2002	Forsteinrichter an der Forstdirektion Tübingen.

Qualifikationen

Mai 2001	Niedersächsische Landesforstverwaltung, Große Forstliche Staatsprüfung für den höheren Forstdienst.
Mai 1999	Universität Göttingen, Fakultät für Forstwissenschaften und Waldökologie. Abschluss als Diplomforstwirt.

Ausbildung

Juni 1999 - Mai 2001	Niedersächsische Landesforstverwaltung, Referendariat.
Okt. 1994 - Mai 1999	Universität Göttingen, Studium an der Fakultät für Forstwissenschaften und Waldökologie.
Mai 1993	Robert-Koch-Gymnasium Clausthal-Zellerfeld, Abitur.