

Technische Universität München
Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt
Department Ökosystem und Landschaftsmanagement – Lehrstuhl für Waldwachstumskunde

**Standortabhängige Ableitung der Höhenwuchsleistung aus
Forstinventurdaten mit Hilfe von Data-Mining-Methoden.
Grundlage für die regionale, standortbezogene Feinjustierung des
forstlichen Wachstumsmodells SILVA**

Hans-Joachim Klemmt

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für
Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung
des akademischen Grades

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender des Prüfungsausschusses
Univ.-Prof. Dr. R. Mosandl

Prüfer der Dissertation

1. Univ.-Prof. Dr. H. Pretzsch
2. Univ.-Prof. Dr. W. Warkotsch
3. Univ.-Prof. Dr. Dr. A. Göttlein

Die Dissertation wurde am 05.04.2007 bei der Technischen Universität München eingereicht
und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung
und Umwelt am 30.07.2007 angenommen.

Inhaltsverzeichnis

0.	VORWORT	1
1.	EINLEITUNG	3
1.1	AUSGANGSSITUATION.....	3
1.2	EINSATZ VON WUCHSMODELLEN IN DER PRAXIS ZUR UNTERSTÜTZUNG FORSTLICHER ENTSCHEIDUNGEN.....	5
1.3	DATENGRUNDLAGEN ZUR PARAMETRIERUNG UND JUSTIERUNG VON EINZELBAUM- BASIERTEN WUCHSMODELLEN.....	8
1.4	BEGRIFFSABGRENZUNGEN	11
1.5	ZIELE DER ARBEIT	12
2.	STAND DES WISSENS.....	16
2.1	KENNTNISSTAND ZUR ABHÄNGIGKEIT DES WALDWACHSTUMS VON STANDÖRTLICHEN FAKTOREN	16
2.1.1	Kenntnisstand zum Einfluss primärer Grundfaktoren auf das Waldwachstum	16
2.1.2	Kenntnisstand zum Einfluss des Standortes auf das Waldwachstum	19
2.2	FUNKTIONSPRINZIP DES WUCHSMODELLS SILVA.....	24
2.3	DATA-MINING	31
2.3.1	Überblick über Data-Mining-Verfahren und deren Anwendung	31
2.3.2	Anwendung von Data-Mining-Verfahren in der Forstwirtschaft.....	36
2.4	KONSEQUENZEN AUS DEM DARGESTELLTEN STAND DES WISSENS FÜR DIE VORLIEGENDE ARBEIT.....	39
3.	MATERIAL UND METHODEN	41
3.1	DATENMATERIAL	41
3.1.1	Beschreibung des Datenmodells der „Zentralen Forsteinrichtungsdatenbank“ der Bayerischen Staatsforsten.....	41
3.1.2	Datenstand zum 01.03.2005	47
3.1.3	Räumliche Verteilung des Datenmaterials.....	48
3.1.4	Auswahl von Forstinventurdaten zur weiteren Auswertung	50
3.1.5	Zahlenmäßige Beschreibung des verwendeten Zahlenmaterials	51
3.1.6	Datenvorverarbeitung	56
3.1.7	Transformation der Daten.....	59
3.2	METHODEN	61
3.2.1	Data-Mining	61
3.2.1.1	Methoden des Data-Mining im Überblick.....	61
3.2.1.2	Assoziationsmethoden	61
3.2.1.3	Segmentations- und Clustermethoden.....	62
3.2.1.4	Diskriminanzanalyse (Klassifikation)	64
3.2.1.5	Beschreibung der Algorithmen CHAID, ID3, C4.5 und CART.....	66
3.2.2	Nichtlineare Regression	71
3.2.3	Anwendung von Data-Mining-Methoden zum Auffinden standörtlicher Muster in Forstinventurdaten.....	74
3.2.4	Online-Befragung zum Vergleich von Forsteinrichtungsdatenquellen in Deutschland.....	79

4.	ERGEBNISSE	83
4.1	<i>REGRESSIONSANALYSEN ZUR EINTEILUNG DER FORSTINVENTURDATEN EINER REGION.....</i>	83
4.2	<i>BEISPIELHAFTE DARSTELLUNG DER ENTSCHEIDUNGSBÄUME FÜR WUCHSBEZIRK 6.2</i>	94
4.2.1	<i>Entscheidungsbäume für die Baumart Buche</i>	95
4.2.2	<i>Entscheidungsbäume für die Baumart Douglasie.....</i>	98
4.2.3	<i>Entscheidungsbäume für die Baumart Eiche</i>	98
4.2.4	<i>Entscheidungsbäume für die Baumart Fichte</i>	101
4.2.5	<i>Entscheidungsbäume für die Baumart Kiefer</i>	105
4.2.6	<i>Entscheidungsbäume für die Baumart Lärche.....</i>	107
4.2.7	<i>Entscheidungsbäume für die Baumart Tanne</i>	109
4.2.8	<i>Zusammenfassung der Ergebnisse zur Entscheidungsbaumgenerierung</i>	111
4.3	<i>POTENZIELLE HÖHENWACHSTUMSGÄNGE DARGESTELLT AM BEISPIEL DER HAUPTBAUMARTEN IM WUCHSBEZIRK 6.2 „SÜDLICHE FRANKENALB UND SÜDLICHER OBERPFÄLZER JURA“</i>	111
4.3.1	<i>Potenzielle Höhenwachstumsgänge der Baumart Buche.....</i>	112
4.3.2	<i>Potenzielle Höhenwachstumsgänge der Baumart Douglasie</i>	113
4.3.3	<i>Potenzielle Höhenwachstumsgänge der Baumart Eiche.....</i>	114
4.3.4	<i>Potenzielle Höhenwachstumsgänge der Baumart Fichte.....</i>	116
4.3.5	<i>Potenzielle Höhenwachstumsgänge der Baumart Kiefer.....</i>	117
4.3.6	<i>Potenzielle Höhenwachstumsgänge der Baumart Lärche</i>	118
4.3.7	<i>Potenzielle Höhenwachstumsgänge der Baumart Tanne.....</i>	119
4.3.8	<i>Zusammenfassung der Ergebnisse zu den dargestellten potenziellen Höhenwachstumsverläufen.....</i>	121
4.4	<i>VALIDIERUNG DER BERECHNUNGSERGEBNISSE DURCH VERGLEICH VON HÖHENZUWÄCHSEN AUS SILVA-SIMULATIONEN MIT REALEN HÖHENZUWÄCHSEN AUF VERSUCHSFLÄCHEN.....</i>	121
4.4.1	<i>Flächenauswahl und Vorgehensweise</i>	121
4.4.2	<i>Vergleichende Darstellung der mittleren Höhenzuwächse.....</i>	124
4.4.3	<i>Statistische Überprüfung der mittleren Höhenzuwächse aus Realität und Simulation.....</i>	131
4.5	<i>AUSWERTUNGSERGEBNISSE DER ONLINE-BEFRAGUNG ZUR VERGLEICHENDEN DARSTELLUNG DER HALTUNG VON FORSTINVENTURDATEN IN DEUTSCHLAND</i>	134
5.	ERSTELLUNG EINER SOFTWARELÖSUNG ZUR FEINJUSTIERUNG VON SILVA.....	141
5.1	<i>GRUNDLEGENDE KONZEPTION</i>	141
5.2	<i>TECHNISCHE KONSTRUKTION</i>	144
5.3	<i>PROGRAMMTECHNISCHE DOKUMENTATION.....</i>	146
5.4	<i>BENUTZERDOKUMENTATION</i>	147
6.	DISKUSSION	155
6.1	<i>ZUGRUNDELIEGENDE DATEN.....</i>	155
6.2	<i>DISKUSSION DER ANGEWENDETEN METHODEN.....</i>	163
6.3	<i>DISKUSSION DER VORGESTELLTEN ERGEBNISSE</i>	166
6.4	<i>AUSBLICK.....</i>	168
7.	ZUSAMMENFASSUNG UND SUMMARY	171
7.1	<i>ZUSAMMENFASSUNG.....</i>	171
7.2	<i>SUMMARY.....</i>	173

8.	LITERATUR	175
9.	ANHANG	191
9.1	<i>TABELLE: STAATSWALDFLÄCHE BAYERNS NACH WUCHSREGIONEN.....</i>	<i>191</i>
9.2	<i>TABELLE: AUFSTELLUNG DER FORSTEINRICHTUNGEN, DEREN DATENSÄTZE FÜR DIESE ARBEIT VERWENDUNG FINDEN</i>	<i>194</i>
9.3	<i>TABELLEN ZUR NUMERISCHEN CHARAKTERISIERUNG DES VERWENDETEN DATENMATERIALS.....</i>	<i>203</i>
9.4	<i>INSTALLATIONSANLEITUNG FÜR DIE SOFTWARELÖSUNG „SILVA-PEM“</i>	<i>210</i>
9.5	<i>QUELLCODE DES SILVAPEM-MODULS „NONLINEARREGRESSION.JAVA“</i>	<i>211</i>
9.6	<i>LEBENS LAUF</i>	<i>213</i>
9.7	<i>ERKLÄRUNG DER SELBSTÄNDIGEN ANFERTIGUNG</i>	<i>215</i>

We are drowning in information,
but starving for knowledge

(John Naisbitt, 2000)

0. Vorwort

Die vorliegende Arbeit entstand in der Zeit von Ende 2003 bis Anfang 2007 am Lehrstuhl für Waldwachstumskunde der Technischen Universität München unter Leitung von Herrn Prof. Dr. Hans Pretzsch. Ich möchte hiermit die Gelegenheit nutzen und mich bei ihm für die interessante und praxisbedeutsame Themenstellung, die wissenschaftliche Anleitung und Diskussion und die Betreuung der Arbeit herzlich bedanken.

Die vorliegende Arbeit ist zu einem Großteil auch den hervorragenden Arbeitsbedingungen am Lehrstuhl für Waldwachstumskunde zu verdanken. Besonders profitiert habe ich von den anregenden Diskussionen mit zahlreichen Kollegen von denen ich hier nur stellvertretend Herrn Dr. Peter Biber, Herrn Dr. Thomas Seifert, Herrn Martin Nickel und Herrn Leonhard Steinacker sowie Herrn Dr. Heinz Utschig aufführen möchte. Nicht vergessen möchte ich auch die stets freundliche und äußerst hilfreiche Unterstützung durch Frau Marga Schmid.

Die Arbeit hat sich u. a. zum Ziel gesetzt, eine Methodik zu entwickeln und diese softwaretechnisch umzusetzen. Die programmtechnische Realisierung wurde dabei durch Herrn Stefan Seifert sowie Herrn Dr. Christian Rössl durch wertvolle Hilfestellungen und Fundstellenhinweise unterstützt, wofür ich mich ebenfalls herzlich bedanken möchte.

Ebenso haben an der Arbeit tatkräftig mitgearbeitet die (z. T. ehemaligen) Zivildienstleistenden und Hilfskräfte des Lehrstuhls, von denen ich stellvertretend Herrn Florian Kohlhuber, Herrn Sebastian Seibold sowie Frau Felicitas Eyrainer benennen möchte.

Weiterhin verdankt die Arbeit ihre Praxisorientierung den intensiven Kontakten und Diskussionen mit Beschäftigten der Bayerischen Landesanstalt für Wald und Forstwirtschaft sowie mit Beschäftigten der ehemaligen Bayerischen Staatsforstverwaltung. Stellvertretend für erstere Gruppe möchte ich hier insbesondere Herrn Dr. Herbert Borchert sowie Herrn Dr. Christian Kölling hervorheben, die die Arbeit durch wertvolle Diskussionen sowie durch die Bereitstellung von Datenmaterial wesentlich beeinflusst haben. Bei den Mitarbeitern der ehemaligen Bayerischen Staatsforstverwaltung möchte ich mich insbesondere bei Herrn Prof. Dr. Andreas Rothe, Frau Karin Gerstendörfer und Herrn Stefan Gerstendörfer für wegbereitende Arbeitsschritte bedanken.

Die vorgelegte Arbeit wurde zu Zeiten begonnen, in denen die Wälder in Bayern durch eine Einheitsforstverwaltung mit der Bezeichnung „Bayerische

Staatsforstverwaltung“ bewirtschaftet wurden. Etwa zur Hälfte der Projektlaufzeit erfolgte durch einen Reformbeschluss der Bayerischen Staatsregierung eine Aufteilung in ein privatwirtschaftlich orientiertes Unternehmen zur Bewirtschaftung des Waldes im Eigentum des Freistaates Bayern mit der Bezeichnung „Bayerische Staatsforsten“ sowie eine mit hoheitlichen Aufgaben betraute Forstverwaltung mit dem neuen Namen „Bayerische Forstverwaltung“. Diese Forstreform hat die vorliegende Arbeit nur wenig beeinflusst, da vor dem Zeitpunkt der Umsetzung der Forstreform bereits alle notwendigen Daten beschafft werden konnten. Nach der Umsetzung der Forstreform konnte insbesondere durch aufgebaute persönliche Kontakte und das rege Interesse von beiden Seiten ein reibungsloser und rascher Arbeitsfortschritt gesichert werden. Hierfür möchte ich mich bei den Mitarbeitern sowohl bei den Bayerischen Staatsforsten als auch bei den Mitarbeitern der Bayerischen Forstverwaltung bedanken. Namentlich möchte ich hier Frau Margret Kolbeck, Herrn Georg Thanbichler sowie Herrn Thomas Zanker benennen.

Die benannte Forstreform hatte lediglich insofern Auswirkungen auf die Arbeit, als dass die Bezeichnung der Stelle, die für die Bewirtschaftung des Staatswaldes und damit für die in dieser Arbeit verwendete Datengrundlage, anfänglich Probleme bereitete. Aus Gründen der Vereinheitlichung wird daher nachfolgend die Bezeichnung „Bayerische Staatsforsten“ verwendet, wohlwissend dass die verwendeten Forstinventurdaten des Staatswaldes in Bayern der ehemaligen Bayerischen Staatsforstverwaltung entstammen und jetzt von den Bayerischen Staatsforsten gehalten und weitergeführt werden.

Bedanken möchte ich mich auch bei der Bayerischen Forstverwaltung, die durch die Einrichtung der sogenannten „Grünen Stellen“ an der Technischen Universität München jungen Forstbeamten und Forstbeamtinnen die Möglichkeit gibt, sich in einem überschaubaren Zeitraum wissenschaftlich weiterzuqualifizieren. In diesem Zusammenhang gilt mein Dank auch dem Kuratorium der Bayerischen Landesanstalt für Wald und Forstwirtschaft für die finanzielle und ideelle Unterstützung des Projektes W 35.

Außerhalb der staatlichen Forstverwaltung verdanke ich auch zahlreiche maßgebliche Impulse für diese Arbeit den Diskussionen mit Herrn Gräfl. Forstdirektor Wolf-Dieter Radike der Gräfl. Arco-Zinneberg'schen Forstverwaltung.

Die Realisierung einer Forschungsarbeit erfordert neben einem produktiven Arbeitsumfeld auch ein privates Umfeld, welches die Arbeit voll unterstützt und mitträgt. In diesem Zusammenhang möchte ich mich bei meinen Schwiegereltern, bei meinen Eltern und insbesondere bei meiner Ehefrau Elke herzlich bedanken.

Freising, April 2007

1. Einleitung

1.1 Ausgangssituation

Tycho Brahe, ein dänischer Adelige und Astronom, ausgerüstet mit den besten Instrumenten seiner Zeit, sammelte in seiner Hauptschaffensphase zwischen den Jahren 1570 - 1580 große Mengen an Daten über das Planetensystem. Er war allerdings nicht in der Lage, sie zu einem einheitlichen Schema zusammenzufassen. Er konnte zwar z. B. genau sagen an welchem Ort des Himmels der Mars an einem bestimmten Tag des Jahres 1584 gestanden hatte, er war aber nicht in der Lage, die verschiedenen Positionen an verschiedenen Tagen so durch eine Theorie in Beziehung zu setzen, dass sie mit seinen hochgenauen Beobachtungen übereinstimmten. Alle Hypothesen, die er testete, ließen sich nicht mit seinen Beobachtungen vereinen. Zwar entwickelte er das tychonische Planetensystem, nach dem sich Sonne und Mond um die Erde, alle anderen Planeten aber um die Sonne bewegen, doch erreichte weder er noch seine wissenschaftliche Arbeit große Bekanntheit. Sein Problem bestand darin, dass er zwar über große Mengen an hochwertigen Daten verfügte, er allerdings das darin verborgene „Wissen“ nicht extrahieren konnte.

Johannes Kepler war ein deutscher Astronom und Gehilfe Brahes, der sein Leben lang versuchte, die Gesetzmäßigkeiten mit mathematischen Methoden zu finden, die die Bewegungen der Planeten bestimmen. Ausgangspunkt für seine Arbeiten bildeten die von Brahe gesammelten Daten. Letztendlich fand er das in den Daten „verborgene Wissen“ und beschrieb es in seinem weltweit bekannt gewordenen Lebenswerk „Astronomica nova“ (Borgelt, 2004).

Diese Überlieferung aus der frühen Neuzeit zeigt bereits erste Tendenzen einer Entwicklung, die heutzutage in vielen Lebensbereichen aufscheint. Der rasche Fortschritt in der Entwicklung informations- und kommunikationstechnologischer Systeme hat sich auf das Verhalten, Daten zu erfassen und zu sammeln ausgewirkt. Digitale Speichermedien und Datenbanken bergen heutzutage riesige Datenmengen. Imation, einer der führenden Anbieter wechselbarer Speichermedien, bezeichnet die Menge der jährlich weltweit produzierten Daten mit etwa zwei ExaByte. Dies entspricht einer Menge von zwei Milliarden GigaByte (iBusiness, 2004).

Auf der anderen Seite hat die Konvergenz der sich immer rascher weiterentwickelnden Informations- und Kommunikationstechnologie weltweit Gesellschaften geschaffen, die von rasch verfügbaren Informationen leben.

Definiert man Daten als aufgezeichnete Fakten, so sind Informationen die Menge von Mustern und Erwartungen die hinter diesen Daten stecken (Witten und Frank, 2005). Aufgrund der stetig und rasch wachsenden Datenmenge auf der einen Seite und der Notwendigkeit der schnellen Verfügbarkeit von entscheidungsrelevanten Informationen andererseits wächst ein dringendes Bedürfnis für datenverarbeitungsgestützte Theorien und Werkzeuge, die die Menschen dahingehend unterstützen, dass sie neue, potenziell nützliche bzw. entscheidungsrelevante Informationen aus den rasch wachsenden digitalen Datenmengen mehr oder weniger autonom extrahieren (Fayyad et al., 1996).

Auch die praktische Forstwirtschaft sowie die nachgelagerte Holzwirtschaft stellen Wirtschaftszweige in unserer Gesellschaft dar, in denen heutzutage viele Daten entstehen und gespeichert werden. Bei Forstinventuren werden ständig große Datenmengen erhoben. In der Bestands- und Materialbuchhaltung der Forstbetriebe wird je nach Wertigkeit für jeden Stamm oder für jedes Verkaufslös ein Datensatz erzeugt und gehalten. Für jeden dieser Datensätze werden im Rahmen des externen und internen Rechnungswesens wiederum neue Datensätze generiert. Beim Eingang der verkauften Holzsortimente in weiterverarbeitende Betriebe werden die Daten direkt übernommen und um Daten, die den eigenen Informationsbedürfnissen genügen, erweitert.

Diese kurze und bei weitem unvollständige Aufzählung zeigt nur einen kleinen Ausschnitt der Datenentstehung im klassischen Produktionsprozess der Forstwirtschaft und Holzwirtschaft. Unterstützt wird diese Entwicklung in diesem Sektor nicht zuletzt durch die zunehmende Leistungsfähigkeit, Robustheit und Verfügbarkeit mobiler Datenerfassungstechnologien.

Die vorliegende Arbeit beschränkt sich in diesem Stadium lediglich auf Datenquellen mit waldwachstumskundlichen Inhalten aus Bayern. Ihr unterliegt die Überzeugung, dass zum Zeitpunkt der Entstehung der Arbeit in Bayern mehrere Datenquellen existieren, die bisher nicht annähernd ausgeschöpfte Informationen zum Waldwachstum enthalten.

Einen Datenfundus, der einen erheblichen Wert für den Erkenntnisprozess der Waldwachstumsforschung besitzt, stellt die Versuchsflächendatenbank des Lehrstuhls für Waldwachstumskunde der Technischen Universität München dar. In dieser Datenbank werden Daten des langfristigen, ertragskundlichen Versuchswesens in Bayern mit größtmöglicher Sorgfalt archiviert und für verschiedenste Auswertungszwecke genutzt. Ein besonderes Charakteristikum stellen die hohen Detailinformationen auf Einzelbaumbene dar, die durch diese Daten vermittelt werden. Daher bildet diese Datenquelle seit vielen Jahrzehnten eine verlässliche Grundlage

für die Herleitung von Gesetzmäßigkeiten des Waldwachstums. Sie besitzt einen unschätzbaren Wert für die Forstwissenschaft und die praktische Forstwirtschaft.

Daneben verfügt die Bayerische Landesanstalt für Wald und Forstwirtschaft mit den Daten der Bundeswaldinventur für Bayern über stichprobenbasierte Daten, die flächenrepräsentativ über alle Wuchsgebiete und Waldbesitzarten Aussagen über den Waldzustand und die Waldentwicklung in Bayern ermöglichen (Bayerische Landesanstalt für Wald und Forstwirtschaft, 2005).

Insbesondere auf die Beantwortung einzelbetrieblicher Fragestellungen zugeschnitten, existiert mit den Daten der Forstbetriebsinventuren der Bayerischen Staatsforsten eine Datensammlung von nicht abschätzbarem Wert (Neft, 2005). Letztere Datenquelle setzt sich ebenso wie die Bundeswaldinventurdaten aus Stichprobendaten zusammen. Im Vergleich zu den Bundeswaldinventurdaten weist das Stichprobenkonzept der Bayerischen Staatsforsten allerdings eine etwa hundertfach höhere Stichprobendichte auf, wodurch detailliertere Aussagen zu einzelbetrieblichen Fragestellungen ermöglicht werden.

1.2 Einsatz von Wachstmodellen in der Praxis zur Unterstützung forstlicher Entscheidungen

Die Sammlung von Daten sowie der damit verbundene Aufbau einer empirischen Informationsbasis stellen den ersten wichtigen Schritt im Erkenntnisprozess der Waldwachstumsforschung dar, der in der Aufdeckung von Gesetzmäßigkeiten und letztendlich in der Entwicklung von Wachstmodellen mündet (Pretzsch, 2001). Die programmtechnischen Umsetzungen von Wachstmodellen werden als Waldwachstumssimulatoren bezeichnet. Die Entwicklung derartiger Programme reicht mittlerweile mehr als drei Jahrzehnte zurück und wird weitgehend durch die rasche technologische Entwicklung im Bereich der elektronischen Datenverarbeitung unterstützt (Vanclay, 1994; Pretzsch, 2001; Hasenauer, 2006).

Mit der Änderung der Art der Waldbewirtschaftung, weg von gleichaltrigen Reinbeständen hin zu einer kleinflächigen bis einzelbaumorientierten Bewirtschaftung, entsprechen klassische Planungshilfen nicht mehr den Informationsbedürfnissen einer zeitgemäßen, nachhaltigen Bewirtschaftung unserer Wälder (Hasenauer, 2004; Sterba, 2002; Pretzsch, 2001).

Eine weitere wichtige Herausforderung für die praktische Forstwirtschaft stellen die sich ändernden Klimabedingungen dar. Auch in diesem Zusammenhang sind flexible Wachstmodelle das einzig adäquate Hilfsmittel zur Abschätzung und Bewertung möglicher Auswirkungen von Klimaveränderungen auf die Wälder, wie verschiedene z. T. interdisziplinäre Forschungsarbeiten der letzten Jahre bewiesen haben (Döbbeler, 2004; Lindner et. al, 2002; Pretzsch und Utschig, 2000).

Neben sich ändernden Waldaufbauformen und veränderten Wuchsbedingungen wachsen in vielen Staaten der Erde die Anforderungen der Gesellschaft an die Wälder. Standen in der Gründerzeit der planmäßigen Forstwirtschaft Nachhaltigkeitsaspekte im Zusammenhang mit der Produktion von Holz im Vordergrund, so ist es heute notwendig, im Sinne einer multifunktionalen, nachhaltigen Forstwirtschaft ein breiteres Spektrum von Leistungen und Funktionen des Waldes in forstliche Planungen mit einzubeziehen (Spellmann, 2003; Pretzsch, 2003a).

Um die Konsequenzen forstlichen Handelns in einer sich immer rascher ändernden Umwelt besser abschätzen zu können und weiterhin konsequent einem multifunktional ausgerichteten Nachhaltigkeitsgedanken Rechnung zu tragen, braucht die moderne Forstwirtschaft zeitgemäße, zuverlässige Planungswerkzeuge, die sich mit ihrem Informationsangebot möglichst optimal in das vorstehend kurz skizzierte Spannungsfeld einpassen. Klassische, meist statische Planungshilfen wie Ertragstafeln können konzeptionell bedingt diese Anforderungen nur noch unzureichend erfüllen. Moderne Planungshilfen wie die eingangs dieses Abschnitts angeführten Wachstumsmodelle sind hingegen aufgrund der zunehmenden Leistungsfähigkeit von EDV-Systemen in der Lage, diese Informationslücke zu schließen. Konzeptionell sind dabei einzelbaumbasierte, managementorientierte Wachstumsmodelle als Planungshilfen für die praktische Forstwirtschaft besonders geeignet, da sie es ermöglichen, das Wachstum von Rein- und Mischbeständen aller Alterszusammensetzungen und Strukturen nachzubilden. Aus diesem Grund beziehen sich die weiteren Ausführungen auf diesen Modelltyp und hier insbesondere auf das forstliche Managementwachstumsmodell SILVA, dessen Entwicklung vor ungefähr 15 Jahren am Münchner Lehrstuhl für Waldwachstumskunde begann und das seit Entwicklungsbeginn stetig weiterentwickelt und verfeinert wird. Neben klassischen Größen der Naturalproduktion können derartige Wachstumsmodelle für jeden einzelnen simulierten Baum neben klassischen ertragskundlichen Größen auch ökonomische, ökologische und sozioökonomische Variablen mitführen und dem Benutzer auf unterschiedlichen Aggregationsebenen vom Einzelbaum über den Bestand oder Betrieb bis hinauf auf Länder- oder Regionalebene ausgeben. In der praktischen Forstwirtschaft sind sie daher in vielen Ländern der Erde bereits ein wichtiges Planungshilfsmittel. Bekannte Beispiele für in der forstlichen Praxis eingesetzte Wachstumsmodelle sind der Forest Vegetation Simulator, der in weiten Teilen der USA Anwendung findet (FVS, 2006) oder das Modell Prognosis (PROGNOSIS, 2006), das in mehreren Regionen British Columbias in Kanada erfolgreich eingesetzt wird. Einen guten, wenngleich nicht ganz aktuellen Überblick über die in den Pazifikküstenstaaten der USA entwickelten und in der Praxis angewendeten Wachstumsmodelle gibt Ritchie (1999).

Aufgrund des Erkenntnisfortschrittes der Waldwachstumsforschung im Zusammenhang mit stark differenzierten Waldaufbauformen werden derartige Modelle zunehmend auch in Mitteleuropa für verschiedenste Einsatzzwecke eingesetzt oder nachgefragt. Pretzsch (2001) sowie Schmidt et al. (2006) sehen drei wichtige Einsatzgebiete in der forstlichen Praxis. Zum einen ermöglichen es einzelbaumorientierte Wuchsmodelle auf Bestands- oder Stratenebene waldbaulich optimierte Behandlungsprogramme zu entwickeln. Zum anderen ermöglichen sie es insbesondere aufgrund der häufig angeschlossenen Visualisierungsroutinen (Pretzsch und Seifert, 2000; Knaufft, 2003), verschiedene Waldbauverfahren und Handlungsoptionen im Rahmen von Schulungen und Ausbildungsveranstaltungen an die Praxis zu vermitteln. Und schließlich stellen sie im Rahmen der kurz- und mittelfristigen Forstbetriebsplanung durch die weitgehend automatisierte Aggregation der Einzelergebnisse von Bestands- oder Stratenbehandlungsvarianten wichtige Entscheidungshilfen zur Verfügung, die neben der zahlenmäßigen Entscheidungsvorbereitung bei geeignetem Einsatz zudem ein erhebliches betriebliches Einsparpotenzial mit sich bringen. Böckmann (2004) berichtet z. B. von einer Kosteneinsparung in der Inventur von 40 % bei nahezu gleichbleibendem Informationsgehalt durch den standardmäßigen Einsatz des Wuchsmodells BwinPro im Rahmen von niedersächsischen Forsteinrichtungen.

Damit Wuchsmodelle eine breite Anwendung in der forstlichen Praxis finden, sind nach Meinung zahlreicher Autoren mehrere Grundvoraussetzungen entscheidend: Zum einen müssen Planungswerkzeuge für eine nachhaltige Forstwirtschaft verlässliche, zutreffende Vorhersagen für Waldbestände jeglicher Alters- und Baumartenzusammensetzung ermöglichen (von Teuffel et al., 2006; Pretzsch, 2001). Um die Prognosegenauigkeit von Waldwachstumsmodellen zu erhöhen, sind nach einer grundlegenden Evaluierung des Modellansatzes rückgekoppelte Validierungs-, Kalibrierungs- und Neujustierungszyklen notwendig.

Neben einer möglichst hohen Prognosegenauigkeit des Modells muss es den Erstellern zudem gelingen, die Anforderungen der Anwender zu erfüllen. Nach Pretzsch (2001) hängt die Akzeptanz, Verbreitung und Ausbaubarkeit von Waldwachstumssimulationsprogrammen im Wesentlichen von der Benutzerfreundlichkeit, der Flexibilität in den Einsatzmöglichkeiten, den vorgesehenen Schnittstellen zu vor- und nachgeschalteten Programmen und Datenbanken und von der Dokumentation der Software ab. Lemm (2004) und von Teuffel et al. (2006) teilen diese Auffassung und führen zudem als wichtige Anforderung der Praxis eine gute Kommunizierbarkeit der Modellergebnisse sowie eine hinreichende Dokumentation des Simulationsprozesses an.

1.3 Datengrundlagen zur Parametrierung und Justierung von einzelbaum-basierten Wachstumsmodellen

Für einen Einsatz in der forstlichen Praxis ist es zweckdienlich, forstliche Wachstumsmodelle in den Informationsfluss der Forstwirtschaft zu integrieren. Hierzu wäre es wünschenswert, wenn einerseits die vom Modell benötigten Start- und Steuerungsgrößen kompatibel mit dem Informationsangebot des Modellanwenders sind und andererseits die vom Modell erzeugten Ergebnisgrößen den Informationsanforderungen des Anwenders entsprechen. Gleichzeitig sollen alle vorhandenen Datenquellen in Verbindung mit dem aktuellen Wissensstand der Waldwachstumsforschung bestmöglich für Modellerstellung und Modellanwendung genutzt werden.

Etwa ab den 60er Jahren des letzten Jahrhunderts wurden bei der Aufnahme von Versuchsflächen zunehmend Einzelbaumgrößen erfasst. So ist mit den Daten des langfristigen forstlichen Versuchswesens eine Datenquelle entstanden, die Informationen zu Dimension, Umgebungsstruktur und Zuwachsverhalten von Einzelbäumen enthält. Durch den langen, kontinuierlichen Beobachtungsverlauf wird hierdurch die Erforschung räumlicher und zeitlicher Aspekte des Waldwachstums insbesondere auf Einzelbaumebene in nahezu allen strukturellen Aspekten möglich. Für die grundlegende Parametrisierung des Wachstumsmodells SILVA 2.2 wurden fast ausschließlich Versuchsflächendaten genutzt. Diese Parametrisierung stützt sich dabei vor allem auf folgende drei wichtigen Datensätze: Das Standort-Leistungs-Modell, welches wie die anderen Teilmodelle des Wachstumsmodells SILVA in Abschnitt 2.2 grundlegend beschrieben wird, kann sich auf ein Datenmaterial aus 330 langfristigen Versuchsflächen aus mehreren Bundesländern Deutschlands sowie der Schweiz stützen (Kahn, 1994). Die Nachbildung der Stammform und Rindenstärke greift auf das Programm BDAT von Kublin und Scharnagl (1988) zurück, für dessen Einstellung Daten von 30.000 sektionsweise vermessenen Bäumen zur Verfügung standen. Die Parametrisierung der weiteren Teilmodelle des Wachstumsmodells SILVA 2.2, wie etwa des Zuwachs-, Kronenform-, Mortalitäts- oder Verjüngungsmodells erfolgte im Wesentlichen mit den Daten des langfristigen Versuchsflächennetzes des Münchner Lehrstuhls für Waldwachstumskunde, welches neben 160 Versuchsflächen mit insgesamt ca. 1.000 Versuchspartzen in Bayern auch aus Flächen in Rheinland-Pfalz und Niedersachsen aufgebaut ist (Pretzsch, 2001).

Neben den Daten des langfristigen ertragskundlichen Versuchswesens für das Bundesland Bayern existieren weitere Datenquellen, die potenziell Informationen zum Waldwachstum in Bayern enthalten.

Dies sind zum einen die Daten der Bundeswaldinventur, die für Bayern 1987 und 2002 durchgeführt wurde. Es handelt sich hierbei um eine Stichprobeninventur, die auf repräsentativ verteilten Stichprobenpunkten auf Waldflächen aller

Waldbesitzarten in Bayern den Zustand des Waldes erfasst. Methodisch ist die Bundeswaldinventur als permanente Stichprobeninventur konzipiert, bei der Waldflächen, die auf ein 4 km * 4 km-Raster treffen, aufgenommen werden, wobei für waldarme Regionen eine Verdichtung des Rasternetzes konzeptionell möglich ist. Die Datenaufnahme an den Inventurpunkten erfolgt über Winkelzählproben bzw. über Probekreise mit unterschiedlichen Radien. An jedem Gitterkreuzungspunkt wurden Probekreise mit Radien von 1,75 m, 5 m, 10 m, und 25 m gezogen. Verjüngung wird an einem seitlich versetzten Punkt in einem Probekreis mit 1 m Radius aufgenommen. Bei der Bundeswaldinventur II (2002) wurden so rund 99.000 Bäume an 7.681 Stichprobenflächen in Bayerns Wäldern erfasst (Bayerische Landesanstalt für Wald und Forstwirtschaft, 2005).

Daneben existiert mit den Daten der betrieblichen Stichprobeninventuren im Staatswald Bayerns ein Datensatz, der für die ca. 800.000 ha Wald im Eigentum des Freistaats Bayern detailliert betriebsweise Aussagen ermöglicht. Die Aufnahme dieser Daten erfolgt im Rahmen des Forstinventurverfahrens als Teil eines Forsteinrichtungsverfahrens. Bei dem im Staatswald Bayerns im Regelfall angewendeten Forstinventurverfahren handelt es um die Aufnahme temporär oder permanent angelegter, konzentrischer Probekreise mit gestaffelten Kluppschwellen. Die Bayerische Staatsforstverwaltung hat sich für dieses Aufnahmeverfahren entschieden, um den Aufwandaufwand für die oft stammzahlreiche Verjüngung auf der relativ kleinsten Probefläche gering zu halten und zugleich das Baumholz in genügend hoher Anzahl aufzunehmen. Die Probekreisgrößen und Kluppschwellen werden für jeden Betrieb vom Inventurleiter bestimmt. Die Aufnahme erfolgt nach dem skizzierten Verfahren seit mittlerweile über 30 Jahren, die Einführung der permanenten Stichprobeninventur erfolgte 1984 im damaligen Bayerischen Forstamt Ebrach (Fuchs, 1993). Sämtliche Daten, die bei einer Forsteinrichtung (Inventur und Begang) erhoben werden, werden in einer Informix-Datenbank auf einem als Sektionsrechner bezeichneten Computersystem abgelegt. Die Synchronisation dieser dezentral gehaltenen Datenbanken erfolgt mit Hilfe nächtlicher Datenfernübertragung mit einer zentral gehaltenen Aufnahmedatenbank. Nach Abschluss eines Forsteinrichtungsverfahrens werden die Daten in eine Zentrale Forsteinrichtungsdatenbank zur Sicherstellung der Datenintegrität übertragen (Bayerisches Staatsministerium für Landwirtschaft und Forsten, 1999). Eine genauere Beschreibung des Datenmodells sowie des Datenstandes der Zentralen Forsteinrichtungsdatenbank der Bayerischen Staatsforsten erfolgt in Abschnitt 3.1.1 dieser Arbeit.

Die kurz vorgestellten waldwachstumskundlichen Datengrundlagen unterscheiden sich sowohl quantitativ als auch qualitativ sehr stark. In Tabelle 1 sind die wesentlichen Unterschiede zusammenfassend gegenübergestellt.

Tab. 1: Gegenüberstellung der wichtigsten Unterschiede von Daten des langfristigen ertragskundlichen Versuchswesens in Bayern, Daten der Bundeswaldinventur für Bayern sowie Daten der forstbetrieblichen Stichprobeninventuren im Staatswald Bayerns.

Kriterium	Daten des langfristigen ertragskundlichen Versuchswesens	Daten der Bundeswaldinventur in Bayern	Daten der betrieblichen Stichprobeninventuren im Staatswald Bayerns
Vorliegende Datenmenge	vergleichsweise viele Daten	vergleichsweise wenige Daten	vergleichsweise viele Daten
Aufnahmeverfahren	Vollaufnahme	Stichprobenaufnahme	Stichprobenaufnahme
Flächenrepräsentanz	gering (vgl. Abb. 10)	mittel-hoch	hoch (vgl. Abb. 9)
Standortrepräsentanz	gering (vgl. Abb. 10)	gering-mittel	hoch (vgl. Abb. 9)
Repräsentanz der Besitzstruktur	Staatswald, (Privatwald, Körperschaftswald)	Staatswald, Privatwald, Körperschaftswald	Staatswald
Messprogramm	zielt auf detaillierte, strukturelle Informationen auf Einzelbaumebene ab	bestimmt durch Aussagemöglichkeit auf Regional- und Landesebene	bestimmt durch Aussagemöglichkeit auf regionaler und organisatorischer Ebene
Zeitreiheneigenschaft	echte Zeitreiheninformationen (Längsschnittdaten, 'longitudinal data')	echte Zeitreiheninformationen (Längsschnittdaten, 'longitudinal data')	(ohne Vorverarbeitungsschritt) unechte und echte Zeitreiheninformationen (Querschnittdaten und Längsschnittdaten, 'panel-data' oder 'time-series cross-sectional data')
Existenz	seit 1870	seit 1987	seit 1984
Zeitlicher Charakter	permanent (kontinuierliche Beobachtung der gleichen Bäume)	permanent (kontinuierliche Beobachtung der gleichen Bäume)	temporär und permanent
Dichte der Probeflächen	unregelmäßige Verteilung der Versuchflächen in Bayern, (Rheinland-Pfalz, Niedersachsen)	gitternetzförmig, i.d.R. 4*4 km Raster im Wald (Verdichtungen möglich)	i.d.R. an einem Gitter ausgerichtete Verteilung der Stichprobenpunkte mit variablen Gitternetzabständen; Punktdichten zwischen 1 Punkt pro Hektar und 0,25 Punkten pro Hektar)
Aufnahmeintervall	alle 5 bis 7 (in Ausnahmefällen 9) Jahre	15 Jahre	alle 10 bis 15 Jahre
Aufnahmeprogramm	durch Versuchsfrage bestimmt		
Detailgrad auf Einzelbaumebene	sehr hoch	mittel, (da räumliche Position der Einzelbäume bekannt (permanenter Charakter))	niedrig (temporär) - mittel (permanent)

1.4 Begriffsabgrenzungen

Beschäftigt man sich mit dem Themenkomplex „Einstellen von Modellgleichungen“, so trifft man im deutschsprachigen Raum zwangsläufig immer wieder auf die Begriffe „Kalibrieren“, „Justieren“ und „Parametrisieren“.

Der Begriff „Kalibrieren“ wird im deutschen Sprachgebrauch kontextuell unterschiedlich angewendet. Nach der 21. Ausgabe des Duden versteht man unter „kalibrieren“ in der Technik „das Kaliber messen“ bzw. [Werkstücke] „auf ein geeignetes Maß bringen“ bzw. [Messinstrumente] „eichen“ (Duden 1996). Die freie Enzyklopädie Wikipedia definiert den Begriff „Kalibrierung“ im Sinne der Metrologie, der Lehre von der Maß- und Gewichtskunde. Danach ist „das Kalibrieren“ in der Metrologie ein Messprozess zur Feststellung und Dokumentation der Abweichung eines Messgerätes oder einer Maßverkörperung vom *richtigen Wert*, dargestellt durch ein Referenznormal. [...] Im Gegensatz zur Justierung wird die Abweichung zwischen Kalibergegenstand und Referenznormal nicht beseitigt. Der Begriff Kalibrierung umfasst keine Aussagen zur Einhaltung vorgegebener Spezifikationen (Messtoleranzen), das ist Gegenstand einer Prüfung.“ (WIKIPEDIA 2004a). Dagegen versteht man nach der gleichen Quelle „unter ‚Justieren‘ das exakte Einstellen einer Messgröße oder Anzeige eines Instrumentes durch einen fachmännischen Eingriff in das Messsystem. [...] Unter Nachjustierung versteht man ein geringfügiges oder verkürztes Verfahren, das z. B. bei periodischen Temperaturänderungen erforderlich ist.“ (WIKIPEDIA 2004b).

Besucht man das englischsprachige Wiki so findet man im metrologischen Sinne eine inhaltsgleiche Übersetzung des Begriffes „calibration“ (WIKIPEDIA 2004c). Im mathematisch-statistischen Zusammenhang findet sich ergänzend folgende Definition: „Calibration in statistics is a reverse process to regression. The *calibration problem* is the use of known data on the observed relationship between a dependent variable and an independent variable to make estimates of other values of the independent variable from new observations of the dependent variable. [...]“ (WIKIPEDIA 2004d).

Im Zusammenhang mit „Modellbildung“ wird ‚Kalibrierung‘ bzw. ‚model calibration‘ vielfach als Prozess verstanden, der die Aufstellung des mathematischen Modells von Anfang an begleitet bzw. der sich aufgrund neuer Erkenntnisse und sich ändernder Umweltbedingungen an dessen Evaluierung immer wieder anschließt. Letzteres wird z. B. von Vanclay (1994) auch als ‚Re-calibration‘ bezeichnet. Er definiert Re-calibration wie folgt: „Re-calibration implies adjusting a growth model so that it provides good predictions for a new population. This may entail estimating new parameters for some or all of the equations in the model, or may use a scaling factor to adjust predictions“.

Im Zusammenhang mit forstlichen Wuchsmodellen werden häufig mit „Kalibrierung“ und „Justierung“ auch die Begriffe „Parametrisierung“, „Parameterschätzung“ bzw. „Parameter estimation“ in Verbindung gebracht (vgl. z. B. Pretzsch 2001, Amaro 2003, Vanclay, 1994). Man versteht hierunter im forstlichen Bereich die Ableitung neuer Parameter mit Hilfe geeigneter statistischer Verfahren aus z. T. neuem Datenmaterial, ohne dass die dem biometrischen Modellaufbau zugrundeliegenden Funktionsgleichungen, die sich auf bewährte Wuchsgesetzmäßigkeiten stützen, geändert werden.

Im Zusammenhang mit der Modellierung von Wasserhaushaltssystemen werden die Begriffe „Parametrisierung“ bzw. „Parameter estimation“ häufig mit der Technik der Inversen Modellierung (Inverse modelling) in Verbindung gebracht. Beispiele hierfür finden sich u. a. in „Methods and guidelines for effective model calibration“ (HILL 1998). Hierbei handelt es sich um eine Technik, die versucht Funktionsgleichungen bzw. deren Parameter zu ermitteln, die eine möglichst gute Anpassung an beobachtete Daten erlauben. Da die ermittelten Funktionsgleichungen i. d. R. nur schwer mit bekannten Gesetzmäßigkeiten vereinbar sind, wird diese Technik in dieser Arbeit nicht weiter verfolgt.

Ausgehend von vorstehenden Erläuterungen soll im weiteren Verlauf der Arbeit unter dem Begriff „Parametrisieren“ eine grundlegende Neuermittlung der Koeffizienten der SILVA im Programmkern (vgl. Abb. 2) zugrundeliegenden Formelgleichungen verstanden werden.

Unter den Begriffen „Justierung“ bzw. „Feinjustierung“ soll hingegen eine Neuermittlung der Koeffizienten von Gleichungen in einzelnen SILVA-Teilmodellen mit dem Ziel der Erhöhung der Prognosegenauigkeit verstanden werden. Vorliegende Arbeit konzentriert sich dabei auf eine Justierung des Standort-Leistungsmodells von SILVA (vgl. Abb. 2) Die begriffliche Unterscheidung nach „Justierung“ und „Feinjustierung“ erfolgt deshalb, weil die Verwendung einer neuen Datengrundlage zur Einstellung des Wuchsmodells eventuell neue regionale oder standortbezogene Erkenntnisse mit sich bringt (Justierung) oder aber lediglich zu einer regionalen, standortbezogenen Präzisierung der Prognosen führt (Feinjustierung). Im weiteren Text wird einheitlich der Begriff „Feinjustierung“ für beide Aspekte verwendet.

1.5 Ziele der Arbeit

Die Grundidee dieser Arbeit besteht darin, für Feinjustierungszwecke des Wuchsmodells SILVA die Forsteinrichtungsdaten der Bayerischen Staatsforsten zu nutzen. Damit soll die Prognosegenauigkeit des Wuchsmodells SILVA insbesondere

für den Einsatz bei kurz- und mittelfristigen Prognosen (z. B. im Rahmen von Forsteinrichtungsverfahren) erhöht und so das Vertrauen von Anwendern aus der forstlichen Praxis gestärkt werden. Erstmals wird damit eine waldwachstums-kundliche Datenquelle der Forstpraxis, die sich quantitativ und qualitativ stark von der zur grundlegenden Parametrisierung der Kernfunktionen von SILVA verwendeten Datengrundlage unterscheidet, für Feinjustierungszwecke auf regionaler Ebene zum Einsatz kommen. Der Rückfluss der Ergebnisse dieser Arbeit entwickelt somit unmittelbaren Nutzen für die Forstpraxis in Bayern und beschränkt sich nicht nur auf den mit der Verwaltung des Bayerischen Staatswaldes betrauten Personenkreis.

Nach Pretzsch und Kahn (1998) ist es möglich, SILVA mit unabhängigen Daten in seiner Prognosegenauigkeit zu verbessern, um eine regionale (oder überregionale) Adjustierung der Modellfunktionen vorzunehmen, ohne tief in die Modellstruktur einzugreifen. Zum Zeitpunkt der Erstellung dieser Arbeit ist das Wachstumsmodell SILVA in der Version 2.2 für die Baumarten „Fichte“, „Kiefer“, „Tanne“, „Douglasie“, „Lärche“, „Eiche“, „Buche“ sowie für die Baumartengruppen „Sonstige Edellaubhölzer“ und „Sonstige Laubhölzer“ im Modellkern (vgl. Abb. 2) parametrisiert. Die Parametrisierung basiert dabei im Wesentlichen auf Daten des langfristigen ertragskundlichen Versuchswesens in Bayern. Wie in Abbildung 10 gezeigt wird, besteht das bayerische Versuchsflächennetz aus einer stark differenzierten Anzahl von Versuchsflächen für die einzelnen Baumarten. Zudem weist dieses Versuchsflächennetz eine regional unterschiedliche Versuchsausstattung auf. Ein Nebenziel dieser Arbeit besteht daher darin, die baumartenweise flächen- und standortrepräsentativen Baumdaten der Forstinventur in Bayern zu nutzen, um das für Prognose- und Szenariorechnungen in SILVA verfügbare Baumartenspektrum zu erweitern. Des weiteren soll es die Flächenrepräsentativität der verwendeten Daten ermöglichen, regionale Wuchsbedingungen treffender als bisher in Simulationen abzubilden.

Derzeit werden die wachstumsbestimmenden Triebkräfte des Wachstums „Standort“ und „Klima“ über einen als „Standort-Leistungsmodell“ bezeichneten Ansatz modellhaft in SILVA eingesteuert. Die Eingangsgrößen in dieses Teilmodell sind hierbei nicht identisch mit den in der Praxis verfügbaren oder real verwendeten Größen. Ein Beispiel hierfür ist die ordinal abgestufte Eingabemöglichkeit der Wasserhaushaltsstufen in SILVA (Seifert et al., 1999), die nicht mit der teils ordinal, teils nominal skalierten Wasserhaushaltsstufenverschlüsselung der Standortkartierung in Bayern übereinstimmt. Auf diese grundlegende Problematik weist unter anderem von Gadow (2003) hin und fordert, dass „die in dieser Form vorhandenen Standortbeschreibungen (...) wenn möglich aufgrund ihrer großräumigen Verbreitung Eingang in forstliche Wachstumsmodelle finden“. Ein weiteres Nebenziel dieser Arbeit besteht daher darin, die standörtlich bedingten,

wachstumsbestimmenden Größen aus der Vielzahl der Inventurdaten zu extrahieren und diese Erkenntnisse dem Wuchsmodell als realitätskonforme Start- und Steuergrößen zur Auswahl zu übergeben. Die Verfolgung dieses Nebenzieles ermöglicht es gegebenenfalls zusätzlich, die in Bayern regional unterschiedliche Bedeutung gleicher Standortkartierungen in ein Relationsniveau zu bringen.

In einem eigenen Arbeitsschritt soll weiterhin die Haltung von Forsteinrichtungsdaten in Deutschland grundlegend verglichen werden. Die dabei gewonnenen Erkenntnisse bilden die Grundlage für die Erarbeitung einer Methodik, die eine Feinjustierung von SILVA auch für Regionen außerhalb Bayerns ermöglichen soll.

Letztlich zielt die Arbeit auf die Entwicklung einer Softwarelösung, die weitgehend autonom bzw. automatisiert die Informationen gemäß den vorstehend formulierten Teilzielen aus einer Forsteinrichtungsdatenquelle extrahiert und dem Benutzer (Experten) zur Überprüfung zur Verfügung stellt. Der Experte soll die Möglichkeit erhalten, diesen Hypothesensatz anzunehmen, zu verwerfen oder abzuändern. Bei Annahme durch den Benutzer werden die mit dem generierten Hypothesensatz verbundenen Parametersätze dem Wuchsmodell SILVA als Steuergrößen übergeben. Die zu erstellende Softwarelösung soll neben der Feinjustierungsfunktion für SILVA zudem einen eigenen Gebrauchs- und Informationswert für einen möglichst großen forstlichen Adressatenkreis aufweisen.

Zum Abschluss werden die Ziele noch einmal stichwortartig zusammenfassend dargestellt.

Oberziel: Semi-automatische Feinjustierung des Wuchsmodells SILVA mit Hilfe von Daten der betrieblichen Forstinventur in Bayern.

Nebenziele:

- Erhöhung der Prognosegenauigkeit für kurz- und mittelfristige Prognose- und Szenariorechnungen mit SILVA
- Erweiterung des in SILVA für Prognose- und Szenariorechnungen verfügbaren Baumartenspektrums
- bessere Berücksichtigung regionaler Wuchsgesetzmäßigkeiten
- Übernahme der in der forstlichen Praxis verwendeten Standortkodierung als Steuerungsgröße in SILVA
- Schaffung einer Methodik, die es ermöglicht, die regional unterschiedlichen Standorteinheiten über Inventurmessungen in ein Relationsniveau zu bringen
- Erweiterung des Kenntnisstandes zur Datenhaltung von Forstinventurdaten in Deutschland
- Schaffung einer Softwarelösung zur Erfüllung des Oberzieles

Vorgelegt wird demnach keine „klassische“ ertragskundliche Arbeit, die versucht mit bewährten oder modernen statistischen Verfahren, den Erkenntnisprozess der Waldwachstumsforschung durch hypothesengetriebene Analysen zu unterstützen. Vorliegende Arbeit bewegt sich vielmehr im Überlappungsbereich der Wissenschaftsbereiche „Waldwachstumskunde“, „Standorterkundung“, „Forsteinrichtung“ und „Informatik“. Sie versucht moderne, in anderen Wirtschaftszweigen bereits bewährte Ansätze aus den Bereichen des Data-Minings und des Maschinellen Lernens mit weitgehend anerkannten Gesetzmäßigkeiten der Waldwachstumslehre, der Standorterkundung und der Forsteinrichtung in Verbindung zu bringen und so neue Erkenntnisse aus einer bislang für derartige Zwecke kaum genutzten Datenquelle zu extrahieren. Abbildung 1 unterstützt noch einmal grafisch die Einordnung vorliegender Arbeit nach dem Untersuchungsproblem, dem Analyseziel, dem Grad der Nutzerinteraktion, der Autonomie des Ansatzes sowie der angewendeten Methodenklassen.

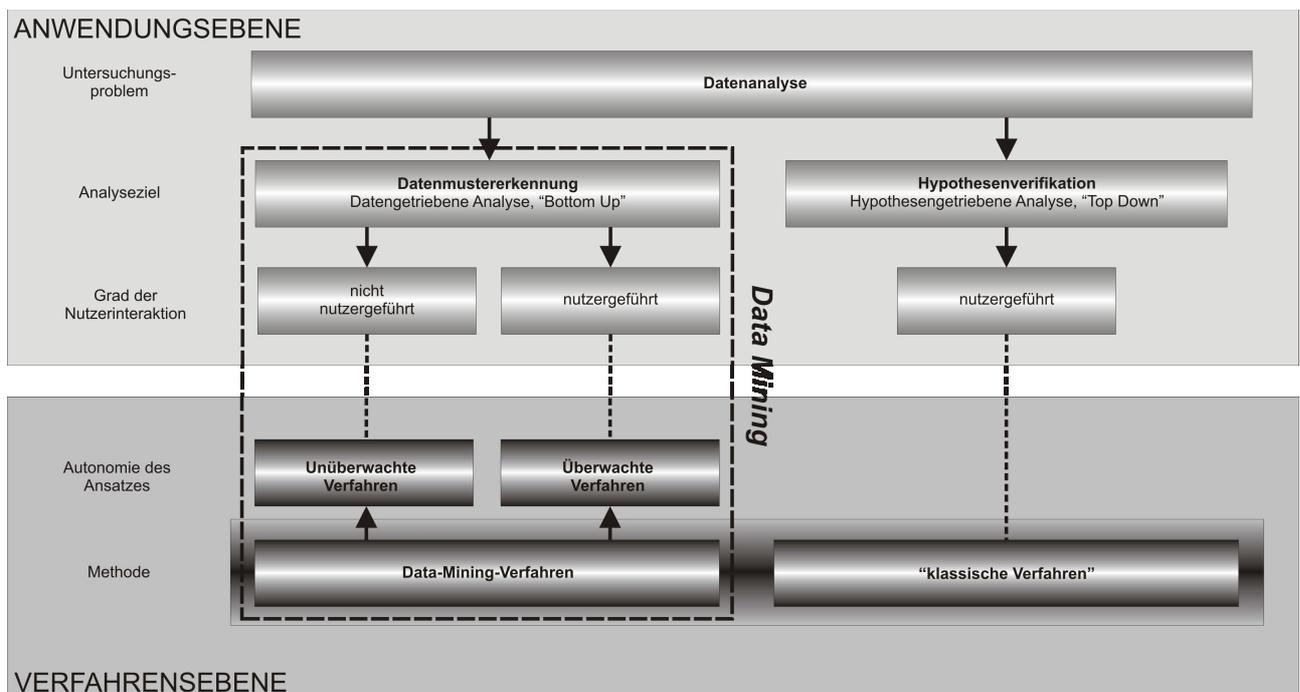


Abb. 1: Datenanalysezyklus (verändert nach Knobloch, 2000)

2. Stand des Wissens

2.1 Kenntnisstand zur Abhängigkeit des Waldwachstums von standörtlichen Faktoren

2.1.1 Kenntnisstand zum Einfluss primärer Grundfaktoren auf das Waldwachstum

Assmann (1961) behandelt die Hauptfaktoren des Assimilationsprozesses Licht, Temperatur, Wasser und CO₂-Gehalt der Luft und weist dabei auf die wichtigsten Ertragsfaktoren des Bodens, nämlich Wärme, Luftsauerstoff, Wasser und Nährelemente hin. Auf die Bedeutung des Bodens geht dieser Autor nur am Rande ein. Kramer (1988) nimmt diese Untergliederung auf und beschreibt den Einfluss der folgenden ökologischen Faktoren auf das Waldwachstum: Strahlung (einschließlich Licht), Temperatur, Wasser, Luft (insbesondere CO₂) sowie die physikalischen und chemischen Bodeneigenschaften und die mechanischen Faktoren. Wenk (1990) geht auf die primären Wachstumsfaktoren nur am Rande ein. Nachstehend wird nach Assmann (1961) und Kramer (1988) der Einfluss der letztgenannten primären Grundfaktoren beschrieben.

Die für das Pflanzenwachstum wichtige Strahlung entstammt im Wesentlichen der Strahlung der Sonne, die ein Strahlungsspektrum zwischen 200 nm und 3.000 nm Wellenlänge umfasst (Häckel, 1993). Die von der Sonne emittierte Strahlung ist verantwortlich für zahlreiche Lebensprozesse auf der Erde. Für das Wachstum von Pflanzen ist sie direkt und indirekt die maßgebende Triebkraft.

Für die CO₂-Assimilation benötigen die grünen Pflanzen die Energie, die der sichtbaren Strahlung entstammt, v. a. aus dem Bereich um 450 nm (blau) und 670 nm (orangerot). Strahlen aus den Wellenbereichen um 550 nm werden dagegen durch das Chlorophyll der Blätter nur gering absorbiert. Bei etwa gleich hoher Respiration erhöhen die grünen Pflanzen die Nettoassimilation mit steigender Lichtintensität. Die Assimilationsorgane mit größerer Schattentoleranz vermögen eine steigende Lichtintensität nicht im gleichen Maße auszunutzen wie solche mit geringerer Lichttoleranz.

Mit nachlassender Belichtung nimmt die Assimilation der Blätter ab, ihre Respiration bleibt allerdings gleich. Der Punkt, bei dem der Betrag beider Primärgrößen gleich ist, wird als Kompensationspunkt bezeichnet. Assmann (1961) hat ermittelt, dass dieser Kompensationspunkt bei den Lichtbaumarten höher liegt als bei den

Schattbaumarten. Außerdem ist er bei Lichtblättern erheblich höher als bei Schattenblättern.

Nach Kramer (1988) vertragen Jungpflanzen eine Beschattung anfänglich relativ gut, bei deutlicher Abnahme der Beleuchtungsstärke ist aber ein Nachlassen des Höhenwachstums zu beobachten. Nach Burschel und Huss (1997) wirkt sich der Lichtentzug umso stärker aus, je ungünstiger die Standortverhältnisse sind.

Auch die Temperatur hängt im Wesentlichen von der Sonnenstrahlung ab. Ferner wirken sich die Höhenlage und das Geländere Relief aus. Der Bodenzustand beeinflusst die Umwandlung der Strahlungsenergie in Wärme.

Waldbäume in Mitteleuropa reagieren unterschiedlich auf Temperatureinflüsse. Zu erklären ist dies dadurch, dass die Nettoassimilation der Blätter von deren Bruttoassimilation und Respiration bestimmt wird. Beide Faktoren weisen jedoch eine andere Abhängigkeit von den Klimafaktoren auf. Die von der Lichtintensität unterschiedliche Wirkung der Wärme kann nach Kramer (1988) damit erklärt werden, dass mit steigender Temperatur die Atmung sehr viel schneller zunimmt als die Bruttoassimilation. Eine durch Lichtmangel stark herabgesetzte Photosynthese führt demnach bei einer durch Temperaturanstieg geförderten Respiration zu einer Senkung der Nettoassimilation.

Zur Charakterisierung der Temperaturverhältnisse eines Wuchsbezirks oder eines Standortes verwendet man nach dem Arbeitskreis Standortkartierung (2003) vor allem die mittleren Lufttemperaturen während des Jahres und während der Vegetationszeit (Mai-September). Kramer (1988) weist in diesem Zusammenhang darauf hin, dass zur Untersuchung von Wachstumsvorgängen Temperaturschwellenwerte, z. B. die mittlere Dauer eines Tagesmittels der Lufttemperatur von mehr als 5° C oder von mehr als 10° C Bedeutung haben. Die Andauer eines Tagesmittels von über 10° C kann als Maß für die Länge der Vegetationszeit verwendet werden, da beim Überschreiten der 10° C-Schwelle einige Lebensvorgänge der Bäume, wie z. B. der Laubausschlag, sichtbar werden.

Der Wachstumsfaktor Wasser besteht aus den drei Elementen Luftfeuchtigkeit, Niederschlag und Bodenfeuchtigkeit. Waldbäume benötigen Wasser für den Aufbau von Wasserstoffverbindungen. Gleichzeitig ist das Wasser Träger des Transpirationsstromes, der von den Wurzeln zu den Blättern aufsteigt und die in ihm gelösten Nährstoffe an die verschiedenen Teile der Pflanze transportiert. Durch die Stomata verdunstet Wasser in die Atmosphäre. Da die Assimilationsleistung der Pflanzen davon abhängt, dass das CO₂ durch die offenen Stomata in das Innere des Blattes gelangen kann, bestehen enge Beziehungen zwischen der Transpiration und der Photosyntheseleistung. Bei ungenügender Wasserversorgung schließen sich die

Spaltöffnungen infolge des sinkenden Turgordrucks in den Schließzellen. Transpiration und Assimilation werden gleichzeitig deutlich gemindert.

Zur Kennzeichnung des Wasserhaushaltes verwendet die forstliche Standortaufnahme insbesondere die mittlere Niederschlagssumme im Jahr und in der Vegetationszeit (Arbeitskreis Standortkartierung, 2003).

Bäume entnehmen der Luft das für die Photosynthese benötigte CO₂ sowie den für die Atmung benötigten Sauerstoff. Für die Photosynthese und damit für die pflanzliche Stoffbildung ist das Kohlendioxid der entscheidende Faktor. Der Gehalt an CO₂ in der Luft betrug 2005 cirka 0,0379 % (IPCC, 2007). Bezugnehmend auf sehr frühe Untersuchungen weist Kramer (1988) darauf hin, dass sich dieser Gehalt in seiner natürlichen Intensität im Minimum seiner Wirkungsmöglichkeit befindet. Eine Erhöhung des Gehaltes kann zu einer erheblichen Steigerung der Assimilation und damit der stofflichen Produktion beitragen. Dieser Erkenntnis bedient man sich bei der Produktionsoptimierung von Gewächshauskulturen. Aktuelle Forschungsarbeiten, die im Auftrag des Intergovernmental Panel on Climate Change (IPCC) durchgeführt wurden, gehen aufgrund der Entwicklung der CO₂-Zuwachsraten der letzten Jahrzehnte davon aus, dass sich der CO₂-Gehalt der Atmosphäre in den nächsten Jahrzehnten deutlich erhöhen wird, was zu erheblichen negativen Konsequenzen für die Menschheit und die Umwelt führen wird (IPCC, 2007).

Die Ansprache des Bodens mit seinen physikalischen und chemischen Eigenschaften wird im Leitfaden des Arbeitskreis Standortkartierung (2003) beschrieben. Eine detaillierte Beschreibung der für das Waldwachstum wichtigen Böden findet sich bei Rehfuess (1990). Das Bodengefüge ist im Wesentlichen verantwortlich für die physikalischen Eigenschaften wie Wasserspeicherefähigkeit, Wasserdurchlässigkeit und Luftgehalt des Bodens. Ganz entscheidend wird nach Kramer (1988) das Wachstum der Bäume von der Bodenfeuchtigkeit bestimmt. Der Wasserhaushalt des Bodens wird insbesondere für wissenschaftliche Untersuchungen durch die nutzbare Wasserkapazität gekennzeichnet. Sie gibt die Wassermenge an, die mit mehr als 0,33 at und weniger als 15 at Saugspannung gebunden ist. Er weist in diesem Zusammenhang darauf hin, dass die in der Literatur häufig zu findende Feststellung, dass die Zuwachsleistung der Waldbäume mit zunehmender nutzbarer Wasserkapazität der Böden steigt, nicht verallgemeinert werden kann.

Auf die Bildung des Bodengefüges und auf die Nährstoffversorgung wirken sich Mächtigkeit und Art des Humus aus. Die unterschiedlichen Humusformen sind durch ihren Zersetzungsgrad und ihren Stickstoffgehalt (ausgedrückt durch das C:N-Verhältnis) gekennzeichnet.

Die zahlreichen chemischen Nährelemente im Waldboden lassen sich nach Kramer (1988) unterteilen in die nichtmineralischen Nährelemente C, O, H sowie in die mineralischen Hauptnährelemente N, P, S, K, Ca, Mg sowie in die mineralischen Spurenelemente Mo, Fe, Mn, Zn, Cu (metallisch) sowie Cl und B (nichtmetallisch). Ein Mangel an bestimmten Elementen führt zu Wuchsstörungen bis hin zum Absterben der Pflanzen. Nährelementmangel ist okular insbesondere an Verfärbungen der Blätter und Nadeln zu erkennen. Mängel in der Nährelementversorgung lassen sich z. T. durch Düngungsmaßnahmen ausgleichen.

Kramer fasst unter dem Oberbegriff mechanische Wachstumsfaktoren die Einwirkungen von Wind, Schnee, Eis und Feuer und Blitzschlag, aber auch biotische Einflüsse durch Insekten, Wild und den Menschen zusammen. Da diese Faktoren als natürliche Störfaktoren angesehen werden, werden sie nicht weiter erläutert.

2.1.2 Kenntnisstand zum Einfluss des Standortes auf das Waldwachstum

Das Wachstum der Waldbäume wird sehr stark durch den Standort geprägt. Nach Kramer (1988) umfasst der Begriff „natürlicher Standort“ die Gesamtheit der Umwelteinflüsse im Wald, die im Wesentlichen von Klima, Boden und Geländeausformung abhängen. Waldstandorte sind das Ergebnis einer Kombination einer Vielzahl primärer Wachstumsfaktoren, von denen die wichtigsten im vorhergehenden Abschnitt kurz angesprochen wurden.

In der forstlichen Praxis liegt meist keine Standortansprache vor, die wissenschaftlichen Anforderungen genügt. Die praktische Standorterkundung stützt sich vielmehr auf vorhandenes Kartenmaterial und Begehungen und einfache Standortansprachen im Gelände. Die bodenkundliche Ansprache erfolgt dabei durch Bohrung oder Spatenaushub gemäß einem im Vorfeld festgelegten, rasterförmigen Stichprobennetz. Die Grenzen der Kartiereinheiten (Standorteinheiten) werden bei der Kartierung im Gelände gezogen, und zwar nicht schematisch nach der Lage der Bohrpunkte, sondern unter strikter Beachtung des Reliefs und anderer wichtiger Merkmale wie z. B. Bodenvegetation, Humusform, Wuchsleistung der aufstockenden Bestände u. a.. Bei der Abgrenzung von Kartiereinheiten im Gelände gelten folgende Grundsätze: Die Grenzen sind umso genauer zu ziehen, je größer der ökologische Unterschied zwischen zwei Kartiereinheiten ist. Die Dichte des Bohrnetzes kann nach den geologisch-morphologischen und bodenkundlichen Verhältnissen variiert werden (Arbeitskreis Standortkartierung, 2003).

In Bayern existiert ein zweistufiges System der Standortkartierung. Auf der übergeordneten Stufe erfolgt eine Unterteilung nach regionalen ökologischen Einheiten zu Wuchsgebieten. Die nach überwiegend geographisch-physiographischen Gesichtspunkten ausgeschiedenen Wuchsgebiete werden weiter in die meist klimatisch abgegrenzten Wuchsbezirke unterteilt. Liegt für Teilbereiche eines Wuchsbezirkes ein von den mittleren Standortverhältnissen klimatisch oder petrographisch abweichendes spezielles Standortmosaik vor, so wird ein Teilwuchsbezirk ausgeschieden. Haben innerhalb eines Wuchsgebietes mehrere Wuchsbezirke einen ähnlichen Landschaftscharakter, so werden sie in Landschaftsgruppen zusammengefasst und eventuell verbleibende eigenständige Wuchsbezirke als Einzellandschaft ausgeschieden. Regionale statistische Flächengrundlage und Bezugsraum für die Ausscheidung der ökologischen Einheiten ist jedoch stets der Wuchsbezirk.

Auf der zweiten, untergeordneten Stufe erfolgt die Ausscheidung der ökologischen Einheiten, die in Bayern als Standorteinheiten bezeichnet werden. Um bei der Kartierung regionalen Besonderheiten bestmöglich nachgehen zu können, wurde für die Ausscheidung und Benennung der Standorteinheiten bundesweit auf allgemein verbindliche Weisungen verzichtet. Im Regelfall werden jedoch die Merkmale Substrattyp, Geländewasserhaushalt, Geländemorphologie, Vegetation und Standörtliche Besonderheiten herangezogen (Arbeitskreis Standortkartierung, 2003). Die einzelnen Standorteinheiten werden nach einem dreistelligen Ziffernsystem verschlüsselt. Dadurch wird dem Benutzer eine rasche Orientierung auch über Wuchsbezirksgrenzen hinweg erleichtert und die Grundlage für eine statistische Erfassung und Auswertung standortkundlicher Daten geschaffen. In dem System beschreibt die erste Ziffer die Substratgruppe. Die zweite Ziffer bringt wesentliche Trophiemerkmale (z. B. „humusreich“) und standörtliche Besonderheiten (z. B. „Hanglage“) zum Ausdruck. Die dritte Ziffer definiert den Wasserhaushalt (Arbeitskreis Standortkartierung, 2003). Gemeinsam mit den übergeordneten landschaftlichen Einheiten Wuchsbezirk und Wuchsgebiet dienen die Standorteinheiten als Entscheidungshilfe für die lang- und mittelfristige Planung in Forsteinrichtungsverfahren sowie im praktischen Waldbau. Die Standortkartierung bietet die Voraussetzung, die standortsabhängigen Wachstumsvorgänge der Waldbäume zu erklären.

Bei der Betrachtung wissenschaftlicher Arbeiten zum Zusammenhang zwischen Waldwachstum und Standort muss primär unterschieden werden, ob die Arbeiten Aussagen zu Höhenwachstumsgängen oder zu Durchmesserzuwachsängen machen. Des Weiteren muss beachtet werden, ob die Aussagen auf Einzelbaum- oder Bestandesebene getroffen werden und ob kleinstandörtliche oder großstandörtliche Faktoren einen Einfluss auf die Wachstumsvorgänge ausüben.

Die drei Standardwerke der Waldwachstumskunde von Assmann (1961), Kramer (1988) und Wenk et al. (1990) stimmen darin überein, dass der Höhenwachstumsgang innerhalb der gleichen Baumart in Abhängigkeit vom Standort sehr unterschiedlich sein kann. Je besser der Standort ist, umso früher kulminiert der Höhenzuwachs und umso schneller sinkt er nach Überschreiten des Kulminationszeitpunktes. Mit zunehmendem Alter vermindern sich daher deutlich die Differenzen des jährlichen Zuwachses von vergleichbaren Bäumen unterschiedlicher Höhenbonität. Die überlegene Höhenwuchsleistung von Bäumen auf guten Standorten ist somit v. a. durch das sehr hohe Jugendwachstum bedingt.

Weiterhin stimmen die drei Autoren darin überein, dass sich in gleichaltrigen Reinbeständen der Wuchsraum des einzelnen Baumes nur wenig auf sein Höhenwachstum auswirkt. Pflanzverbandsversuche weisen allerdings darauf hin, dass dichte Kulturen und Naturverjüngungen das Höhenwachstum der Bäume hemmen können. Kramer (1988) berichtet hierzu, dass derartige Wuchshemmungen besonders auf armen und trockenen Standorten auftreten. Wenk et al. (1990) berichten, dass nach Beendigung der Jugendphase Standraumerweiterungen durch Durchforstungen das Höhenwachstum nicht mehr positiv beeinflussen können. Der Höhenzuwachs nimmt nach diesem Ansatz die größten Werte bei leichten Konkurrenzverhältnissen im Bestand an und fällt erst wieder bei sehr starken Konkurrenzverhältnissen. Der Durchmesserzuwachs hingegen fällt kontinuierlich mit zunehmender Konkurrenzsituation der Einzelbäume.

Auch wird von Assmann, Kramer und Wenk einheitlich berichtet, dass in gleichaltrigen Reinbeständen sowie in Mischbeständen der Höhenwachstumsgang der mitherrschenden und beherrschten Bäume in Folge Lichtmangels stark beeinträchtigt wird. Durch lange Überschirmungsphasen erfolgt Wachstum und Entwicklung der unterdrückten Baumart zeitverzögert. Dies hat allerdings keine Auswirkungen auf die erreichbare Maximalhöhe. Abhängig von verschiedenen inneren und äußeren Faktoren sind die Bäume in der Lage, nach mehr oder weniger langen Überschirmungsphasen eine andere soziale Stellung im Sozialgefüge einzunehmen. Dieses Phänomen wird als „Umsetzen“ bezeichnet. Assmann (1961) schreibt hierzu, dass sich (in Mischbeständen) die Höhenwuchsleistung und die spezifische Höhenentwicklung umso mehr zugunsten einer Baumart verändert, je mehr die Standorteigenschaften ihren individuellen Bedürfnissen entsprechen oder je mehr sie befähigt ist, sich diesen anzupassen.

Bei allen drei Autoren herrscht weiterhin Einigkeit darüber, dass die Durchmesserentwicklung weit mehr als die Höhenentwicklung durch bewirtschaftende Maßnahmen in Waldbeständen gelenkt werden kann. Wenk et al. (1990) benennen hierzu als die drei wesentlichen Faktoren und Maßnahmen die Ausgangsstammzahlen, die Stammzahlregulierung in Jungbeständen sowie die

Ausnutzung des Lichtungszuwachses. Alle benannten waldwachstumskundlichen Standardwerke stimmen in dem Punkt überein, dass bei gleicher bestandesgeschichtlicher Entwicklung das Durchmesserwachstum auf guten Standorten größer ist als auf schlechten Standorten.

In den benannten forstlichen Standardwerken finden sich eine Vielzahl von Arbeiten, die sich in den letzten Jahrzehnten sehr intensiv mit dem Wachstumsgang der einzelnen Baumarten in Abhängigkeit von natürlichen Standortfaktoren befasst haben. Nahezu allen diesen Arbeiten ist gemein, dass sie sich auf Auswertungen von Versuchsflächendaten beziehen. Eine gute Übersicht über derartige Arbeiten findet sich bei Kramer (1988). Aus diesem Werk werden nachfolgend zahlreiche Ergebnisse dieser Arbeiten nach Baumarten getrennt kurz dargestellt.

Für die Baumart *Fichte* stellen die Arbeiten im Wesentlichen übereinstimmend fest, dass Leistung und Entwicklung von Fichtenbeständen v. a. vom lokalen Wasser- und Wärmehaushalt, daneben auch vom Nährstoffreichtum des Bodens insbesondere vom Stickstoffgehalt bestimmt werden. Bei ausreichenden Niederschlägen sinkt die Höhenwuchsleistung mit abnehmender Jahresmitteltemperatur. Mangelnde Niederschläge können z. T. durch den Grundwasserstand sowie durch die Wasserhaltekapazität und den Nährstoffgehalt des Bodens ausgeglichen werden. Moosmayer und Schöpfer (1972) haben hierzu in einer Untersuchung in Baden-Württemberg einen engen Zusammenhang zwischen Regenfaktor im Sommerhalbjahr und Veränderung der Ertragsklassen ermittelt. „So ist es verständlich, dass Fichtenbestände auf sehr ähnlichen Böden, aber bei unterschiedlichem Lokalklima verschiedene Bonitätsentwicklungen haben“.

Maßgebliche historische Arbeiten zur Baumart *Kiefer* in Bayern stammen von Schmidt. In einer Arbeit aus dem Jahr 1971 ermittelte er im Vergleich zur Ertragstafel Wiedemann einen deutlichen Bonitätsanstieg. Dieser kann auch durch den verstärkten Stickstoffeintrag der letzten Jahrzehnte verursacht worden sein. Das Ausmaß dieses Anstiegs ist jedoch im Einzelfall sehr unterschiedlich und hängt wesentlich von der Wasserversorgung des Standortes ab. Für das Bundesland Baden-Württemberg existiert eine Untersuchung von Laatsch (1963) zu dieser Thematik. Nach dieser Arbeit erwiesen sich der Nährstoffgehalt und die Wasserversorgung des Bodens als die am stärksten begrenzenden Ertragsfaktoren für die Wachstumsleistung der Kiefer.

Nach Kramer (1988) hängt auch bei der Baumart *Douglasie* der Höhenwachstumsgang sehr stark vom Wasser- und Wärmehaushalt des Bodens ab. Das Jugendwachstum dieser Baumart wird nach der gleichen Quelle v. a. durch den

Wärmehaushalt, das Wachstum im späteren Alter durch eine gute Wasserversorgung gefördert.

Zur Untersuchung der Standortabhängigkeit des Wachstums bei *Buche* hat Kennel (1973) die Buchendurchforstungsversuche des langfristigen forstlichen Versuchswesens in Bayern untersucht. Auch in dieser Arbeit konnte eine standortabhängige Änderung der Höhenbonität festgestellt werden, wenngleich die Änderungen nicht so einheitlich und übereinstimmend waren wie z. B. für die Baumart Fichte. Röhe (1985) hat eine hohe Wuchsleistung der Buche in Baden-Württemberg bei einer Vegetationszeit von mindestens 150 Tagen festgestellt, wobei als Abgrenzungskriterium eine Tagesdurchschnittstemperatur von mindestens 10° C angenommen wurde. Bei höherer Wasserversorgung bzw. einer längeren Vegetationszeit hängt ihr Wachstum wesentlich von der Höhe der Jahresniederschläge ab. Zuwachshemmende Effekte wurden bereits ab einer Niederschlagsmenge unter 800 mm pro Jahr festgestellt. Zudem wird nach der selben Arbeit der Zuwachs der Buche stärker durch die Nährstoffverhältnisse als bei der Baumart Fichte beeinflusst. Moosmayer (1987) weist einen deutlichen regionalklimatischen Einfluss auf das Wachstum der Buche nach. Bei ausreichender Wasserversorgung wirkt sich nach dieser Arbeit besonders die Länge der Vegetationszeit positiv auf die Wuchsleistung aus.

Die Baumart *Eiche* zeigt dagegen unterschiedliche, standortbedingte Höhenentwicklungen. Bei guter Wärmeversorgung ist für ihre Bonitätsentwicklung der Wasserhaushalt des Standortes entscheidend. So konnte z. B. Kramer für den Nordspessart nachweisen, dass die Eiche im Laufe des Bestandeslebens auf mäßig frischem bis frischem Standort einen Bonitätsanstieg, auf trockenem Standort jedoch ein Gleichbleiben oder Absinken der Ertragsklasse zeigt (Kramer, 1988). Eine aktuellere Arbeit für Furniereichen stammt von Göttlein (1994). Er hat regressions- und varianzanalytische Untersuchungen für Furniereichen aus der Oberschicht diverser Spessartbestände durchgeführt. Er kommt zu dem Ergebnis, dass für die erreichte Baumhöhe der Standort nur geringe Bedeutung hat, während die Topographie und hier insbesondere die Hangrichtung einen wesentlichen Einflussfaktor darstellt. Deutlich hat die Baumhöhe des untersuchten Bestandesteilkollektivs auch auf zunehmende Bodenfrische reagiert, wobei jedoch wechselfrische Standorte eine Sonderstellung einnahmen.

Wenk et al. (1990) weisen darauf hin, dass eine Verallgemeinerung dieser wertvollen Ergebnisse dadurch erschwert wird, weil in den einzelnen Ländern unterschiedliche Methoden der Standortkartierung angewendet werden. Des Weiteren muss kritisch gesehen werden, dass die Standorteinheiten der Standortkartierung nicht

unveränderlich sind. Ältere Arbeiten gehen mehr oder weniger von einer Konstanz der Standortbedingungen aus (vgl. z. B. Assmann, 1961). Aufgrund zahlreicher aktuellerer waldwachstumskundlicher (vgl. z. B. Pretzsch, 1985; Spelsberg, 1994, Röhle, 1995), waldernährungskundlicher (vgl. z. B. Prietzel und Kölling, 1998, v. Mammen et al., 2003) und bodenkundlicher Arbeiten (vgl. z. B. Rehfuss, 1990, Prietzel et al., 1997) hat sich allerdings gezeigt, dass sich forstliche Standorte, v. a. aufgrund atmosphärischer Einträge in die Waldböden, verändern. Daher gelten Standorteinheiten mittlerweile selbst nicht als unveränderlich. Diese neuen Erkenntnisse haben mittlerweile zu Wiederholungen der Standorterkundung geführt, wobei sich hierbei Standortgrenzen und Standorteinheiten ändern können. Nach Wenk et al. (1990) wirkt sich auf die Aufklärung der Zusammenhänge zwischen Standort und Ertrag darüber hinaus erschwerend aus, dass auf grundsätzlich unterschiedlichen Standorten der Wachstumsverlauf ähnlich sein kann.

Aufgrund der sich ändernden Standortverhältnisse wurden für Bayern in den letzten Jahren zahlreiche aktuelle Forschungsarbeiten zu den Wachstumstrends bei den Baumarten Fichte (Pretzsch und Utschig, 2000), Kiefer (Küsters, 2001), Buche (Harrer, 2004) und Eiche (Herling, 2005) angefertigt.

Mehrere Autoren haben in der Vergangenheit versucht, den Zusammenhang zwischen Höhenwachstum und Standortvariablen zu quantifizieren. Stellvertretend hierfür wird die Arbeit von Moosmayer und Schöpfer (1972) angeführt.

Wegen der außerordentlichen Schwierigkeit, das Wachstum als Funktion von Standortfaktoren zu beschreiben, benutzt man seit über 100 Jahren die Höhe als indirekten Standortweiser. Gewöhnlich wird die Mittelhöhe oder Oberhöhe im Alter von 100 Jahren zur Bonitierung des Standortes benutzt. Aufgrund der sich ändernden Wachstumsbedingungen verliert diese Weisergröße in der forstlichen Praxis allerdings zunehmend an Bedeutung.

2.2 Funktionsprinzip des Wachstumsmodells SILVA

Mit SILVA wurde am Münchner Lehrstuhl für Waldwachstumskunde ein einzelbaumorientiertes, distanzabhängiges und standortsensitives Managementwachstumsmodell entwickelt, das für den Einsatz in der forstlichen Praxis sowie als Lehr- und Ausbildungsinstrument in Forstpraxis und Wissenschaft als auch als Forschungsinstrument konzipiert ist. Eine detaillierte Beschreibung zu Konzeption und Konstruktion findet sich bei Pretzsch und Kahn (1998).

Wegen der grundsätzlichen Bedeutung für das Verständnis dieser Arbeit wird zunächst das Funktionsprinzip anhand des Prognosealgorithmus dargestellt. Anschließend werden wichtige Charakteristika des Wachstumsmodells SILVA erläutert.

Das Wachstumsmodell SILVA geht in der vorliegenden Version 2.2 von nur wenigen Start- und Steuergrößen aus. Diese charakterisieren die Ausgangssituation eines Bestandes sowie seine Standortbedingungen. Über ein standortabhängiges Funktionensystem wird die Bestandesdynamik in Fünf-Jahres-Zyklen von der Bestandesbegründung bis zum Generationswechsel nachgebildet (Pretzsch, 2001).

Der SILVA 2.2 zugrundeliegende Prognosealgorithmus ist eine Abbildung 2 dargestellt.

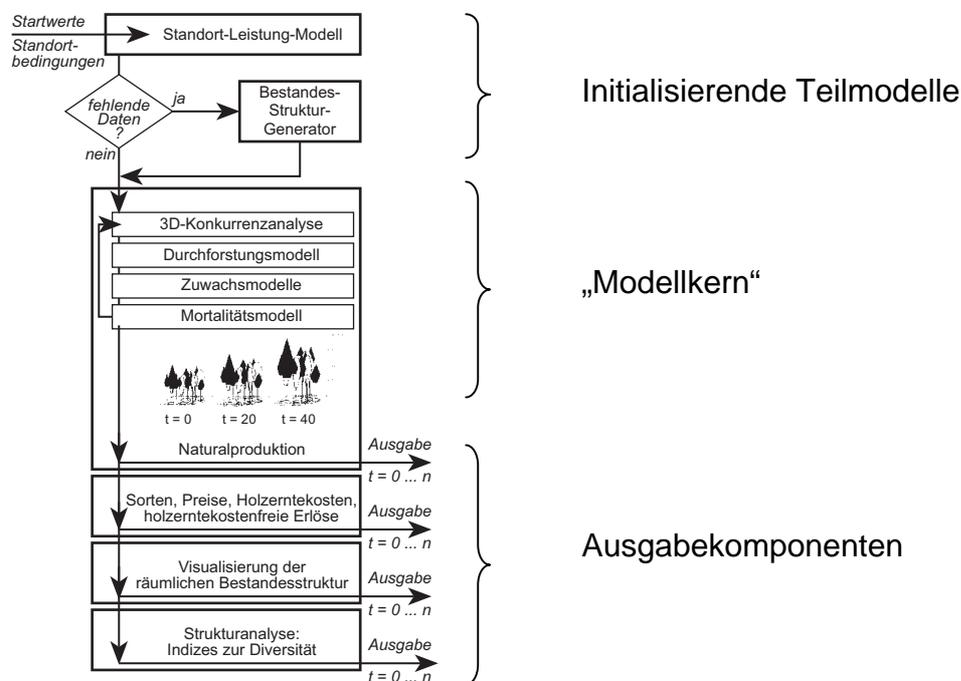


Abb. 2: Prognosealgorithmus des Wachstumsmodells SILVA 2.2 (verändert nach Pretzsch, 2001)

Als Startgrößen werden zu Beginn eines Prognoselaufes die Dimension und die Position der Einzelbäume eingelesen. Als Steuergrößen gehen in das sogenannte Standort-Leistungsmodell die wachstumsbestimmenden Standortparameter ein. Allerdings liegen nur in seltenen Fällen die Eingangsgrößen vollständig vor. Aus diesem Grund wurde als Teilmodell ein Bestandes-Struktur-Generator (Pretzsch, 1993) entwickelt und implementiert, der fehlende Größen nach dem statistischen Prinzip wirklichkeitsnah ergänzt.

Sind die Ausgangsgrößen vorhanden, so beginnt im „Kern des Modells“ die eigentliche Wachstumsprognose, die sich aus einer vom Benutzer anzugebenden Anzahl von Fünf-Jahres-Zyklen aufbaut. In einem ersten Schritt wird ein räumliches Bestandesmodell aufgebaut und die individuelle Wuchskonstellation jedes Baumes über sog. Konkurrenzindizes erfasst. Hierbei kommen die Maßzahlen KKL als Indikatorgröße für die Kronenkonkurrenz, NDIST zur Beschreibung von Symmetrie und Asymmetrie der Konkurrenzwirkung und KMA für den Mischungsanteil in der Nachbarschaft von Einzelbäumen zum Einsatz. Wie mehrere Autoren (vgl. z. B. Filla, 1981; Pretzsch, 2001) betonen, stellt die modellhafte numerische Charakterisierung der Konkurrenzsituation der Einzelbäume im Bestand einen wesentlichen Teilschritt im gesamten Prognoseprozess dar. Im darauffolgenden Prognoseschritt wird für jedes Bestandesindividuum geprüft, ob es unter dem vom Benutzer vorgegebenen Durchforstungskonzept im Bestand verbleibt oder durch simulierte Durchforstungseingriffe entnommen wird. Im nächsten Schritt erfolgt die Zuwachsallokation auf die verbliebenen Bestandeselemente mit Hilfe baumindividueller Zuwachsmodele. Die Steuerung übernehmen hierbei wiederum die zuvor berechneten Konkurrenzmaßgrößen. Im vierten Schritt entscheidet das Mortalitätsmodell, welche Bäume den zurückliegenden fünfjährigen Zeitraum aufgrund von Konkurrenzeinwirkung nicht überleben. Optional kann ein weiteres Teilmodell, welches die Verjüngungsentwicklung prognostiziert, für jeden 5-Jahreszyklus aktiviert werden.

Am Ende eines SILVA-Prognoselaufes stellt das Wachstumsmodell thematisch unterschiedliche Ausgabegrößen zur Verfügung.

Bei dem Wachstumsmodell SILVA handelt es sich um ein **einzelbaumbasiertes**, **distanzabhängiges** und **standortsensitives** Wachstumsmodell, welches die Zuwachsschätzung nach einem als **Potenzial-Modifizierungs**-Prinzip bezeichneten Verfahren vornimmt. Diese charakteristischen Merkmale werden nachstehend in veränderter Reihenfolge kurz erläutert.

Nach Pretzsch (2002) haben sich sowohl die direkte regressionsanalytische Schätzung des Zuwachses als ein Ansatz als auch die indirekte Zuwachssteuerung nach der Methode der Potenzial-Modifizierung bewährt. SILVA arbeitet nach dem zweiten Ansatz, der erstmalig in einer Arbeit von Newnham (1964) vorgestellt wurde. Bei der Methode der Potenzial-Modifizierung, geht die Zuwachssteuerung von einem potenziellen Zuwachs z_{pot} aus. Der im konkreten Fall erwartete Zuwachs z_{erw} des Einzelbaumes ergibt sich durch Multiplikation des potenziellen Zuwachses, der ohne Konkurrenzeinfluss zu erwarten wäre, mit einem Reduktionsfaktor R , der einen Wert zwischen 0 und 1,0 einnimmt. Der Reduktionsfaktor ist in SILVA konzeptionell funktional abhängig von Konkurrenz, Vitalität und Standort (vgl. Abb. 3), die

wesentliche Einflussgröße stellt allerdings die Konkurrenz dar, die in SILVA quantitativ durch die drei Maßgrößen KKL, NDIST und KMA gefasst wird. Aufgrund der Steuerung sowohl des Höhenzuwachses als auch des Durchmesserzuwachses nach der Potenzial-Modifizierungs-Methode kommt der Ermittlung und Definition des Durchmesser- und Höhenpotenzials erhebliche Bedeutung zu, weil damit die maximal erreichbaren Werte feststehen (Hasenauer, 1999). Nach Hasenauer (1997) wird der potenzielle bzw. maximal erreichbare Durchmesserzuwachs bei Solitären erreicht. Da nur lückenhaft Daten über das Solitärwachstum von Waldbäumen vorliegen, kann eine Abschätzung potenzieller Höhenzuwächse bzw. Baumhöhen entweder aus dem bonitätsabhängigen Oberhöhenverlauf des Bestandes (Hasenauer, 1999) oder mit Hilfe eines modellhaft hergeleiteten Standort-Leistungspotenzials (vgl. z. B. Kahn, 1994) erfolgen. Weitergehende Überlegungen zur Ableitung potenzieller Zuwächse stellen Reed et al. (2001) dar. In dieser Arbeit wird die Ableitung der potenziellen Durchmesser- und Höhenentwicklung für Eucalyptus globulus-Plantagenwälder in Portugal vorgenommen bzw. wird über eine destruktive Ermittlung der Potenziale in Pinus resinosa-Beständen in den USA berichtet. Daher sind die Ergebnisse dieser Arbeit nur bedingt für den weiteren Fortgang der vorliegenden Arbeit verwertbar.

Einzelbaumbasierte Wachstumsmodelle lösen einen Bestand mosaikartig in seine Einzelbäume auf und bilden ihr Miteinander im Bestand als räumlich-zeitliches System nach. Hierdurch entsteht eine Modellflexibilität, welche die Nachbildung verschiedenster Mischungs- und Strukturformen, Pflegeregime und Verjüngungsverfahren erlaubt. Während der eingangs beschriebenen Prognoseläufe werden für jeden Baum zu jedem Prognosezeitpunkt Einzelbauminformationen mitgeführt. Die Verwaltung der Einzelbäume auf Bestandesebene erfolgt über Baumlisten. Durch die Aggregation der Informationen für die einzelnen Bestandesglieder können Ausgabegrößen räumlich und zeitlich auf verschiedenen Auflösungsebenen erzeugt werden (Pretzsch, 2002).

Distanzabhängige Wachstumsmodelle berücksichtigen die räumliche Position der einzelnen Bäume im Bestand, d. h. für jeden Baum werden neben seinen Dimensionsgrößen seine horizontalen (und vertikalen) Koordinaten für die Zuwachsschätzung benötigt. Ein wesentlicher Vorteil dieses Ansatzes liegt darin, dass dadurch das Zuwachsverhalten von Mischbaumarten in gemischten Beständen, welches maßgeblich von der räumlichen Struktur ihrer Vergesellschaftung abhängt, explizit berücksichtigt werden kann.

Neben der Einzelbaumbasiertheit und der Distanzabhängigkeit besteht ein weiteres Charakteristikum von SILVA in seiner Standortsensitivität. In SILVA wird diese Eigenschaft durch ein Modell zur Schätzung des Standort-Leistungsbezuges erreicht,

welches in seinen Grundzügen auf dem von Kahn (1994) entwickelten Verfahren zur Schätzung der Oberhöhe in Abhängigkeit vom Standort beruht. Die Bestandeshöhe bzw. die daraus abgeleitete potenzielle Einzelbaumhöhe ist in SILVA 2.2 die wichtigste Leitgröße für die Modellierung des Waldwachstums. Die Höhe über dem Alter stellt somit die Maßgröße für die Leistungsfähigkeit des Standortes dar. Neben dem Höhenzuwachs werden Durchmesserzuwachs sowie die Einzelbaummortalität in Abhängigkeit vom Standort modelliert. Pretzsch und Kahn (1998) führen aus, dass es mit diesem Standort-Leistungsmodell möglich wird, die baumartenspezifische Bonität eines Standortes, unabhängig von eventuell verfügbaren Altersangaben und Mischungsstrukturen, einzuschätzen. Um die Alters-Höhenentwicklung standortabhängig zu modellieren, wurde von Kahn die dreiparametrische Wachstumsfunktion nach Chapman-Richards (Formel 1) gewählt, da diese bereits mehrfach mit gutem Erfolg auf Daten des langfristigen Versuchsflächennetzes in Bayern angewendet werden konnte (Pretzsch, 2001).

$$h_{100} = A * (1 - e^{-k*t})^P \quad (\text{Formel 1})$$

h_{100}	Bestandesoberhöhe
A	Asymptote
k, P	Steigungs- und Formparameter
t	Bestandesalter in Jahren

Um für einen gegebenen Standort die Alters-Höhenentwicklung zu schätzen, werden nach dem Standort-Leistungsmodell die neun Standortvariablen „Nährstoffversorgung des Bodens“, „NO₂-Gehalt der Luft“, „CO₂-Gehalt der Luft“, „Länge der Vegetationszeit (Tage über 10°C)“, „Jahrestemperaturamplitude in °C“, „mittlere Temperatur in der Vegetationszeit in °C“, „Ariditätsindex nach DE MARTONNE“, „Niederschlagssumme in der Vegetationszeit“ sowie „Bodenfrische“ zu den drei komplexen ökologischen Faktoren „Nährstoffversorgung“, „Wärmeversorgung“ und „Wasserversorgung“ zusammengefasst. In Anhalt an die Formeln 2 und 3 wird die Asymptote A der Wachstumskurve sowie der Zeitpunkt der Zuwachskulmination t_{kulm} geschätzt.

$$A = A_0 * A_1 * \left(\prod_{j=1}^3 KF_j \right)^{1-\gamma_1} * \left(1 - \left(\prod_{j=1}^3 (1 - KF_j) \right)^{\gamma_1} \right) \quad (\text{Formel 2})$$

$$t_{kulm} = t_0 + t_1 * \left(\prod_{j=1}^3 KF_j \right)^{1-\gamma_2} * \left(1 - \prod_{j=1}^3 (1 - KF_j) \right)^{\gamma_2} \quad (\text{Formel 3})$$

Es bedeuten

A	Asymptote
A_0	minimale Asymptote
A_1	maximale Asymptote
t_{kulm}	Bestandesalter, in dem der Höhenzuwachs kulminiert
t_0	Minimalwert von t_{kulm}
t_1	Maximalwert von t_{kulm}
KF	Komplexer ökologischer Faktor
$\gamma_1.. \gamma_5$	Aggregationsoperatoren
j	Laufvariable

Die Aggregationsoperatoren wurden aus Wachstums- und Standortdaten der langfristigen Versuchsflächen regressionsanalytisch geschätzt. Der Parameter A der Chapman-Richards-Funktion ist damit bekannt. Zur Bestimmung der Parameter k und P wird auf eine Hilfsbeziehung zwischen dem Zeitpunkt der Zuwachskulmination t_{kulm} und der Baumhöhe zum Zeitpunkt der Zuwachskulmination zurückgegriffen.

$$h_{t_{kulm}} = B * \left(1 - e^{-c * t_{kulm}}\right) = A * \left(1 - \frac{1}{P}\right)^P \quad (\text{Formel 4})$$

Die Parameter B und c werden ebenfalls aus Versuchsflächendaten geschätzt. Nach Auflösen der überführten Gleichung nach P ist der Kurvenparameter P bekannt. Dieser dient neben t_{kulm} zur Bestimmung des Steigungsparameters k.

$$k = - \frac{- \ln \left(\frac{1}{P} \right)}{t_{kulm}} \quad (\text{Formel 5})$$

Um von der Bestandesoberhöhe h_{100} zur potenziellen Höhenkurve h_{pot} für Einzelbäume zu gelangen, werden ausgehend von Ergebnissen des langfristigen ertragskundlichen Versuchswesens in Bayern die Bestandesoberhöhenkurven der Baumarten durch Multiplikation mit baumartenspezifischen Faktoren angehoben.

Die so modellhaft im initialisierenden „Standort-Leistungsmodell“ ermittelten potenziellen Wachstumskurven (vgl. Abb. 3) stellen die Grundlage für die Kalkulation der erwarteten Einzelbaumzuwächse in den Zuwachsmodellen im „Modellkern“ (vgl. Abb. 2) dar.

Strukturell ist das Durchmesserzuwachsmodell in SILVA nahezu identisch mit dem Höhenzuwachsmodell. Während jedoch beim Höhenzuwachsmodell der Standort bereits im Höhenpotenzialwert berücksichtigt wird, kommt beim Durchmesserzuwachsmodell ein standortabhängiger Modifikator dazu. Nachstehend sind die Formelgleichungen für die beiden Zuwachsmodelle nach Pretzsch und Kahn (1998) aufgeführt.

Durchmesserzuwachsmodell (gesteuert über den Grundflächenzuwachs)

$$zg_{erw} = zg_{pot} * a_6 * WSto * WKrone * WKonkurrenz \quad (\text{Formel 6})$$

Es bedeuten für das Durchmesserzuwachsmodell:

zg_{erw}	prognostizierter Grundflächenzuwachs
zg_{pot}	Grundflächenzuwachspotenzial
$WSto$	standortabhängiger Zuwachsmodifikator
$WKrone$	Wirkung der Kronenmantelfläche auf den Zuwachs
$WKonkurrenz$	integrierte Wirkung der Konkurrenz auf den Zuwachs
a_6	baumartenspezifischer Parameter

Höhenzuwachsmodell

$$zh_{erw} = zh_{pot} * a_6 * WKrone * WKonkurrenz \quad (\text{Formel 7})$$

Es bedeuten für das Höhenzuwachsmodell:

zh_{erw}	prognostizierter, erwarteter Höhenzuwachs
zh_{pot}	standortabhängiges Höhenzuwachspotenzial
$WKrone$	Wirkung der Kronenmantelfläche auf den Zuwachs
$WKonkurrenz$	integrierte Wirkung der Konkurrenz auf den Zuwachs
a_6	baumartenspezifischer Parameter

Abbildung 3 verdeutlicht noch einmal das grundlegende Funktionsprinzip der Herleitung des potenziellen Höhenzuwachses zh_{pot} im Wuchsmodell SILVA aus der auf den Werten der Parameter A, k und P beruhenden Wachstumskurve nach Chapman-Richards. Im weiteren Verlauf dieser Arbeit werden diese Parameterwerte gelegentlich verkürzend als AkP-Werte bezeichnet.

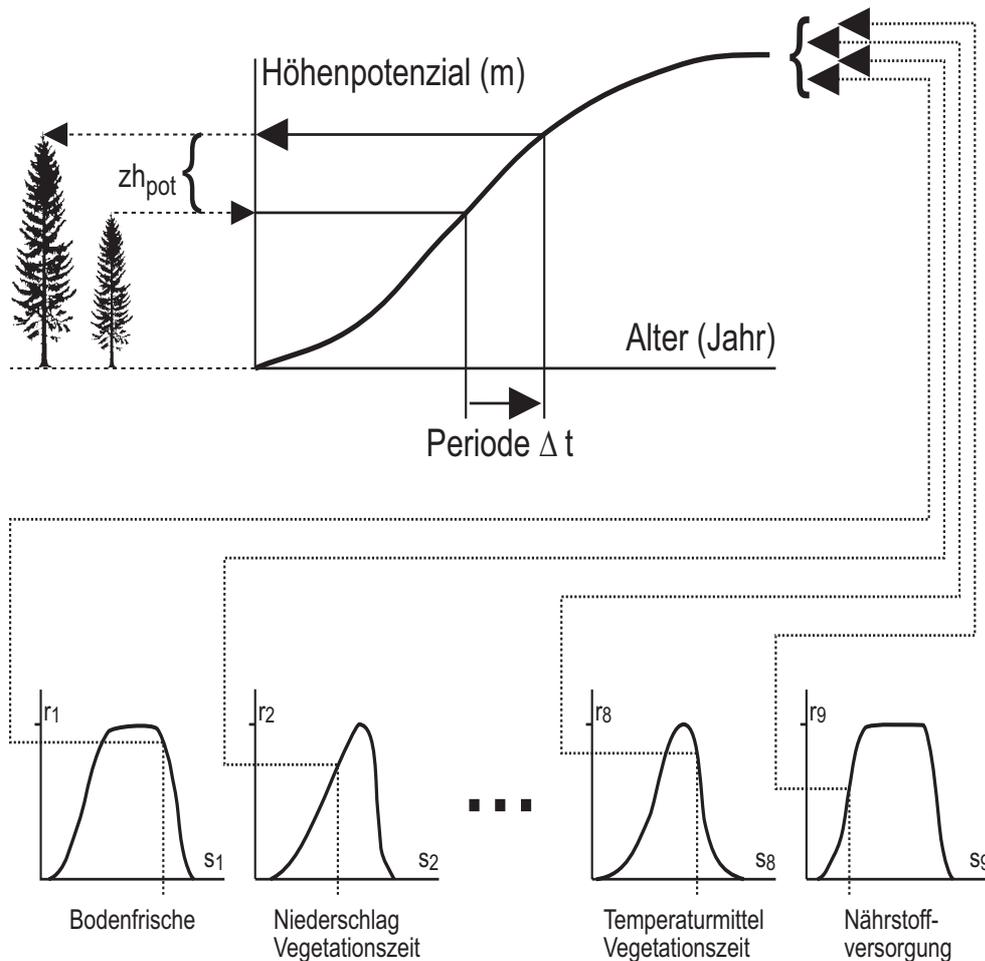


Abb. 3: Herleitung des potenziellen Höhenzuwachses zh_{pot} im Wuchsmodell SILVA (verändert nach Pretzsch, 2001) aus der potenziellen Wachstumskurve, die mit Hilfe der Koeffizienten der Parameter A, k und P beschrieben werden kann.

2.3 Data-Mining

2.3.1 Überblick über Data-Mining-Verfahren und deren Anwendung

Data-Mining ist ein Sammelbegriff für eine Vielzahl von teilweise sehr unterschiedlichen Ansätzen aus verschiedenen Fachdisziplinen. Bevor eine Definition des Begriffspaares Data-Mining sowie ein Überblick über die Verfahren folgt, wird nachfolgend ein Anwendungsbeispiel aus der forstlichen Praxis von Kollegen aus Slovenien (Dzeroski et al., 2006) vorgestellt.

Slovenien ist ein waldreiches Land mit einem Waldflächenanteil von 57,4 %. Aufgrund aktueller Fragestellungen der Gesellschaft im Zusammenhang mit sich ändernden Klima- und Umweltbedingungen, müssen auch dort forstliche Forscher

neue Wege beschreiten. Eine aktuelle Fragestellung ist die Frage nach der Biomassebindung in den Wäldern Sloveniens, an deren Beantwortung mehrere Gruppen arbeiten. Für die Beantwortung dieser Frage ist es unter anderem wichtig, möglichst präzise waldwachstumskundliche Größen, wie die Bestandeshöhe oder den Überschirmungsgrad zu kennen, um darauf aufbauend Biomasseschätzungen vornehmen zu können.

Um für Waldflächen einer Region an diese Größen zu gelangen, verfolgen Dzeroski et al. folgenden Ansatz: Von flugzeuggetragenen Laserscannern können sehr gute Daten gewonnen werden, die eine präzise Ermittlung der Bestandeshöhe und der Überschirmung von Beständen zulassen. Leider ist die Anwendung dieses Verfahrens sehr teuer (660 US\$/ km²) und daher nicht auf regionaler Ebene einsetzbar. Gleichzeitig sind für Slovenien flächig und kostengünstig (0,1 US\$/ km²) Landsat ETM+-Satellitenbilddaten verfügbar, in denen eventuell verborgenes Wissen im Zusammenhang mit den zwei zu ermittelnden waldwachstumskundlichen Größen steckt. Es existiert also ausschnittsweise eine vergleichsweise kleine Datenquelle mit einem hohen Detailgrad und gleichzeitig eine aus vielen Datensätzen bestehende Massendatenquelle mit einem niedrigeren Detailgrad. Für kleine Teilregionen überschneiden sich beide Datenquellen räumlich.

Dzeroski et al. (2006) versuchen mit Hilfe der Laserscanningdaten verschiedene Klassifikationsalgorithmen (M5, CUBIST) in der freien Data-Mining-Software WEKA (Witten und Frank, 2005) zu trainieren, um nach Anwendung auf die Landsat-Testdaten entsprechende Muster zu finden, die Rückschlüsse auf die Bestandeshöhe sowie den Überschirmungsgrad zulassen. Zur Klassifikation haben sie verschiedene Zielvariablen wie den Anteil der Vegetationsbedeckung (CC) innerhalb eines Bildausschnittes oder den Farbwert bei der größten gemessenen Baumhöhe innerhalb eines Bildausschnittes (FSH) eingeführt.

Die Anwendung der Klassifikationsalgorithmen brachte sehr hohe Übereinstimmungen (Korrelationskoeffizienten für CC zwischen 0,73 und 0,86 und zwischen 0,78 und 0,88 für FSH) zwischen den in den Satellitenbilddaten aufgefundenen Mustern und den (nicht zu Trainingszwecken) verwendeten Laserscanningdaten. Die Ergebnisse werden von Dzeroski et al. (2006) zum Anlass genommen, auf diesem Gebiet praxisorientiert weiter zu forschen.

In der Literatur findet sich für mittlerweile eine Vielzahl von Definitionen für das Begriffspaar „Data-Mining“. Eine vielfach zitierte, weit gefasste Definition stammt von Fayyad et al. (1996). Sie verstehen unter „Data-Mining“ (...) „die nicht-triviale Entdeckung neuer, gültiger, potenziell nützlicher und verständlicher Muster in Datenbeständen“. Data-Mining wird dabei als Teilprozess innerhalb einer als Knowledge-Discovery in Databases (KDD) bezeichneten Prozesskette verstanden. Wilde (2001) weist darauf hin, dass heute viele Autoren die Begriffe Data-Mining und

Knowledge-Discovery in Databases synonym verwenden. Im Folgenden soll jedoch explizit zwischen beiden Begriffen unterschieden werden.

Abbildung 4 veranschaulicht die Stellung des Data-Mining im KDD-Prozess und zeigt die Phasengrenzen.

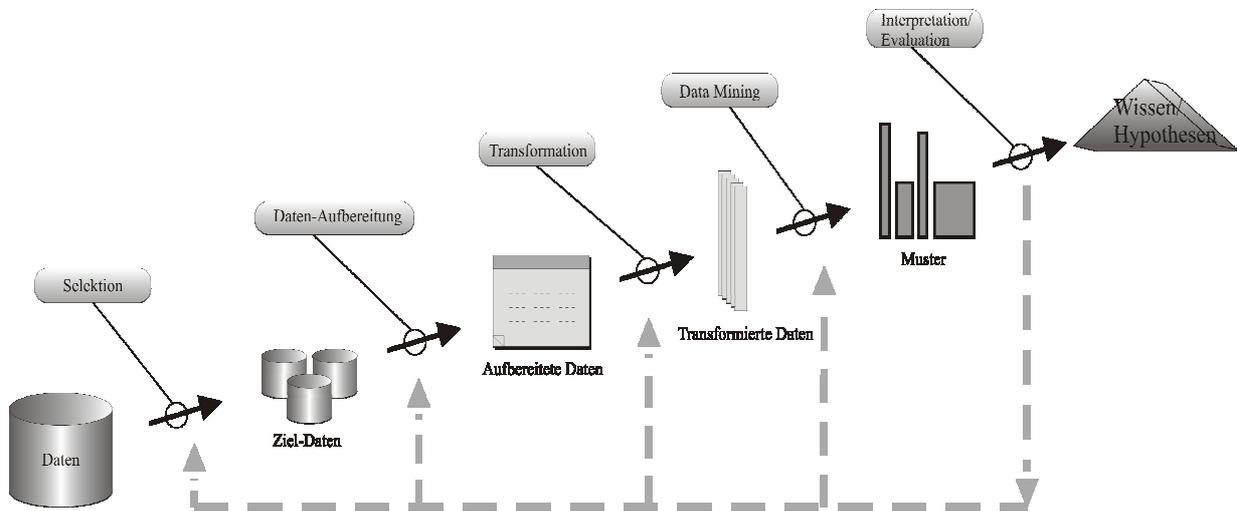


Abb. 4: Stellung des Data-Mining innerhalb des gesamten als Knowledge-Discovery in Databases (KDD) bezeichneten Prozesses (verändert nach Fayyad et. al. 1996)

Nach Hippner und Wilde (2001) verlaufen die einzelnen Phasen mit zahlreichen Rückkopplungen sowie in intensiver Interaktion mit dem Anwender.

Knobloch und Weidner (2000) berichten, dass gerade die Anwendung des Data-Mining Ansatzes zur Unterstützung und Vorbereitung betriebswirtschaftlicher Entscheidungen in der betriebswirtschaftlichen Praxis bislang häufig hinter den Erwartungen zurückgeblieben ist. Als Grund hierfür wird z. B. die mangelnde Effizienz bei den Teilschritten der Datenvorbereitung angeführt. Hierfür werden v. a. Datenmängel sowie Probleme bei der Datenharmonisierung ursächlich gemacht. Beide Autoren weisen daher darauf hin, dass die Ergebnisse des gesamten Prozesses sehr stark von der Datenvorbereitung beeinflusst werden. Witten und Frank (2005) beziffern den zeitbezogenen Aufwand für die Datenvorbereitung bzw. Datenaufbereitung auf 60 %, Knobloch und Weidner (2000) benennen ihn sogar mit bis zu 80 %. Da der größte Teil des Aufwandes für Data-Mining-Projekte auf diese Arbeitsphase entfällt, berichten Wilde (2001), Grob und Bensberg (1999), Lusti (2002), Witten und Frank (2005) sowie zahlreiche andere Autoren darüber, dass die Verwendung eines Data-Warehouse-Systems als Datenquelle erhebliche Synergieeffekte mit sich bringt, da sich die Verfügbarkeit einer bereits konsolidierten und bereinigten Datenbasis aus zahlreichen Gesichtspunkten wesentlich günstiger als der direkte Zugriff auf Rohdaten erweist.

Im Zusammenhang mit den Ergebnissen des KDD-Prozesses weisen Knobloch und Weidner (2000) darauf hin, dass die allein durch die Anwendung von Data-Mining-Methoden entdeckten Muster noch kein Wissen darstellen. Erst durch ergänzende Interpretations- und Bewertungsmaßnahmen kann entschieden werden, ob die ermittelten Informationen tatsächlich „interessant“, d. h. gültig, neuartig, nützlich und verständlich sind. Auch Witten und Frank (2005) sowie Lusti (2002) betonen die Bedeutung des Experten- oder Domänenwissens, durch welches erst eine Überführung der abgeleiteten Muster in Hypothesen oder letztendlich Wissen möglich wird.

Knobloch und Weidner (2000) formulieren weiterhin die Anforderungen an KDD-Systeme und leiten daraus Designkriterien für den Aufbau von Entscheidungsunterstützungssystemen ab. Nach beider Ansicht muss ein Datenanalyzesystem eine möglichst große Auswahl an „traditionellen“ und an Data-Mining-Methoden zur Verfügung stellen, die derart integriert sind, dass sie auf einfache Weise miteinander gekoppelt werden können. Außerdem sollen sie einfach und einheitlich bedienbar sein. Zur Ableitung von Handlungsmaßnahmen sind gängige Werkzeuge zur Präsentation und Visualisierung der Analyseergebnisse nicht ausreichend. Hieraus resultiert die Forderung nach einer ergonomischen Arbeitsumgebung, die über Präsentations- und Visualisierungswerkzeuge weit hinaus geht.

In Data-Mining-Anwendungen herrschen vier grundsätzlich verschiedene Arten des „Lernens“ vor. Beim *klassifizierenden Lernen* nimmt ein Lernverfahren eine Menge klassifizierter Beispiele entgegen, aus denen es lernen soll, unbekannte Beispiele zu klassifizieren. Beim *assoziiierenden Lernen* werden alle Assoziationen zwischen den Attributen berücksichtigt, nicht nur diejenigen, die einen bestimmten Klassenwert vorhersagen. Beim *Clustering* werden Gruppen von zusammengehörigen Beispielen gesucht. Bei der *numerischen Vorhersage* ist das vorhergesagte Ergebnis keine diskrete Klasse sondern eine numerische Größe (Witten und Frank, 2005). Lusti (2002) bezeichnet die verschiedenen Arten des Lernens als Anwendungsklassen und fügt darüber hinaus noch eine fünfte Klasse mit der Bezeichnung *Text Mining* hinzu. Tabelle 2 stellt tabellarisch zusammengefasst diese Anwendungsklassen noch einmal dar.

Einzelne Autoren benennen über die in Tabelle 2 hinaus genannten Anwendungsklassen zusätzlich Neuronale Netze als Klasse, die im Data-Mining Anwendung findet. Diese besitzen allerdings im Vergleich zu Klassifikationsalgorithmen und den ggf. daraus resultierenden Entscheidungsbäumen eine geringere Transparenz (Grob und Bensberg, 1999), weshalb diese Zuordnung im Wesentlichen erst in jüngeren Arbeiten vorgenommen wird.

Tab. 2: Klassen von Data-Mining Anwendungen nach Lusti (2002).

Klasse	Aufgabe	Anwendung (Beispiel)	Methodenbeispiel
Klassifikation	Individuen bereits bekannten Klassen zuordnen	OCR, Bonitätsbeurteilung	Regelinduktion, Neuronale Netze
Vorhersage	Kontinuierliche zukünftige Werte aus unabhängigen Variablen berechnen	Bonitätsbeurteilung	Regression, Neuronale Netze
Clustering	Gruppen aufgrund von Ähnlichkeiten zwischen Individuen identifizieren	Werbeadressaten einteilen	Neuronale Netze, 'konventionelle' Clusteranalyse
Assoziation	Abhängigkeiten entdecken und quantifizieren	Marktkorbanalyse	statistische Zusammenhangsanalyse
Text Mining	Textmuster suchen	Information Retrieval	Suchalgorithmen

Data-Mining Techniken werden mittlerweile in vielen Produktivsystemen im realen Wirtschaftsleben zur Findung und Unterstützung betrieblicher Entscheidungen eingesetzt. Einen umfassenden theoretischen Überblick über den Einsatz von Data Mining Methoden zur Unterstützung betrieblicher Entscheidungen liefert Tillmans (2003).

Ein konkretes Beispiel ist der Einsatz von Data-Mining Techniken im Bereich der Kundensegmentierung. Banken beurteilen beispielsweise standardmäßig die Kreditwürdigkeit von Kunden mit Hilfe geeigneter Data-Mining-Techniken (Lusti, 2002). Weitere Beispiele sind der Einsatz im Bereich des Managements von Kundenbeziehungen (Customer Relationship Management, CRM). Gerade im Teilbereich des analytischen CRM wird hier versucht, wichtige Informationen über den Kunden auf Basis der Kundenstamm- und Transaktionsdaten zu gewinnen, mit den Zielen neue Kunden zu akquirieren, bestehende Kunden bestmöglich zu pflegen bzw. verloren gegangene Kunden zurückzugewinnen (mySAP CRM, 2006).

Ein weiterer, zunehmend bedeutender werdender Bereich, ist das sog. Web Mining. Hierbei werden Techniken des Data-Mining zur (teil)automatisierten Extraktion von Inhalten aus dem Internet verwendet. Werden Inhalte durchsucht, spricht man von Web-Content-Mining. Wird die Struktur der Verbindungen im Netz analysiert, spricht man von Web-Structure-Mining, während man bei einer Analyse des Zugriffsverhaltens beispielsweise aus Logfiles von Web-Usage-Mining spricht.

Die aktivste Anwendung haben Data-Mining Techniken im praktischen Wirtschaftsleben in den Bereichen Marketing und Verkauf erfahren. Vielfach zitiert sind Anwendungen im Bereich der Warenkorbanalyse, wobei hier insbesondere Assoziationstechniken zum Einsatz kommen. Hierbei besteht das Hauptziel der Analyse letztendlich darin, Produktgruppen, die in Transaktionen kombiniert auftreten, zu finden. Diese Vorgehensweise wird häufig auf Daten von Kassierstellen (Point of Sales) in Supermärkten angewendet. Die dabei gewonnenen Informationen

besitzen für Warenanbieter einen erheblichen Wert und fließen nicht zuletzt in die Planung der Ladenarchitektur sowie in die Planung von Werbemaßnahmen ein (Witten und Frank, 2005). Ein weiterer wichtiger Bereich ist der Einsatz von Data-Mining-Techniken im Bereich des Direktmarketings. Hier werden Data-Mining-Techniken genutzt, um möglichst zielgruppengerecht kostenintensives Werbematerial zu versenden mit dem Ziel, die Rücklaufquote zu optimieren.

Im Bereich der Biowissenschaften finden sich zahlreiche Arbeiten zum Einsatz von Data-Mining-Techniken in der Genetik. Beispiele sind Arbeiten von Iratxeta et al. (2002) oder von King et al. (2000). Mittlerweile hat sich mit der sog. Bioinformatik eine eigene wissenschaftliche Teildisziplin entwickelt, in der das Erlernen von Data-Mining-Techniken einen Ausbildungsschwerpunkt darstellt.

2.3.2 Anwendung von Data-Mining-Verfahren in der Forstwirtschaft

In der Forstwissenschaft bzw. in der praktischen Forstwirtschaft ist über den Einsatz von Data-Mining-Techniken bisher wenig bekannt.

Eine grundlegende theoretische Arbeit, die sich mit der Verwaltung und Exploration waldökologischer Daten beschäftigt, stammt von *Rajanowa (1997)*. Diese Arbeit entstammt einer Zeit, in der die Arbeit mit relationalen Datenbankmanagement-Systemen noch vergleichsweise wenig entwickelt war. Die Arbeit beschäftigt sich mit dem Aufbau der Datenbank „Waldökosystem“ bzw. mit grundlegenden Schritten der Datenexploration. Einen Schwerpunkt hierbei stellen Visualisierungstechniken dar. Nach Rajanova bezeichnet Data-Mining eine zusammenfassende Methodik zum Extrahieren von Hypothesen aus großen Datenbanken. Neben der Visualisierung gehören dazu Techniken zur Ableitung von verschiedenartigen Modellen (neuronal, IF .. THEN-Regel-Sätzen, Entscheidungsbäume, Assoziationsregeln etc.). Für die Waldökologie sind das neue Ansätze, die zu mehreren Anwendungsideen führen. Data-Mining auf der Etappe des Konzipierens wird nach Rajanova (1997) die Bereicherung der Methoden auf den darauf folgenden Datenanalyseetappen fördern. Sie führt aus, dass z. B. die Generierung von IF .. THEN-Regel-Modellen aus den Daten die Erstellung einer soliden Wissensbasis und dadurch den verstärkten Einsatz von wissensbasierten Systemen ermöglichen wird. Die Waldökosystemforschung ist nach Rajanova (1997) ein Gebiet, auf dem neue methodische Ansätze nicht zu schnell aufgegriffen werden. „Eine wichtige Aufgabe der Angewandten Informatik bleibt daher die Präsentation von Pilotapplikationen, um einen ersten Schritt zum Zusammenführen und zur Analyse der vielen waldökologischen Informationen zu tun“.

Savnik et al. (2000) haben Erkennungsverfahren zum Auffinden von Mustern innerhalb von Zeitreihendaten entwickelt und implementiert. Die entwickelten Algorithmen, welche in Pseudocode-Notation angegeben sind, wurden an Jahrringmessungen an 50 Buchen aus zwei unterschiedlichen Regionen Deutschlands getestet. Nach Entwicklung, Implementierung und erfolgreichem Test wurden die Algorithmen zur Datierung sowie zum Auffinden kurzer Muster innerhalb der Zeitreihendaten nach Datenlage erfolgreich angewendet.

Ludwig et al. (2001) nutzen Daten des US Forest Services um im Rahmen der Informatik-Ausbildung im Fach „Knowledge Discovery in Databases“ an der Morris University of Minnesota, praxisrelevante Fragen mit Studenten zu bearbeiten. Über den Kurs im Sommer 2001 wird berichtet, dass die Studenten mit verschiedenen Data-Mining-Techniken das Datenmaterial auf die Frage des US Forest Service untersucht haben, ob Zusammenhänge zwischen der geographischen Lage und dem Lebensalter von Aspen bestehen. Es wird berichtet, dass keine gesicherten Zusammenhänge sowie keine speziellen Regionen gefunden werden konnten, die diese Annahme rechtfertigen.

Borrie et al. (2002) haben Data-Mining-Techniken im „Public Purpose Recreation Marketing“ eingesetzt. Ziel der Untersuchung war es, Informationen über die Besuchergruppen zu gewinnen, um diese in Zukunft zielgruppengerechter ansprechen zu können. Hierzu haben sie eine telefonische Befragung bei ca. 5000 Besuchern von Freizeit- und Erholungseinrichtungen des US Forest Service in den Bundesstaaten der Pazifikküste durchgeführt. Das vorliegende Datenmaterial wurde mit Hilfe von Entscheidungsbaumverfahren analysiert. Anhand der Prädiktoren Vertrauen (Trust), Bindung (Commitment), Sozialverantwortung (Social responsibility) und Unterstützung für von der öffentlichen Hand vorgehaltene Einrichtungen (support for public purpose) konnten vier typische Besuchergruppen ermittelt werden. Die erste isolierte Gruppe wird als „Unterstützer des Forest Service und dessen Anliegen bezeichnet, die aber keine aktiven Erholungssuchenden in der freien Natur sind“, die zweite wird als „Uninteressierte am US Forest Service“ bezeichnet, die dritte isolierte Gruppe wird als „Fans‘ des US Forest Service bezeichnet und umfasst Menschen, die die Anliegen des US Forest Service unterstützen und sich gerne in der freien Natur erholen. Die vierte Gruppe ist dominiert von Menschen der Landbevölkerung und hier insbesondere durch Männer“. Diese Gruppe zieht eine Erholung in der freien Natur vor und unterstützt den US Forest Service, ist aber nur bedingt bereit, die Anliegen der Verwaltung der Erholungseinrichtungen zu unterstützen.

Fei (2005) hat verschiedene mathematisch-statistische Techniken angewendet, um Verbreitungsschwerpunkte der Verjüngung von „Chestnut oak“ (*Quercus montana*) zu ermitteln. Datengrundlage waren 1433 Messpunkte in 52 Eichen-Mischbeständen in den zentralen Appalachen. Alle angewendeten Verfahren (ranking, SaTScan, ClusterSeer, classification tree) lieferten ähnlich gute Ergebnisse, „classification tree ermittelte allerdings nicht nur die Verbreitungsschwerpunkte der Verjüngung sondern stellte gleichzeitig die damit verbundenen biotischen und abiotischen Variablen zur Verfügung“.

Corona et al. (2005) haben die Einflüsse der standörtlichen Faktoren auf die Korkqualität in Sardinien mit Hilfe von Klassifikationsbaumalgorithmen untersucht. Datengrundlage waren 63 Probebäume aus sieben verschiedenen Straten, wobei die Stratenabgrenzung nach kleinstandörtlichen Gesichtspunkten vorgenommen wurde. Nach der Korkqualität wurde für jeden Probebaum eine Dummy-Variable eingeführt, die über- bzw. unterdurchschnittliche standörtliche Verhältnisse kennzeichnet. Aus einer Anzahl von 17 möglichen Prädiktoren wurden mit Hilfe des CART-Algorithmus als entscheidende standörtliche Größen für eine hochwertige Korkproduktion die Höhenlage, der Phosphorgehalt sowie der Sandgehalt des Bodens herausgestellt.

Fukuda und Pearson (2006) haben Klassifikationsalgorithmen angewendet, um aus Luftbildaufnahmen biotisch verursachte Waldschäden in einer Testregion in British Columbia automatisch zu extrahieren und deren Lage zu bestimmen. Die Datengrundlage hierfür stellten Luftbilder dar, die bei den jährlich stattfindenden Befliegungen des Canadian Forest Services entstanden sind. Die digitalen Luftbildaufnahmen wurden hierzu im Vorfeld der Anwendung der Klassifikationsalgorithmen in Pixelfelder gleicher Größe zerlegt. Anhand von Trainingsdatensätzen wurden mehrere Klassifikationsalgorithmen des Data-Mining-Frameworks WEKA nach vier Klassen der Entnadelung bzw. Entlaubung trainiert. Anschließend wurde die Erkennungsrate sowie die Abgrenzungsgenauigkeit bei der Anwendung an Testdaten mit terrestrischen Aufnahmen sowie mit Satellitenaufnahmen verglichen. Die Arbeit hatte zum Ergebnis, dass bei einer ausreichenden Zahl an Trainingsdatensätze sehr hohe Erkennungsraten von bis zu 84 % der Fälle gegenüber der terrestrischen Erhebung gefunden werden konnten. Die Erkennungsrate stieg dabei mit der Anzahl und der Definition der Heterogenität der Trainingsdatensätze an.

Auf die von Dzeroski et al (2006) durchgeführte Arbeit, die sich mit der Auswertung von Satellitenbildaufnahmen mit Hilfe von Data-Mining-Techniken beschäftigt, wurde bereits im einführenden Beispiel eingegangen.

Wie in Abschnitt 2.3.1 dieser Arbeit erwähnt, wird die Anwendung Neuronaler Netze zu Klassifikationszwecken zum Teil als Data-Mining-Anwendung gesehen. Als Beispiele für forstliche Arbeiten, die diese Techniken anwenden, sind die Arbeiten von Hanewinkel (2005) und Hasenauer et al. (2001) angeführt.

2.4 Konsequenzen aus dem dargestellten Stand des Wissens für die vorliegende Arbeit

In Abschnitt 2.1 dieser Arbeit wurde versucht, aus waldwachstumskundlicher Standardliteratur und neueren Forschungsarbeiten den Wissensstand zur Abhängigkeit des Wachstums von Einzelbäumen oder Waldbeständen von primären Grundfaktoren des Waldwachstums, von standörtlichen Faktoren und z. T. von waldbaulichen Maßnahmen darzustellen. Nach den zitierten Arbeiten hängt das Wachstum von Einzelbäumen und Waldbeständen stark von standörtlichen Faktoren ab. Das Höhenwachstum von Einzelbäumen im Wald bzw. von Waldbeständen kann dabei weniger durch aktive waldwirtschaftliche Maßnahmen beeinflusst werden als das Dickenwachstum. Es ist damit stärker von standörtlichen Faktoren abhängig als das Dickenwachstum der Waldbäume. Für eine optimale Höhenentwicklung von Einzelbäumen im Wald ist weiterhin eine baumartenspezifische, leichte Konkurrenzsituation günstig.

Das Wissen um die genaue Wirkung einzelner primärer Grundfaktoren auf das Höhen- oder Dickenwachstum von Einzelbäumen im Wald ist noch unvollständig. Ebenso ist das Wissen um die Wachstumsverhältnisse von Waldbäumen auf den von der Forstpraxis ausgeschiedenen Standorten äußerst lückenhaft. Vergleiche von Forschungsarbeiten, die sich letzterem Themenkomplex zugewandt haben, werden dadurch erschwert, dass z. T. unterschiedliche Verfahren der Standortkartierung angewandt werden oder gleiche Standorte z. T. unterschiedlich codiert werden.

All diese Gründe sind maßgeblich dafür, dass vorliegende Arbeit versucht, aus den Forstinventurdaten Rückschlüsse auf die maßgeblichen, wachstumsbestimmenden Triebkräfte am Standort zu ziehen. Dieser Rückschluss soll über die gemessenen Baumhöhen der Bäume der Oberschicht am Inventurpunkt erfolgen. Die erreichte Höhe dieser Bäume in einem bestimmten Alter soll also als Grundlage für die Beurteilung der Wuchsleistung einer bestimmten Baumart auf einem bestimmten Standort dienen. In Anhalt an die Arbeit von Kahn (1994) wird im weiteren Verlauf für diesen Zusammenhang auch von Standort-Leistungsbezug gesprochen.

Aufgrund der großen Anzahl der Datensätze in der Zentralen Forsteinrichtungsdatenbank der Bayerischen Staatsforsten sollen zur Extraktion der

wachstumsbestimmenden Triebkräfte des Wachstums am Standort Methoden angewandt werden, die sich in anderen Wirtschafts- und Wissenschaftsbereichen bereits bewährt haben, die in der Forstwirtschaft und Forstwissenschaft allerdings bisher kaum Anwendung gefunden haben. Charakteristisch für diese Methoden ist, dass sie stets auf große, z. T. neue Datenquellen angewendet werden und sie „datengetrieben“ (vgl. Abb. 1) automatisch neue Hypothesen generieren, deren Weiternutzung Expertenwissen voraussetzt (vgl. Abschnitt 2.3). Die stetig wachsende Größe der benutzten Datenquelle sowie die regionale Flächen- und Standortrepräsentativität der darin gehaltenen Daten ist ursächlich dafür, dass in dieser Arbeit für die Ableitung eines Standort-Leistungsbezuges kein modellhafter Ansatz, wie in Abschnitt 2.2 beschrieben, verfolgt wird sondern ein empirisch-statistischer Ansatz Anwendung finden soll. Gemäß den in Abschnitt 1.5 beschriebenen Teilzielen der Arbeit soll eine Übertragbarkeit auf andere Regionen sowie eine Wiederholbarkeit der Anwendung des neuen Verfahrens durch die Entwicklung einer eigenständigen Softwarelösung gewährleistet werden.

3. Material und Methoden

3.1 Datenmaterial

Die in dieser Arbeit anzuwendenden Methoden sollen auf die bei Forstinventur- und Forsteinrichtungsverfahren im Staatswald Bayerns erhobenen Daten angewendet werden. Diese Inventurdaten werden mit Hilfe eines relationalen Datenbankmanagementsystems zentral gespeichert. Bei der hierzu verwendeten Softwarelösung handelt es sich zum Stichtag der Datengewinnung (01.03.2005) um Informix 4.11 der Firma IBM (Bayerisches Staatsministerium für Landwirtschaft und Forsten, 1999). Nachfolgend werden einführend das Datenmodell sowie der Datenstand beschrieben. Daran anschließend wird versucht, über die Beschreibung der räumlichen Verteilung des Datenmaterials sowie über eine kurze numerische Beschreibung einen Eindruck über die Daten zu vermitteln.

3.1.1 Beschreibung des Datenmodells der „Zentralen Forsteinrichtungsdatenbank“ der Bayerischen Staatsforsten

Die bei der Forstinventur bzw. bei Forsteinrichtungsverfahren im Staatswald Bayerns erhobenen Daten werden mit Hilfe eines relationalen Datenbankmanagementmodells verwaltet. Die theoretischen Grundlagen derartiger Systeme werden bei Connolly et al. (2002) bzw. in kurzer Form bei Herbolsheimer (2002) dargestellt.

In relationalen Datenbanksystemen werden die Datensätze in Tabellen abgelegt. Die Tabellen stehen dabei untereinander in Beziehung. Jede Zeile einer Tabelle entspricht dabei einem Datensatz, jede Spalte dient der Aufnahme eines Attributes. Die maximal mögliche Anzahl der Tabellen einer mit Hilfe eines relationalen Datenbankmanagementsystems gehaltenen Datenbank ist dabei prinzipiell unbegrenzt und lediglich durch Hardwarerestriktionen limitiert.

Zur Sicherstellung der Datenintegrität werden die Forstinventur- und Forsteinrichtungsdaten bei den Bayerischen Staatsforsten mit Hilfe von zwei Datenbanken gehalten. Diese befinden sich auf zwei physisch getrennten Rechnern. Zur Haltung der Daten während des laufenden Forstinventur- und Forsteinrichtungsverfahrens dient die sogenannte „Aufnahmedatenbank“. Nach Abschluss eines Inventur- bzw. Einrichtungsverfahrens werden die Daten in eine als „Zentrale Forsteinrichtungsdatenbank“ bezeichnete Datenbank übertragen. Die Tabellen beider Datenbanken stimmen im strukturellen Aufbau weitgehend überein (Ausnahme: z. B. Tabelle „sdwao“).

Gegenstand der weiteren Arbeit sind die in der „Zentralen Forsteinrichtungsdatenbank“ gehaltenen Daten. Diese wurde zum 01.03.2005 vollständig kopiert und zur weiteren Verarbeitung auf einem Datenbankmanagementsystem der Firma Microsoft (SQL Server 2000, Developer Edition) abgelegt. Übertragen wurden hierbei 194 Tabellen. Auf sieben Tabellen der originären Informix-Datenbank konnte zum Stichtag nicht zugegriffen werden.

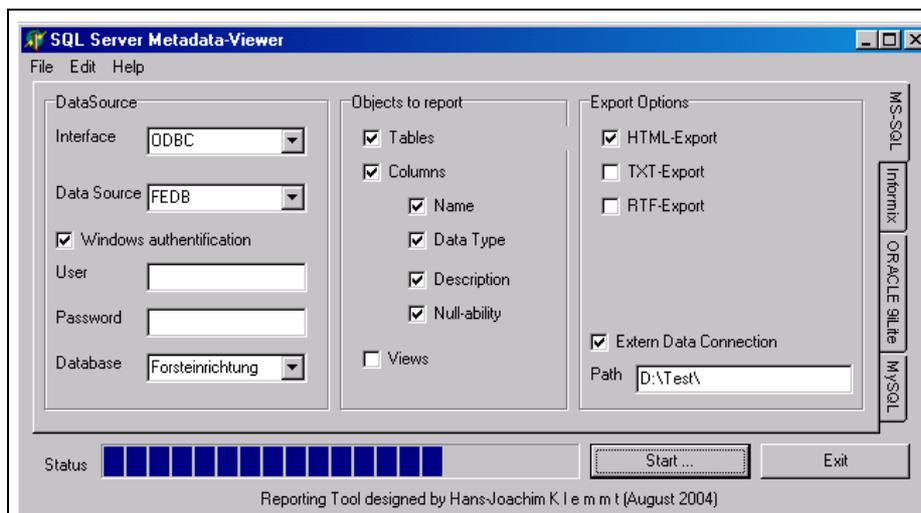


Abb. 5: Bildschirmabzug der Benutzeroberfläche des im Rahmen dieser Arbeit erstellten Programms zur Dokumentation der Metadaten einer Datenbank.

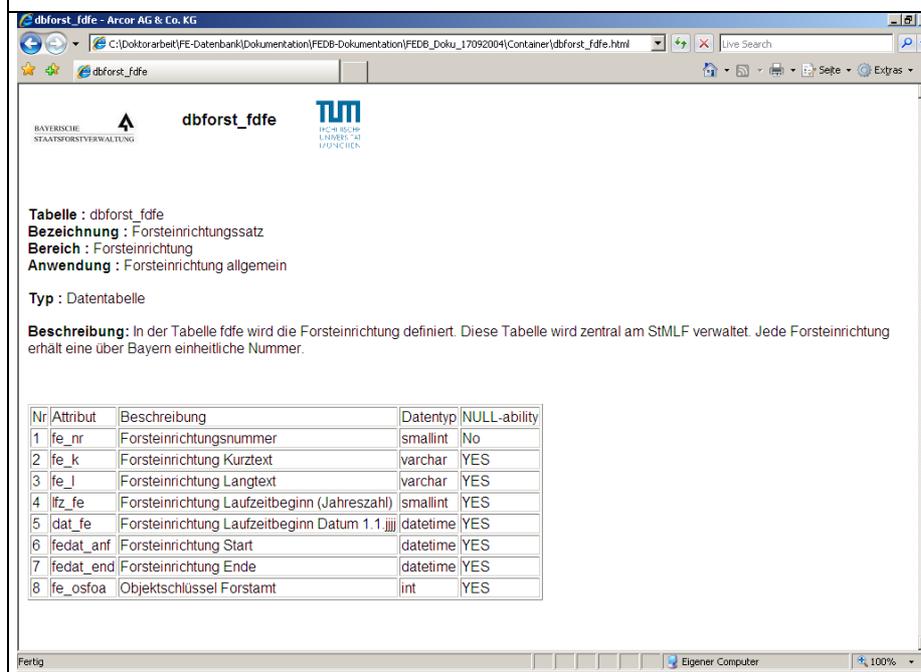


Abb. 6: Beispielhafte Darstellung einer mit dem nach Abb. 5 dargestellten Programm erzeugten HTML-Seite für die Tabelle fdfc der Zentralen Forsteinrichtungsdatenbank.

Für die „Aufnahmedatenbank“ sowie für die „Zentrale Forsteinrichtungsdatenbank“ existiert eine Dokumentation (Bayerisches Staatsministerium für Landwirtschaft und Forsten, 1999), die allerdings zum benannten Stichtag nicht mehr in allen Attributsbezeichnungen mit den kopierten Tabellen übereingestimmt hat. Um eine Analyse des Datenmodells und eine Dokumentation mit aktuellen

Attributsbezeichnungen zu erstellen, wurde daher zu Beginn dieser Arbeit eine einfache Softwarelösung erstellt, die die Metadaten einer vom Benutzer festzulegenden Datenbank ausliest. Diese Metadaten können als Textdatei oder als HTML-Datei dargestellt werden. Abbildung 5 zeigt einen Bildschirmabzug der Benutzeroberfläche des Programmsystems, Abbildung 6 zeigt exemplarisch das Ergebnis des Programmlaufes für die Tabelle „fdfe“ der Zentralen Forsteinrichtungsdatenbank. Das vollständige Ergebnis eines Programmlaufes befindet sich auf dem Datenträger, der ausgewählten schriftlichen Versionen dieser Arbeit beigelegt ist. Die Ergebnisdarstellung erfolgt dabei über einen Browser. Startpunkt ist die Datei „index.html“ im Verzeichnis „FEDB-Dokumentation“.

Die Tabellen sowohl der „Aufnahmedatenbank“ als auch der „Zentralen Forsteinrichtungsdatenbank“ lassen sich zu folgenden Gruppen zusammenfassen (Bayerisches Staatsministerium für Landwirtschaft und Forsten, 1999):

- *Tabellen der Verfahrenssteuerung:* Diese Tabellen enthalten Daten, die vom Inventurleiter vor Beginn der Aufnahmen auf die Belange des jeweiligen Forstbetriebes anzupassen sind. Beispiele hierfür sind die Festlegung von Kreisgrößen der Inventur, die Anlage von Baumartengruppen o. ä.. In Abbildung 7 sind die Tabellen der Verfahrenssteuerung braun dargestellt.
- *Schlüsseltabellen:* Schlüsseltabellen legen zulässige Wertebereiche der Attribute fest. Ein Beispiel hierfür ist die Tabelle „fcbaumart“, in der alle bei der Inventur aufnehmbaren Baumarten aufgelistet sind. In Abbildung 7 sind Tabellen dieser Gruppe gelb dargestellt.
- *Arbeitstabellen:* Arbeitstabellen enthalten die bei der Erstaufnahme enthaltenen Daten. In Abbildung 7 sind diese Tabellen rot dargestellt.
- *Urdatentabellen:* In Urdatentabellen befinden sich die Rohdaten für die Wiederholungsaufnahme. In Abbildung 7 sind Tabellen dieser Gruppe gepunktet dargestellt.
- *Datentabellen:* Datentabellen enthalten Datensätze, die nur von der zentral datenhaltenden Stelle angelegt oder verändert werden können. Ein Beispiel hierfür ist die Tabelle „fdfe“ (Forsteinrichtungssatz), in die vor Beginn eines Inventur- oder Forsteinrichtungsverfahrens ein Datensatz mit einer eindeutigen Forsteinrichtungsnummer eingetragen wird. In Abbildung 7 ist nur eine einzige Tabelle dieses Typs enthalten, die grün eingefärbt ist.
- *Produktionstabellen:* Die Produktionstabellen beinhalten Ergebnisse ertragskundlicher Berechnungen, in die Daten aus anderen Tabellengruppen eingehen. In Abbildung 7 sind diese Tabellen blau dargestellt.

Neben diesen Tabellengruppen enthält die Zentrale Forsteinrichtungsdatenbank der Bayerischen Staatsforsten Tabellen, die Informationen nach dem Forstsaatgutrecht enthalten. Diese Tabellen sind am Buchstabenkürzel „ez“ (= Erntezulassung) im Tabellennamen zu erkennen.

Abbildung 7 zeigt einen Ausschnitt der Tabellen der Zentralen Forsteinrichtungsdatenbank, die Daten der Forstinventur enthalten, bzw. die im Rahmen des Inventurverfahrens Verwendung finden.

In den blau dargestellten Produktionstabellen befinden sich die für diese Arbeit verwendeten Baum- und Probekreisdaten. Daher werden die drei Tabellen „fdinvkrs“, „fdinvba“ und „fdinvbhd“ nachstehend kurz verbal beschrieben. Abbildung 8 zeigt abschließend noch einmal ein Entity-Relationship-Diagramm für diese drei Tabellen, das die Beziehungen zwischen diesen Tabellen grafisch veranschaulicht.

Für Einzelbäume erhobene Daten werden in der Tabelle „fdinvbhd“ (Inventur BHD-Stufe) gespeichert. Für jeden in einem Probekreis stehenden Baum, dessen Brusthöhendurchmesser (BHD) die in der Verfahrenssteuerung festgelegte Kluppschwelle überschreitet, wird ein Datensatz angelegt, der ggf. die Polarkoordinaten, den BHD, die gemessene oder die aus der Bestandeshöhenkurve entnommene Höhe, die auf den Hektar hochgerechneten Werte für Stammzahl, Grundfläche und Vorrat sowie die errechneten Jahreszuwächse für BHD, Höhe, Grundfläche und Vorrat beinhaltet.

Einzelbaumdaten aus der Tabelle „fdinvbhd“ sind in der Tabelle „fdinvba“ (Inventur Baumarten) aggregiert zu Sätzen, die je Probekreis Daten für alle Bäume enthalten, die der selben Art angehören, gleich alt sind und der selben Bestandesschicht zugeordnet werden. Jeder „fdinvbhd“-Satz einer Forsteinrichtung kann durch eine gemeinsame Probekreiskoordinate und eine eindeutige laufende Nummer mit einem Baumartensatz verknüpft werden. Der Tabelle „fdinvba“ sind Informationen über Baumart, Bauartenanteil, Alter, Schichtzugehörigkeit, Höhe und BHD des Grundflächenmittelstamms, die auf Hektar hochgerechneten Werte von Stammzahl, Vorrat und Zuwachs, Bonität, Bestockungsgrad sowie die Ertragstafelprognose des Volumenzuwachses für die folgenden 10 Jahre zu entnehmen.

Die Tabelle „fdinvkrs“ enthält Informationen zu jedem Probekreis. Hierzu zählen z. B. die Repräsentationsfläche, die Geländeneigung sowie Daten zum Standort.

Vereinfachtes Datenmodell der FE-Datenbank (Teilbereich "Inventur") der Bayerischen Staatsforstverwaltung

LEGENDE

- Schlüssel Tabellen
- Arbeitstabellen
Altdaten EG 20
- Arbeitstabellen EG
20
- Datentabellen zur
Verfahrenssteuerung
- Produktionstabellen

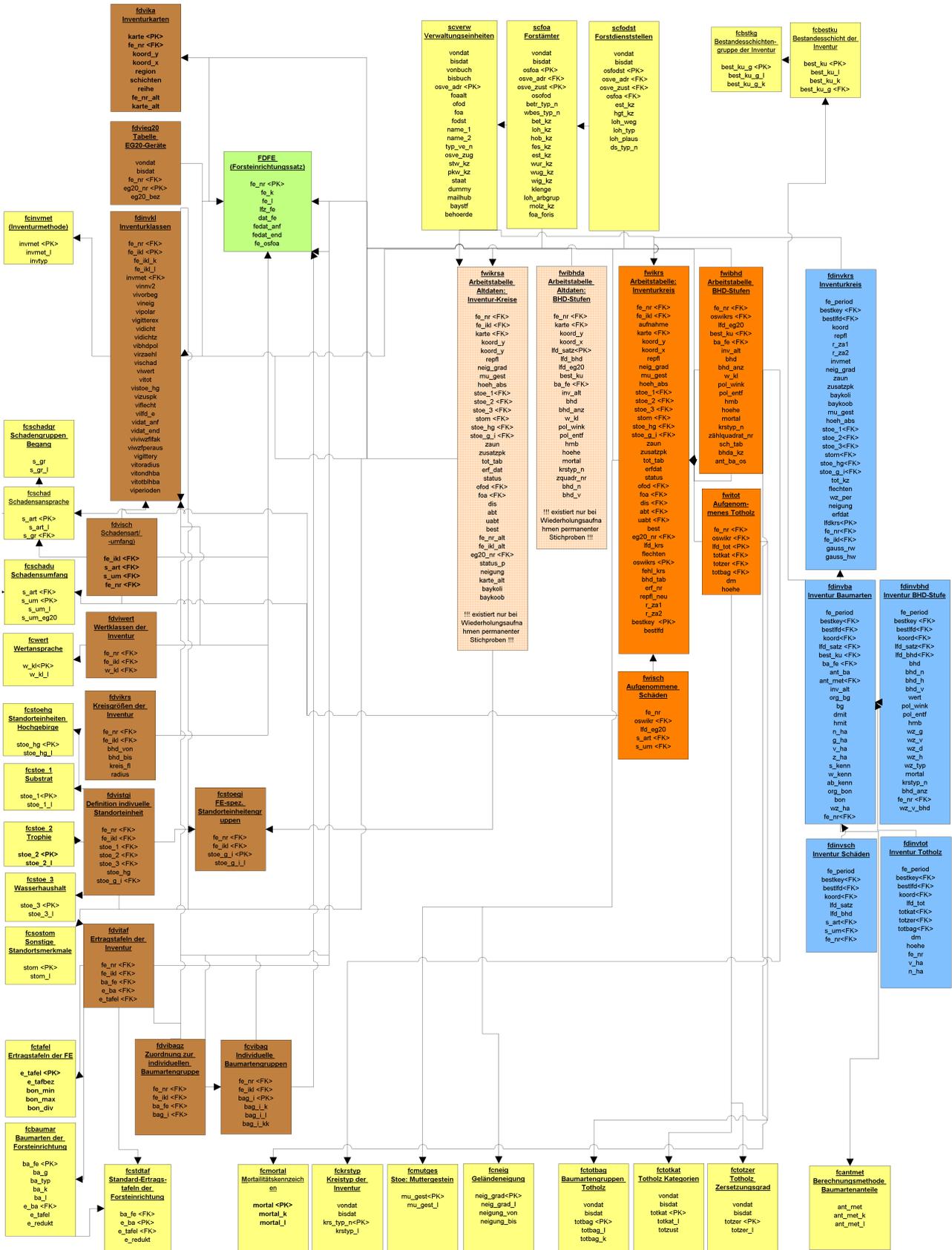


Abb. 7: Vereinfachtes Datenmodell eines Teiles der Tabellen der Zentralen Forsteinrichtungsdatenbank, die im Rahmen von Inventurverfahren verwendet werden.

Tabelle „fdinvkrs“ (Inventur Kreis)

Nr	Attribut	Beschreibung	Datentyp	NULL-ability
1	fe_period	Forsteinrichtungsperiode	smallint	YES
2	bestkey	Bestandsschlüssel	decimal	YES
3	bestlfd	Bestandsteifläche	smallint	YES
4	koord	Inv.Kreis Koordinatenwerten ka+koordY+koordX	bigint	No
5	repfl	Repräsentationsfläche	decimal	YES
6	r_zal	Einreihungsfläche 1. Zeitabschnitt	decimal	YES
7	r_zal2	Einreihungsfläche 2. Zeitabschnitt	decimal	YES
8	invmet	Inventurmethode	smallint	YES
9	neig_grad	Geländeneigungsklasse	smallint	YES
10	zaun	Zäunungskennzeichen	varchar	YES
11	zusatzpk	Zusatzpunkt vorhanden? J/N	varchar	YES
12	baykoli	Bayernkoordinaten Links	bigint	YES
13	baykoob	Bayernkoordinaten Oben	bigint	YES
14	mu_gest	Muttergestein	smallint	YES
15	hoeh_abs	Absolute Höhenlage	smallint	YES
16	stoe_1	Substrat	smallint	YES
17	stoe_2	Trophie	smallint	YES
18	stoe_3	Wasserhaushalt	smallint	YES
19	stom	Sonstige Standortsmerkmale	smallint	YES
20	stoe_hg	Standortseinheit im Hochgebirge	smallint	YES
21	stoe_g_i	Standortseinheitengruppe individuell	smallint	YES
22	tot_kz	Totholzzeichen	varchar	YES
23	flechten	Flechten in %	varchar	YES
24	wz_per	Zuwachsperioden seit letzter Inventur	smallint	YES
25	neigung	Geländeneigung in Grad	smallint	YES
26	erfdat	Aufnahmedatum	datetime	YES
27	lfd_krs	Kreisnummer bei EG20-Aufnahme	bigint	YES
28	fe_nr	Forsteinrichtungsnummer	smallint	No
29	fe_ikl	Forsteinrichtungs-Inventurklasse	smallint	YES
30	gauss_rw	Gauss-Krüger-Koordinaten Rechtswert	bigint	YES
31	gauss_hw	Gauss-Krüger-Koordinaten Hochwert	bigint	YES

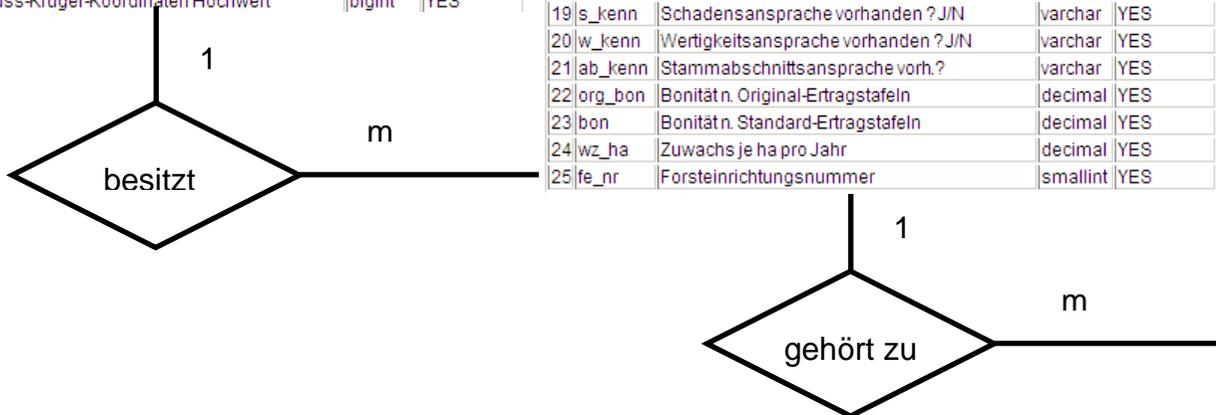
Abb 8: Entity-Relationship-Diagramm der Produktionstabellen „fdinvkrs“, „fdinvba“ und „fdinvbhd“ der Zentralen Forsteinrichtungsdatenbank (verändert nach Bayerisches Staatsministerium für Landwirtschaft und Forsten, 1999)

Tabelle „fdinvba“ (Inventur Baumarten)

Nr	Attribut	Beschreibung	Datentyp	NULL-ability
1	fe_period	Forsteinrichtungsperiode	smallint	No
2	bestkey	Bestandsschlüssel	decimal	No
3	bestlfd	Bestandsteifläche	smallint	No
4	koord	Inv.Kreis Koordinatenwerten ka+koordY+koordX	bigint	No
5	lfd_satz	BA-Satznr. innerhalb des Probekreises	smallint	No
6	best_ku	Bestandsschicht bei der Inventur (Schlüsselwert)	smallint	YES
7	ba_fe	Baumart	smallint	YES
8	ant_ba	Baumartenanteil in %	decimal	YES
9	ant_met	Anteil-Ba Berechn.-Methode	smallint	YES
10	inv_alt	Inventuralter	smallint	YES
11	org_bg	Bestockungsgrad n. Original-Ertragstafeln	decimal	YES
12	bg	Bestockungsgrad n. Standard-Ertragstafeln	decimal	YES
13	dmit	Mittendurchmesser	decimal	YES
14	hmit	Mittelhöhe	decimal	YES
15	n_ha	Stammzahl je ha	bigint	YES
16	g_ha	Grundfläche je ha	decimal	YES
17	v_ha	Vorrat je ha	decimal	YES
18	z_ha	Zuwachs je ha pro Jahr	decimal	YES
19	s_kenn	Schadensansprache vorhanden? J/N	varchar	YES
20	w_kenn	Wertigkeitsansprache vorhanden? J/N	varchar	YES
21	ab_kenn	Stammabschnittsansprache vorh.?	varchar	YES
22	org_bon	Bonität n. Original-Ertragstafeln	decimal	YES
23	bon	Bonität n. Standard-Ertragstafeln	decimal	YES
24	wz_ha	Zuwachs je ha pro Jahr	decimal	YES
25	fe_nr	Forsteinrichtungsnummer	smallint	YES

Tabelle „fdinvbhd“ (Inventur BHD-Stufen)

Nr	Attribut	Beschreibung	Datentyp	NULL-ability
1	fe_period	Forsteinrichtungsperiode	smallint	No
2	bestkey	Bestandsschlüssel	decimal	No
3	bestlfd	Bestandsteifläche	smallint	No
4	koord	Inv.Kreis Koordinatenwerten ka+koordY+koordX	bigint	No
5	lfd_satz	BA-Satznr. innerhalb des Probekreises	smallint	No
6	lfd_bhd	BHD-Stufe innerhalb des Probekreises	smallint	No
7	bhd	Brusthöhendurchmesser	decimal	YES
8	bhd_n	Stammzahl der BHD-Stufe	decimal	YES
9	bhd_h	Höhe der BHD-Stufe	decimal	YES
10	bhd_v	Vorrat der BHD-Stufe	decimal	YES
11	wert	Wertklasse	varchar	YES
12	pol_wink	Polarkoordinaten Winkel	varchar	YES
13	pol_entf	Polarkoordinaten Entfernung	varchar	YES
14	hmb	Höhenmessbaum? J/N	varchar	YES
15	wz_g	Grundflächenänderung pro Jahr	varchar	YES
16	wz_v	Vorratsänderung pro Jahr	varchar	YES
17	wz_d	Durchmesseränderung pro Jahr	varchar	YES
18	wz_h	Höhenänderung pro Jahr	varchar	YES
19	wz_typ	Fraktion/ Typ des Zuwachses	varchar	YES
20	mortal	Mortalitätskennzeichen perm. Inventur	varchar	YES
21	krstyp_n	Inventur-Kreistyp	varchar	YES
22	bhd_anz	Anzahl der aufgenommenen Bäume mit diesem BHD	varchar	YES
23	fe_nr	Forsteinrichtungsnummer	smallint	YES
24	wz_v_bhd	Volumenzuwachs des Einzelbaums	varchar	YES



3.1.2 Datenstand zum 01.03.2005

Tabelle 3 führt die Anzahl der Datensätze in den in Abbildung 8 dargestellten Tabellen auf, wobei die Anzahlen für die Arbeitstabellen und die Schlüsseltabellen nicht enthalten sind.

Zum 01.03.2005 waren 418 Datensätze in der Tabelle „fdfe“ (Forsteinrichtungssatz, vgl. Abb. 7) angelegt. Diese Zahl ist nicht identisch mit tatsächlich durchgeführten Forstinventur- und Forsteinrichtungsverfahren, da z. B. auch für Schälsschadensinventuren o. ä. Datensätze in dieser Tabelle angelegt wurden.

In der Tabelle „fdinvbhd“ fanden sich zum benannten Datum 8.123.296 Datensätze. Die Aufnahme ist dabei in 626.147 Probekreisen erfolgt (Datenstand Tabelle „fdinvkrs“). Die 8.123.296 Einzelbaumdaten wurden über die Berechnungsprogramme der Forsteinrichtung zu 2.105.845 Datensätzen für jede Baumart in einem gegebenen Alter pro Bestandesschicht verdichtet (Datenstand Tabelle „fdinvba“).

Tab. 3: Datenstand in den inventurrelevanten Tabellen der Zentralen Forsteinrichtungsdatenbank gemäß Abbildung 7 zum 01.03.2005.

Tabellenname	Anzahl Datensätze	Tabellenname	Anzahl Datensätze
fcantmet	3	fctotzer	3
fcbaumar	45	fcvibag	1.613
fcinvmet	7	fcwert	3
fckrstyp	5	fdfe	418
fcmortal	3	fdinvba	2.105.845
fcmutges	10	fdinvbhd	8.123.296
fcneig	9	fdinvkrs	626.147
fcschad	9	fdinvsch	1.100.283
fcschadgr	8	fdinvtot	105.947
fcschadu	20	fdvieg20	80
fcsostom	10	fdvika	818
fcstdtaf	45	fdvikl	594
fcstoe1	10	fdvikrs	808
fcstoe2	10	fdvisch	938
fcstoe3	10	fdvistgi	7.180
fcstoegi	997	fdvitaaf	22.581
fcstoehg	21	fdviwert	543
fctafel	14	scfoa	604
fctotbag	4	scfodst	2.892
fctotkat	4	scverw	3.764

3.1.3 Räumliche Verteilung des Datenmaterials

Abbildung 9 zeigt die Lage des Staatswaldes in Bayern nach Wuchsgebieten und Wuchsbezirken. Die Grundlage hierfür stellt die Wuchsgebietsgliederung für Bayern (Bayerische Landesanstalt für Wald und Forstwirtschaft, 2001). Zum Vergleich dazu ist neben Abbildung 9 in Abbildung 10 die Lage der Versuchsfelder des langfristigen ertragskundlichen Versuchswesens in Bayern eingetragen. Es erfolgt hierbei eine Untergliederung nach Baumarten und Versuchsarten.

Ein Vergleich beider Abbildungen verdeutlicht die in Abschnitt 1.3 getroffenen Aussagen zur unterschiedlichen Flächenrepräsentativität der Versuchsfelder- und Forstinventurdaten in Bayern. Ergänzend hierzu findet sich in der Tabelle in Anhang 9.1 die absolute Staatswaldfläche in Hektar nach Wuchsgebieten, Wuchsbezirken und Teilwuchsbezirken. Dieser Tabelle können auch die Bezeichnungen für die in Abbildung 9 dargestellten Wuchsbezirke entnommen werden.

Aus Abbildung 9 und Anhang 9.1 geht hervor, dass die größten absoluten Staatswaldflächen im Wuchsbezirk 15.5 („Mittlere Bayerische Kalkalpen“) liegen. Danach folgt der Wuchsbezirk 15.6 („Chiemgauer Alpen und Salforstamt St. Martin“). Letztere Flächenberechnung wurde nicht um die Staatswaldflächen auf österreichischer Landesfläche (Salforsten) bereinigt. An dritter Stelle folgt die Staatswaldfläche im Wuchsbezirk 6.2 („Südliche Frankenalb und Südlicher Oberpfälzer Jura“), gefolgt von den Staatswaldflächen im Wuchsbezirk 9.1 („Oberpfälzer Becken- und Hügelland“) und 15.9 („Berchtesgadener Hochalpen und Salforstamt St. Martin“). Bei der Berechnung letzterer Fläche wurden wiederum nicht die auf österreichischer Landesfläche liegenden Staatswaldflächen abgezogen. In den benannten fünf Wuchsbezirken liegen insgesamt ca. 27,5 % der Staatswaldfläche Bayerns. Die Tatsache, dass unter den fünf benannten Wuchsbezirken drei Wuchsbezirke zu finden sind, die dem Wuchsgebiet 15 („Bayerische Alpen“) angehören, unterstreicht die Bedeutung des Bergwaldes für die Bewirtschaftung des Staatswaldes in Bayern.

In der Tabelle in Anhang 9.1 grau markiert sind Wuchsbezirke, für die Staatswaldflächen unter 100 ha errechnet wurden. Diese Wuchsbezirke werden aufgrund der zu geringen Flächenausstattung in Verbindung mit der zu geringen Anzahl an Probekreisen im weiteren Verlauf der Arbeit nicht ausgewertet.

Weiterhin wird in den folgenden Abschnitten 3.1.4 bis 3.1.6, dem in Abschnitt 2.3.1 als „Knowledge-Discovery-in-Databases (KDD)“ bezeichneten Prozesses folgend, die weitergehende Datenvorbereitung erläutert.

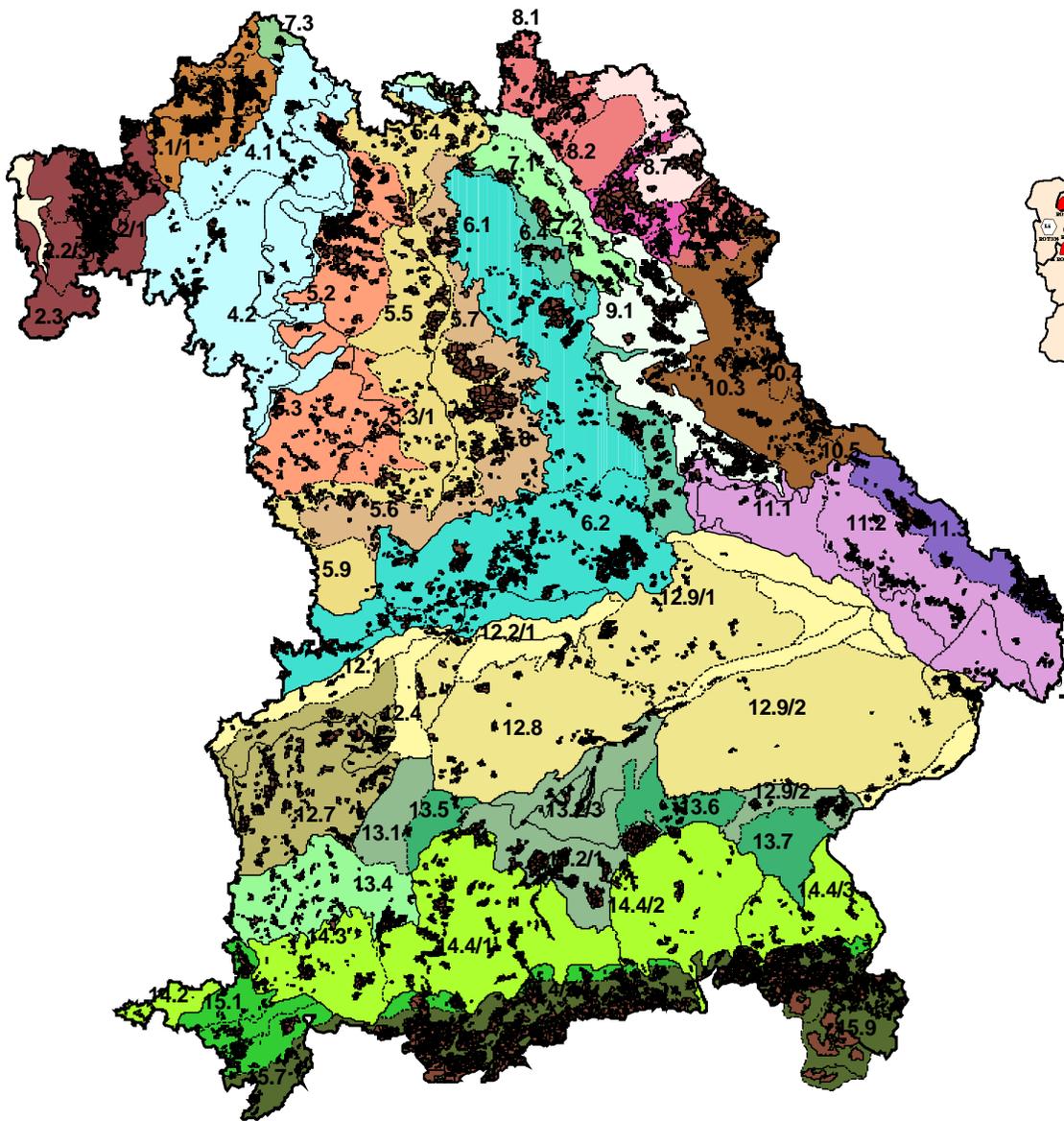


Abb. 9: Lage des Staatswaldes in Bayern nach Wuchsbezirken.
Kartengrundlage: Bayerische Landesanstalt für Wald und Forstwirtschaft, 2001

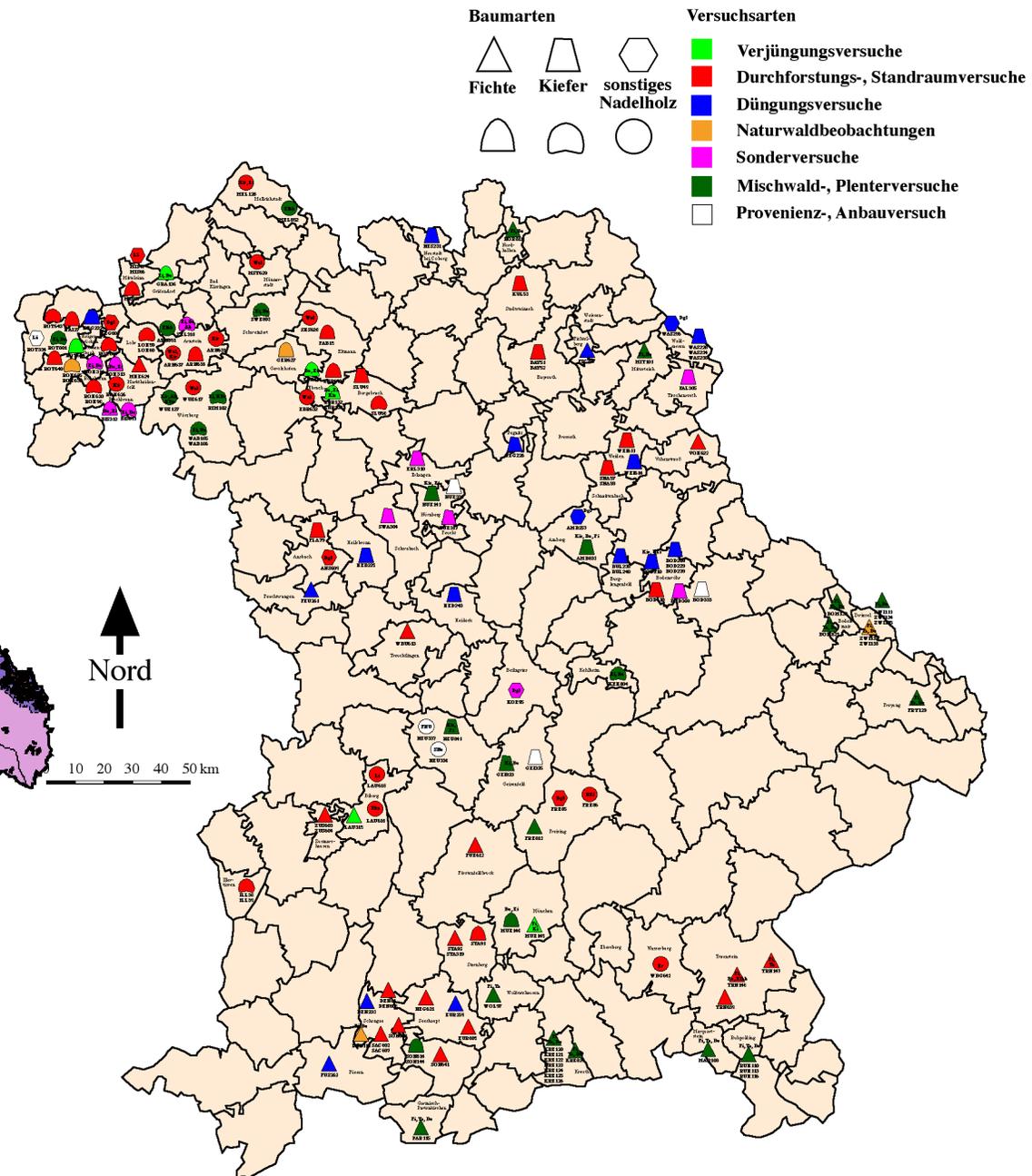


Abb. 10: Verteilung der Versuchsflächen des langfristigen ertragskundlichen Versuchswesens in Bayern. Kartengrundlage: Bayerisches Staatsministerium für Landwirtschaft und Forsten, 2004

3.1.4 Auswahl von Forstinventurdaten zur weiteren Auswertung

In der „Zentralen Forsteinrichtungsdatenbank“ der Bayerischen Staatsforsten befanden sich zum 01.03.2005 Baumdaten aus 626.147 Probekreis-aufnahmen. Wie aus Abbildung 8 ersichtlich ist, kann jeder Probekreis bezüglich seiner räumlichen Lage aufgrund der angegebenen Gauss-Krüger-Koordinaten lokalisiert werden. Diese Lageinformation stellt die Grundlage für die weitere Datenaufbereitung dar. Zur Überprüfung der Güte der Lageangaben wurden die Probekreise mit Hilfe der GIS-Applikation ArcView 3.3 mit den dazugehörigen Forstbetriebskarten verschnitten. Bereits eine visuelle Überprüfung hat dabei gezeigt, dass zahlreiche Probekreise bezüglich ihrer Lage nicht richtig zugeordnet waren. Da eine richtige Lagezuordnung allerdings ein wesentliches Qualitätskriterium für die weitere Datenaufbereitung darstellt, wurden alle in der Tabelle „fdfe“ eingetragenen Forsteinrichtungen bezüglich der räumlichen Lage der Probekreise überprüft. Datensätze von Forsteinrichtungen, deren Probekreise nicht zu 95 % innerhalb der Grenzen der Forstbetriebskarten lagen und deren Probekreise nicht auf einem erkennbaren Rasternetz lagen, wurden für die weitere Verarbeitung nicht verwendet (vgl. Abb. 11, „Überprüfung Schritt 1“).

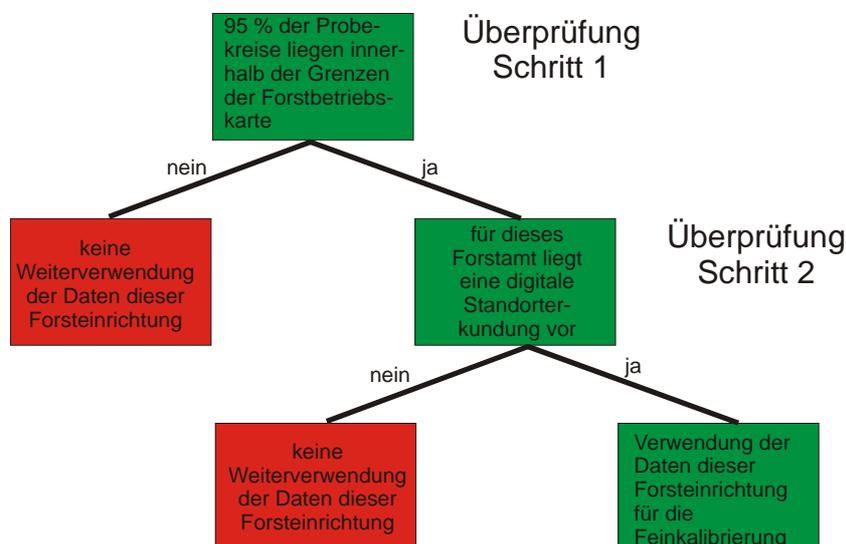


Abb. 11: Entscheidungsschema für die Auswahl der Datensätze dieser Arbeit

Aus Abbildung 8 geht hervor, dass für jeden Probekreis eine Standorteinheit angegeben wird. Die Angabe erfolgt dabei nach den drei Standortmerkmalen Substrat (stoe_1), Trophie (stoe_2) und Wasserhaushalt (stoe_3). Da die Standorterkundung im Bayerischen Staatswald einer in unregelmäßigen Zeitabständen erfolgenden Überprüfung und ggf. Überarbeitung unterliegt, soll durch

eine geeignete Verschneidung der Probekreisdaten mit aktuellen Standortkarten sichergestellt werden, dass alle Probekreise aktuelle, vergleichbare Standortinformationen besitzen. Da für den Bayerischen Staatswald zum 01.03.2005 noch nicht für alle Forstämter eine digitale Standortkartierung vorlag, wurden Forsteinrichtungsdatensätze entfernt, für die keine digitale Verschneidung durchgeführt werden konnte (vgl. Abb. 11, „Überprüfung Schritt 2“).

In Anhang 9.2 zu dieser Arbeit findet sich eine Tabelle, die die Forsteinrichtungen auflistet, die gemäß dem Entscheidungsbaum nach Abbildung 11 für diese Arbeit Verwendung finden.

Aus diesen Forsteinrichtungsdatensätzen wurden diejenigen Probekreise entfernt, die nach der Lageverschneidung nicht innerhalb der Forstbetriebskarte lagen. Nach der Entfernung dieser Probekreise wurden für die Baumarten Fichte, Kiefer, Tanne, Douglasie, Lärche sowie Buche und Eiche alle als Höhenmessbäume (Tabelle „fdinvbhd“, Wert des Attributes „hmb“=“j“) der Oberschicht (Tabelle „fdinvba“, Wert des Attributes „best_ku“=“1“) bzw. alle als Höhenmeßbäume aus „ungeschichtet“ gekennzeichneten Beständen (Tabelle „fdinvba“, Wert des Attributes „best_ku“=“0“) selektiert.

3.1.5 Zahlenmäßige Beschreibung des verwendeten Zahlenmaterials

Im Anhang 9.3 sind in den Tabellen für die Wuchsgebiete 2 bis 14 für die Nadelbaumarten Fichte, Tanne, Douglasie, Kiefer sowie für die Laubbaumarten Eiche und Buche die Anzahlen der höhengemessenen Bäume sowie die primären Messgrößen „Brusthöhendurchmesser“ (Quelltabelle in der Zentralen Forsteinrichtungsdatenbank: „fdinvbhd“, Attribut „bhd“) und „Baumhöhe“ (Quelltabelle in der Zentralen Forsteinrichtungsdatenbank: „fdinvbhd“, Attribut „bhd_h“) eingetragen. Des Weiteren enthalten die Tabellen Angaben zum mittleren Alter der Bäume. Es handelt sich hierbei um eine mittlere Altersangabe pro Baumart und Schicht (Quelltabelle „fdinvba“, Attribut „inv_alt“).

Bei der Betrachtung der Tabellen fällt auf, dass für mehrere Wuchsbezirke nur geringe Anzahlen an Höhenmeßbäumen gefunden werden konnten. Weitergehende Auswertungen wurden daher nur durchgeführt, wenn mindestens 100 höhengemessene Bäume vorhanden waren. Diese Einschränkung wurde vorgenommen um sicherzustellen, dass das gesamte Altersspektrum einer Baumart in einem Wuchsbezirk mit einer ausreichenden Zahl an Höhenmessbäumen

abgedeckt werden kann. Dies ermöglicht die Ableitung einer plausiblen, standortbezogenen Alters-Höhenbeziehung aus dem Datenmaterial mit Hilfe geeigneter regressionsanalytischer Methoden.

Nachfolgend sind in den Abbildungen 12 bis 14 die Spannweiten sowie die Mittelwerte der Inventuralter sowie der Brusthöhendurchmesser und der Baumhöhen für die Laubbaumarten Buche und Eiche (Abb. 12), Fichte und Kiefer (Abb. 13) sowie Lärche und Douglasie (Abb. 14) dargestellt. Um die Vergleichbarkeit zu erleichtern, sind bei den gleichen Werten auf der X-Achse Hilfslinien eingetragen.

Auffällig an den Daten für die Wuchsgebiete 2 („Spessart-Odenwald“), 3 („Rhön“) bis 4 („Fränkische Platte“) ist einerseits das relativ hohe Durchschnittsalter der Eichen im Spessart, was auf die unterschiedliche Behandlung und die höhere Bedeutung der Eichenwirtschaft im Spessart zurückzuführen ist (Bayerisches Staatsministerium für Landwirtschaft und Forsten, 1999). Weiterhin auffällig sind die unterdurchschnittlichen Baumhöhen für alle Baumarten in der Fränkischen Platte. Dies ist ein Folge der unterdurchschnittlichen Niederschlagsversorgung dieser Region Bayerns (BayFORKLIM, 1996).

Weiterhin zeigt eine Betrachtung der Daten für die Wuchsbezirke 5.1 („Hassberge“), 5.2 („Steigerwald“) und 5.3 („Frankenhöhe“), dass die unteren Altersgrenzen und komplementär die unteren Durchmessergrößen relativ hohen Werte aufweisen. Der Grund hierfür liegt darin, dass nur vergleichsweise alte Forsteinrichtungen berücksichtigt wurden, in denen nur Bäume des äußeren Messkreises als Bäume der Oberschicht gekennzeichnet wurden.

Für das Wuchsgebiet 9 („Oberpfälzer Becken- und Hügelland“) fällt auf, dass lediglich für die Baumart Kiefer die gesetzte Mindestanzahl an Höhenmessbäumen erreicht wurde. Grund hierfür waren zum Auswertungszeitpunkt nicht vorliegende digitale Standortkartierungen für die ehemaligen Bayerischen Forstämter Weiden und Pressath. Um Auswertungen für diesen Wuchsbezirk durchführen zu können, wurden nur Daten aus dem Forstamt Sulzbach-Rosenberg herangezogen, dessen Waldflächen nach der digitalen Verschneidung allerdings nur zu einem geringen Teil im Wuchsgebiet 9 lagen.

Weiterhin auffällig an den Datenwerten sind die jeweils vergleichsweise großen Spannweiten für Alter, Durchmesser und Höhe bei den Baumarten für den Wuchsbezirk 14.4 („Oberbayerische Jungmoräne und Molassevorberge“). Der Grund hierfür liegt in der zusammengefassten Darstellung der drei Teilwuchsbezirke, wodurch sich für den ganzen Wuchsbezirk hohe Anzahlen an Höhenmessbäumen mit jeweils breiten Wertespektren ergeben.

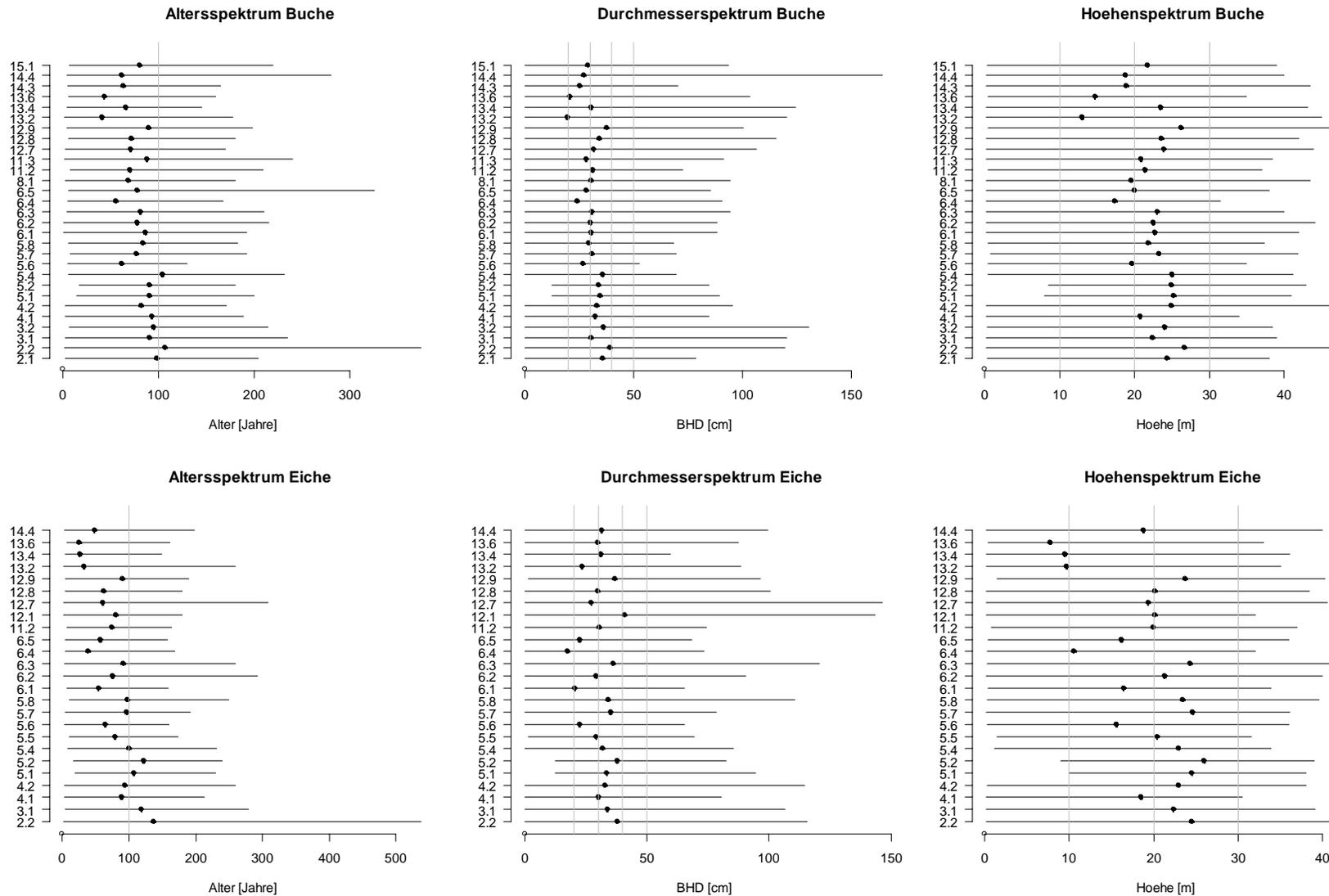


Abb. 12: Spannweite und Mittelwerte für das Alter, den Durchmesser sowie die Höhe für die Laubbaumarten Buche (oben) und Eiche (unten) nach Wuchsbezirken. Die Datengrundlage bilden die Forstinventurdaten des Staatswaldes in Bayern.

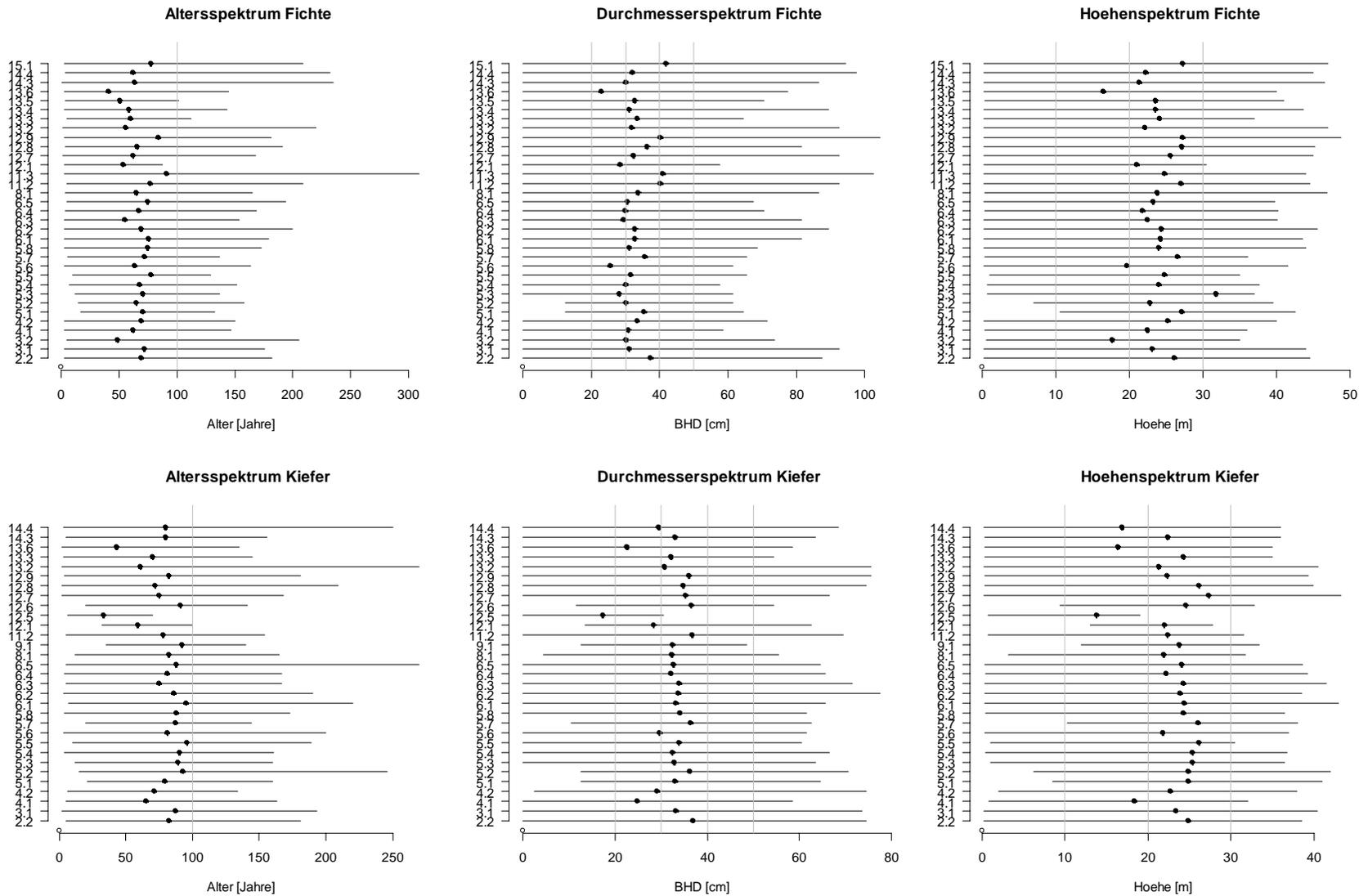


Abb. 13: Spannweite und Mittelwerte für das Alter, den Durchmesser sowie die Höhe für die Nadelbaumarten Fichte (oben) und Kiefer (unten) nach Wuchsbezirken. Die Datengrundlage bilden die Forstinventurdaten des Staatswaldes in Bayern.

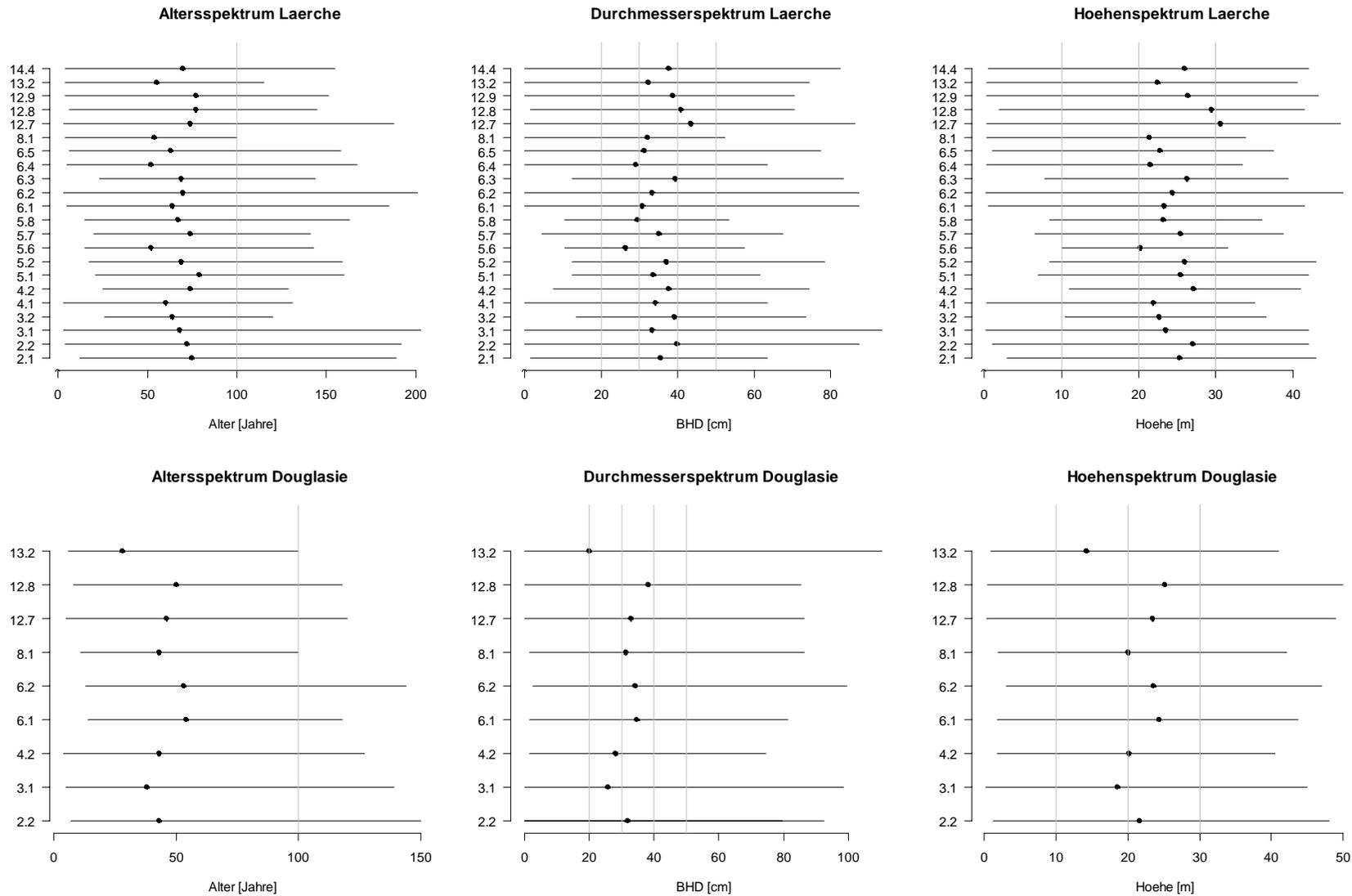


Abb. 14: Spannweite und Mittelwerte für das Alter, den Durchmesser sowie die Höhen für die Nadelbaumarten Lärche (oben) und Douglasie (unten) nach Wuchsbezirken. Die Datengrundlage bilden die Forstinventurdaten des Staatswaldes in Bayern.

Zusammenfassend lässt sich nach den Tabellen in Anhang 9.3 bzw. nach den Abbildungen 12 bis 14 feststellen, dass für weitergehende Auswertungen 309.060 Höhenmessbäume der Nadelbaumarten Fichte, Tanne, Kiefer, Lärche und Douglasie sowie für die Laubbaumarten Buche und Eiche in dieser Arbeit Verwendung finden. Dies sind lediglich 29,52 % der 1.046.930 in der Zentralen Forsteinrichtungsdatenbank zum 01.03.2005 gehaltenen Datensätze für Höhenmessbäume, die als Bäume der Oberschicht oder als Höhenmessbäume aus ungeschichteten Beständen gekennzeichnet waren. Die große Reduktion ist insbesondere eine Folge der strengen Lageprüfung auf Forsteinrichtungsebene (vgl. Abb. 11), nach der ganze Forsteinrichtungen unberücksichtigt blieben, wenn nicht mindestens 95 % der Probekreise bezüglich ihrer räumlichen Lage ordnungsgemäße Koordinateneintragungen in der Tabelle „fdinvkrs“ der Zentralen Forsteinrichtungsdatenbank hatten.

Die weiterhin verwendeten Daten entstammen den Wuchsgebieten 2 bis 14, wobei nicht für alle Baumarten in den Wuchsbezirken die festgelegte Mindestanzahl von 100 Höhenmessbäumen gefunden werden konnte. Aufgrund der eingeführten Kriterien für die Datenvorauswahl (vgl. Abb. 11) konnten für die Wuchsgebiete 7 („Fränkisches Tertiärhügelland“) sowie 10 („Oberpfälzer Wald“) keine weitergehenden Auswertungen durchgeführt werden. Für das Wuchsgebiet 10 konnten deswegen keine weitergehenden Auswertungen durchgeführt werden, weil, trotz einer großen Zahl an Datensätzen, zu wenige als Höhenmessbäume markierte Bäume in der Zentralen Forsteinrichtungsdatenbank gefunden werden konnten.

Die geringen Unterschiede bezüglich mittlerer Durchmesser und Höhen resultieren z. T. aus der Berücksichtigung von Forsteinrichtungen mit unterschiedlichem Laufzeitbeginn bzw. Laufzeitende. Dies ist verbunden mit z. T. unterschiedlichen Aufnahmeanweisungen. Beispielsweise wurden für die Wuchsbezirke 5.1 bis 5.3 nach der Datenvorauswahl lediglich ältere Forsteinrichtungen berücksichtigt, bei denen nur Bäume als Bäume der Oberschicht gekennzeichnet wurden, wenn diese sich im äußersten Messkreis befanden.

3.1.6 Datenvorverarbeitung

Für weitergehende Auswertungen wird im Rahmen dieser Arbeit ein Data Warehouse aufgebaut.

Sowohl bei der Aufnahmedatenbank als auch bei der Zentralen Forsteinrichtungsdatenbank der Bayerischen Staatsforsten handelt es sich um transaktionale Systeme, wobei der transaktionale Charakter bei der Aufnahmedatenbank stärker

ausgeprägt ist. Diese Systeme sind dazu gedacht, einen hohen Durchsatz an Transaktionen zu verarbeiten, die normalerweise kleine Veränderungen an den operativen Daten der Organisation vornehmen, d. h. dass es sich um Daten handelt, die ein Unternehmen zur Erledigung der täglichen Arbeit braucht. Systeme dieser Art werden als Online Transaction Processing-Systeme bezeichnet (Connolly et al., 2002).

Entscheidungsträger eines Unternehmens benötigen allerdings Zugriff auf alle entscheidungsrelevanten Daten, egal wo sich diese befinden. Umfassende Analysen der Organisation setzen zudem nicht nur den Zugriff auf die aktuellen Werte sondern auch auf Daten der Vergangenheit voraus. Um Analysen dieser Art zu vereinfachen, wurde das Data Warehouse geschaffen, welches Daten aus verschiedenen Quellen sowie historische und z. T. zusammenfassende Darstellungen enthält (Connolly et al., 2002). Inmo (zitiert nach Bauer und Günzel, 2004) definiert den Begriff Data Warehouse als „(...) subject-oriented, integrated, time-varying, non-volatile collection of data that is used primarily in organizational decision making“. Ein Data Warehouse ist demnach durch die vier Merkmale „Themenorientierung“, „Integration“, „Zeit-Raum-Bezug“ und „Nicht-Volatilität“ gekennzeichnet. In Anhalt an Mantel (2000) werden diese Merkmale nachstehend kurz erläutert. Unter dem Merkmal „Themenorientierung“ versteht man eine datenorientierte Vorgehensweise bei der Konzeption eines Data Warehouse. Im Gegensatz zu operativen Anwendungssystemen, bei deren Konzeption innerbetriebliche Abläufe und Funktionen maßgeblich sind, eröffnet diese Vorgehensweise dem Entscheider die Möglichkeit, Unternehmensdaten aus verschiedenen Blickwinkeln zu betrachten. Weiterhin sind in einem Data Warehouse alle Daten integriert. Da diese Daten aus verschiedenen, heterogenen Quellen stammen, sind Datenredundanzen und damit Inkonsistenzen unvermeidbar. In operativen Systemen werden jeweils nur die aktuell gültigen Datenwerte gespeichert. Unter dem Merkmal des „Zeit-Raum-Bezuges“ versteht man hingegen, dass sämtliche Daten zu einem bestimmten Zeitpunkt gültig bzw. gültig gewesen sind. Data Warehouses speichern historische Daten, die von Entscheidungsträgern benötigt werden, um Trends und Entwicklungen zu erkennen und angemessen reagieren zu können. Der Begriff der „Volatilität“ beschreibt den Grad, mit dem sich Daten im Laufe der normalen Nutzung ändern. Während in operativen Systemen laufend Daten eingegeben werden, werden alle Daten in einem Data Warehouse nach der einmaligen Übernahme aus dem Quellsystem normalerweise nicht mehr verändert. Es existieren daher nur zwei grundlegende Operationen auf den Daten in einem Data Warehouse – der initiale Ladevorgang und der ausschließliche lesende Zugriff auf die Daten. Durch diese Nicht-Volatilität lassen sich sämtliche Analysen jederzeit reproduzieren.

Zum Aufbau des Data Warehouse für diese Arbeit wurde die Zentrale Forsteinrichtungsdatenbank der Bayerischen Staatsforsten zum 01.03.2005 kopiert und, nur mit Leserechten versehen, auf einem MS-SQL-Server-2000 basierten Datenbanksystem neu angelegt. Dies stellt die Non-Volatilität des Systems sicher. Weiterhin wurden in das Data Warehouse Geodaten aufgenommen, mit deren Hilfe über die Verbindung mit den Forstinventurdaten Rückschlüsse auf die maßgeblichen Triebkräfte des Waldwachstums in Bayern gezogen werden sollen. Die Verbindung der Forstinventurdaten und der Geodaten erfolgt über Spatial-Joins. Hierbei handelt es sich um feste Beziehungen zwischen den verschiedenen Sachdaten-Tabellen, wobei die Sachdaten des Flächenthemas (z. B. Klima- oder Witterungsinformationen) an die Sachdaten des Punkthemas (Probekreisdaten bzw. zugehörige Baumdaten) angehängt werden.

In Tabelle 4 sind die Geodaten eingetragen, die mit den Probekreisinformationen für diese Arbeit verbunden wurden.

Abschließend ist in Abbildung 15 ergänzend zu Tabelle 4 dargestellt, welche wachstumsbestimmenden Informationen mit den Probekreisinformationen verbunden worden sind.

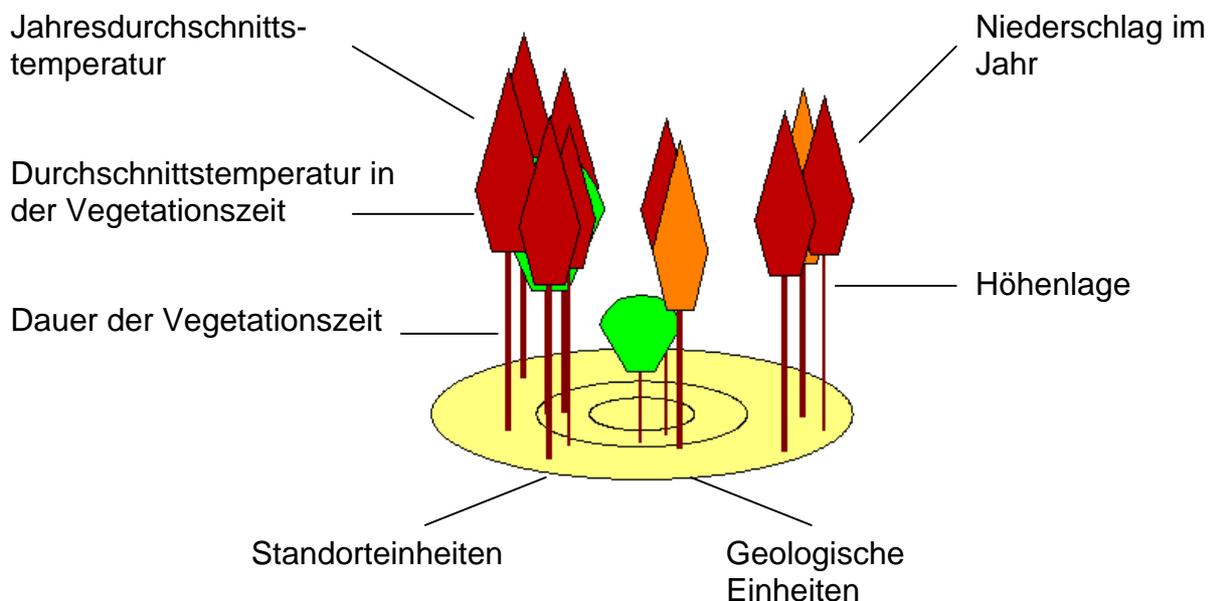


Abb. 15: Wachstumsbestimmende Triebkraftdaten, die im Rahmen dieser Arbeit mit den Probekreisdaten der Inventur verbunden wurden.

Tab. 4: Geodaten, die im Rahmen dieser Arbeit mit den Inventurpunktdaten verbunden wurden.

Nr.	Bezeichnung	Information	Skalen-niveau	Herkunft
1	Digitalisierte Standortkartierung für die Forstämter der ehemaligen Bayerischen Staatsforstverwaltung (Stand: 28.04.2005)	Standorteinheiten (Ziffern 1-3, Substrat, Trophie, Wasserhaushalt)	nominal	Bayerische Landsanstalt für Wald und Forstwirtschaft
2	Digitalisierte Karte der Wuchsgebietsgliederung für Bayern	Wuchsgebiets-, Wuchsbezirks- bzw. Teilwuchsbezirkszuordnung	nominal	Bayerische Landsanstalt für Wald und Forstwirtschaft
3	Digitalisierte geologische Übersichtskarte für Bayern	Geologische Einheiten mit ihren Eigenschaften „Stratigraphie“, „Petrographie“ und „Genese“	nominal	Bayerisches Geologisches Landesamt
4	Digitalisierte Karte "Relief/ Höhe über NN" (Karte 1)	klassenweise Höhenschichtinformationen (Auflösung in 100m Höhenstufen)	ordinal	Klimaatlas für Bayern (Projekt BayFORKLIM)
5	Digitalisierte Karte "Lufttemperatur/ Jahr" (Karte 2)	klassenweise Informationen zur mittleren Lufttemperatur/ Jahr (Mittelwerte 1951 - 1980)	ordinal	Klimaatlas für Bayern (Projekt BayFORKLIM)
6	Digitalisierte Karte "Lufttemperatur/ Vegetationsperiode" (Karte 7)	klassenweise Informationen zur Lufttemperatur in der Vegetationsperiode (Mittelwerte 1951 - 1980)	ordinal	Klimaatlas für Bayern (Projekt BayFORKLIM)
7	Digitalisierte Karte "Andauer einer Temperatur von mindestens 10° C/ Tag" (Karte 17)	klassenweise Information zur Anzahl der Tage mit einer Tagesdurchschnittstemperatur von mindestens 10°C/ Tag	ordinal	Klimaatlas für Bayern (Projekt BayFORKLIM)
8	Digitalisierte Karte "Niederschlagssumme/ Jahr" (Karte 25)	klassenweise Information zur mittleren Niederschlagssumme/ Jahr (Mittelwerte 1961-1990)	ordinal	Klimaatlas für Bayern (Projekt BayFORKLIM)

3.1.7 Transformation der Daten

Die Verbindung der Tabellen des Data Warehouse mit den Daten der Forstinventur wurde mit Hilfe von Sichten vorgenommen. Diese Vorgehensweise erspart den notwendigen Ladeschritt der Forstinventurdaten in das Data Warehouse und vermeidet eine redundante Datenhaltung.

Das aufgebaute Data Warehouse wurde in einem Transformationsschritt dahingehend erweitert, dass jeder einzelne Baum mit einer Klassenzuordnung versehen wurde, die angibt, ob der Baum bezüglich eines mittleren Wachstumsganges in dieser Region ein über- oder unterdurchschnittliches Wachstum mit Blick auf seine Höhe aufweist.

Zur Zuweisung der Klassen wurden alle ausgewählten Höhenmessbäume über den gegebenen Inventuraltern mit ihren Höhen aufgetragen. Die so entstandene Datenmenge wurde mit Hilfe der Wachstumsfunktion von Chapman-Richards (vgl. Formel 1) nichtlinear regressionsanalytisch ausgeglichen. Alle Bäume, die oberhalb der Ausgleichsline zu liegen kamen wurden als überdurchschnittlich hohe Bäume angesehen und mit der Klassenbezeichnung 1 versehen. Alle Bäume unter oder auf der Ausgleichsline wurden als unterdurchschnittliche hohe Bäume angesehen und mit der Klassenbezeichnung 2 versehen. Abbildung 16 zeigt diese Klassenzuordnung noch einmal grafisch am Beispiel der Fichteninventurdaten für den Wuchsbezirk 4.2 „Südliche Fränkische Platte“.

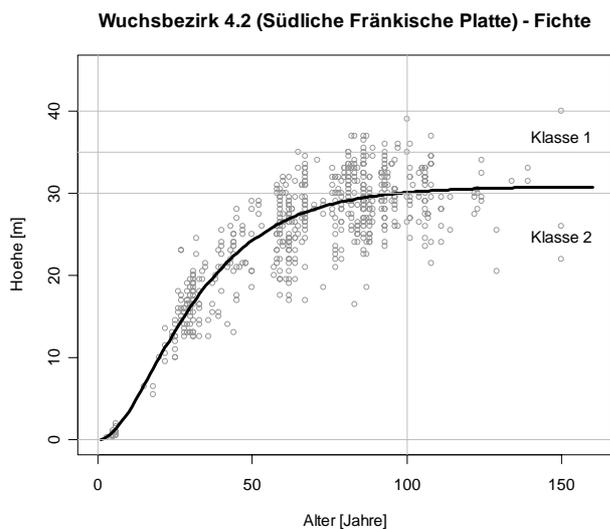


Abb. 16: Schematische Darstellung der Zuordnung der Höhenmessbäume zu zwei Klassen am Beispiel der Inventurdaten für die Baumart Fichte im Wuchsbezirk 4.2 („Südliche Fränkische Platte“)

Die vorgenommene Einteilung nach zwei Klassen generiert eine Verteilungsinformation, nach der Aussagen getroffen werden können, ob Baumarten auf bestimmten Standorten o. ä. signifikant unterschiedliche Verteilungsmuster im Hinblick auf die eingeführte, datenhalbiende Linie aufweisen oder nicht.

Die Beschränkung auf lediglich zwei Klassen wurde eingeführt, da a priori keine Information vorhanden ist, wie viele Klassen für einen optimalen Informationsgewinn anzulegen sind.

3.2 Methoden

3.2.1 Data-Mining

3.2.1.1 Methoden des Data-Mining im Überblick

Nachfolgend werden die Methoden des Data-Mining im engeren Sinne (Assoziationsmethoden, Segmentations- oder Clustermethoden, Diskriminanz- oder Klassifikationsmethoden sowie Regression) in Anhalt an Küsters (2001), Krahl et. al (1998), Borgelt und Kruse (1998) sowie Witten und Frank (2005) kurz beschrieben, wobei ein Schwerpunkt auf die Theorie zu Entscheidungsbäumen gelegt wird. Abgeschlossen wird dieser Textabschnitt mit Kapitel 3.2.1.5, in dem gängige Entscheidungsbaumalgorithmen dargestellt werden.

3.2.1.2 Assoziationsmethoden

Die seit Anfang der 90er Jahre vor allem von Agrawal et al. (1994) entwickelten Assoziationsmethoden gehören zu den wenigen originären Verfahren des Data-Mining. Ziel der Assoziationserkennung ist die Aufdeckung von strukturellen Zusammenhängen in Datenbanken. Diese weisen verallgemeinert folgende Form auf:

Wenn A und B, dann folgt mit Wahrscheinlichkeit $conf(idence)\%$ die Aussage C. Diese Assoziation gilt mit einer Wahrscheinlichkeit von $sup(port)\%$, d. h. $sup(port)\%$ aller Fälle weisen die Ereignisse A, B, und C auf.

Assoziationsmuster werden im Marketing v. a. für die Warenkorbanalyse verwendet, um „interessante“ Transaktionsmuster aufzudecken. Ein vielzitiertes Beispiel hierfür ist: Männliche Käufer (A) von Chips (B) kaufen mit einer Wahrscheinlichkeit von xx % auch Bier (C). Derartige Informationen werden u. a. bei der optimalen Supermarkt-Layout-Planung, beim Design von Versandhauskatalogen usw. genutzt.

Im Grunde ist das Auffinden von Assoziationsregeln nicht schwierig, da man im Prinzip eine Lösung durch Enumeration finden kann. In großen Datenbanken besteht das reale Problem jedoch in der kombinatorischen Explosion des Lösungsraumes und den damit verbundenen Anforderungen an die Rechnerkapazitäten.

Vielfach zitierte moderne Algorithmen für Assoziationsanalysen sind Apriori sowie AprioriTid von Agrawal et. al. (1994). Für den erstgenannten Algorithmus liegt eine frei verfügbare JAVA-Implementierung in WEKA (Abk. für „Waikato Environment for Knowledge Achievement“) der University of Waikato vor (Witten und Frank, 2005),

für letzteren liegt darüber hinaus eine frei verfügbare JAVA-Variante im KDD-Framework YALE (Abk. für „Yet Another Learning Environment“) der Universität Dortmund vor (<http://yale.cs.uni-dortmund.de>).

3.2.1.3 Segmentations- und Clustermethoden

Bei der Clusterbildung liegt die Klassenzugehörigkeit eines Objektes nicht fest sondern wird durch das Verfahren festgelegt bzw. definiert. Die Aufgabenstellung der Clusteranalyse besteht damit in der explorativen Konstruktion von unterschiedlichen Klassen (Cluster, Gruppen), die sich möglichst ähnlich (homogen) sind, sich aber ansonsten möglichst stark voneinander unterscheiden.

Methodisch besteht nach Küsters (2001) jede Clusteranalyse aus zwei Schritten:

1. Zunächst berechnet man mit Hilfe von Ähnlichkeits- oder Unähnlichkeitsmaßen auf der Basis aller Variablen $x_{i1}, x_{i2}, \dots, x_{ip}$ Abstände (Distanzen) zwischen allen Objekten einer Stichprobe. Das Hauptdifferenzierungsmerkmal zwischen den diversen Maßen ist dabei das Skalenniveau der Variablen.
2. Anschließend werden die Objekte mit Hilfe von Algorithmen verschiedenen, in sich möglichst homogenen Teilgruppen zugeordnet. Manche Verfahren verfolgen dabei nicht die Maximierung der Homogenität der einzelnen Gruppen, sondern auch die Maximierung der Unterschiede (Heterogenität) zwischen den Gruppen.

Distanz- und Ähnlichkeitsmaße lassen sich sowohl für metrische als auch für ordinale und nominale Merkmale definieren. Bei metrischen Variablen dominieren Distanzmaße auf der Grundlage geometrischer Abstandskonstrukte wie der euklidischen Distanz oder der Blockmetrik (Manhattan-Distanz). Bei nominalskalierten Variablen werden hingegen primär Übereinstimmungsmaße herangezogen, bei denen man die Übereinstimmungen der Ausprägungen bei den verschiedenen Merkmalen durch Auszählung zugrunde legt. Dabei muss bei der Auswahl des Maßes darauf geachtet werden, ob das Nichtvorhandensein eines Merkmales bei zwei Objekten bereits eine Indikation für Ähnlichkeit ist oder nicht. Distanzmaße D_{ij} und Ähnlichkeitsmaße A_{ij} zwischen zwei Objekten i und j werden daher oft durch $A_{ij} = 1 - D_{ij}$ in einer normierten Weise in Beziehung gesetzt.

Die meisten Clusteralgorithmen benutzen eine Distanz-Matrix D oder eine Ähnlichkeitsmatrix A zwischen allen Objekten $i, j = 1..N$ als Grundlage. Damit wächst die Matrix quadratisch mit der Beobachtungszahl.

Clusteranalyseverfahren lassen sich im Wesentlichen in hierarchische und in partitionierende Verfahren einteilen.

Bei hierarchischen Verfahren dominieren die agglomerativen Verfahren, bei denen die einzelnen Objekte sequentiell und paarweise zu Klassen zusammengefasst werden, so dass ein hierarchischer binärer Baum entsteht. Dabei benötigt man zur Zusammenfassung von Clustern auch Funktionen zur Berechnung der Ähnlichkeiten bzw. Distanzen zwischen zwei Klassen und zwischen einer Beobachtung und einer Klasse. Eine einfache Technik ist das „single-linkage“, bei dem die Distanz zwischen zwei Clustern durch den Minimalabstand von zwei Objekten berechnet wird. Dadurch neigt diese Methode zur Brückenbildung zwischen Clustern. Eine von vielen weiteren Alternativen ist das „complete-linkage“, bei dem die Distanz zwischen zwei Clustern durch den Maximalabstand ihrer Objekte bestimmt wird.

Als Gegensatz zu agglomerativen Verfahren wurden divisive Clustermethoden entwickelt, bei denen die Stichprobe sukzessiv in Teilgruppen zerlegt wird. Agglomerative und divisive Verfahren weisen den Nachteil auf, dass die Bäume mit wachsendem Stichprobenumfang exponentiell anwachsen. Daher wurden als Alternative nichthierarchische, partitionierende Methoden entwickelt, bei denen man zumindest bei den einfachen Varianten zunächst von einer festen Klassenzahl ausgeht. Für metrische Variablen dominiert hier die k-Means-Clusteranalyse, bei der die Klassen durch Minimierung der euklidischen Distanz zu Gruppenmittelwerten gebildet werden.

Traditionelle Verfahren der Clusteranalyse unterstellen, dass ein Objekt genau einer Klasse zugeordnet werden kann. Bei einer hohen Homogenität innerhalb einzelner Gruppen und einer hohen Heterogenität zwischen den Gruppen ist dies aus praktischer Sicht nicht problematisch. In der Praxis findet man aber immer wieder Objekte, die sich nur sehr schwierig einer Gruppe zuordnen lassen. Der Einsatz eines streng partitionierenden Verfahrens führt hier u. U. dazu, dass relativ inhomogene Gruppen entstehen können. Ein Ausweg ist die unscharfe Clusteranalyse (fuzzy clustering), bei der den einzelnen Objekten Zugehörigkeitsgrade bzw. Wahrscheinlichkeiten der Klassenzugehörigkeiten zugeordnet werden. Ein weiteres gängiges und mit der Clusteranalyse eng verwandtes Verfahren ist die multidimensionale Skalierung (MDS). Diese Technik, die primär eine Visualisierungsmethode darstellt, bildet Ähnlichkeiten bzw. Distanzen zwischen Objekten typischerweise auf ein zweidimensionales Koordinatensystem ab. Ziel der diversen MDS-Algorithmen, die nahezu vollständig auf den gleichen Ähnlichkeits- und Distanzmaßen wie die Clusteranalyse beruhen, ist es, Objekte aus

einem hochdimensionalen Merkmalsraum (d. h. mit einer hohen Anzahl von Merkmalen) so auf die X-Y-Ebene abzubilden, dass im Ursprungsraum unähnliche (distante) Objekte eine hohe Distanz in der X-Y Darstellungsebene ausweisen. Damit kann man mit der multidimensionalen Skalierung ähnliche und entfernte Objekte durch eine visuelle Inspektion identifizieren.

Bei der Anwendung im Marketing, die – wie eingangs der Arbeit erwähnt – einen Anwendungsschwerpunkt von Data Mining Methoden darstellt, werden Clusteranalysen vor allem zur Kundensegmentation verwendet (Wilde, 2005).

An dieser Stelle sei erwähnt, dass in der Literatur der Mustererkennungsverfahren die Schätzung von Diskriminanzfunktionen, wie sie im nachfolgenden Textblock beschrieben wird, als überwachtes Lernen (supervised learning) bezeichnet wird, da der Ausgang der geschätzten Diskriminanzfunktion mit der tatsächliche realisierten Gruppenzugehörigkeit verglichen werden kann. Die Schätzung von Clustern (Klassenzugehörigkeiten) wird hingegen als nichtüberwachtes Lernen (unsupervised learning) bezeichnet, da keine Referenzklasse bekannt ist.

3.2.1.4 Diskriminanzanalyse (Klassifikation)

Die grundlegende Fragestellung der Diskriminanzanalyse besteht darin, aufgrund einer Menge von p Variablen $x_{i1}, x_{i2}, \dots, x_{ip}$ auf die Klassenzugehörigkeit $c = 1, \dots, C$ eines Objektes zu schließen. Ein vielfach in der Data-Mining-Literatur zitiertes Beispiel ist das Kredit-Scoring (vgl. z. B. Wilde, 2005 oder Witten und Frank, 2005), bei dem man mit Hilfe von Variablen wie z. B. dem Monatseinkommen, Familienstand, Beruf, beantragter Kredithöhe usw. auf die Kreditwürdigkeit (im einfachsten Zweigruppenfall nur „Ja“ und „Nein“) schließt.

Die verschiedenen diskriminanzanalytischen Verfahren unterscheiden sich im Wesentlichen in der Art und Weise, wie die Diskriminanz-Funktion

$$c = d(x_1, x_2, \dots, x_p) : X_1 \times X_2 \times \dots \times X_p \rightarrow \{1, 2, \dots, C\}$$

als Zuordnungsfunktion von der i -ten Beobachtung $x_{i1}, x_{i2}, \dots, x_{ip}$ zur Klasse c geschätzt (kalibriert) wird. Man unterscheidet hierbei zwischen dem Stichprobenparadigma und dem diagnostischen Paradigma (Rudolph, 2006).

- a) Beim Stichprobenparadigma werden die gruppenspezifischen Verteilungen $f_c(x_1, x_2, \dots, x_p)$ für $c \in \{1, 2, \dots, C\}$ der Variablen (x_1, x_2, \dots, x_p) geschätzt und anschließend mit Hilfe des Bayesschen Theorems in eine Diskriminanzfunktion nach dem Prinzip der maximalen Posterior-Dichten oder minimalen Kosten (Risikofunktionen) umgesetzt.

- b) Nach dem diagnostischen Paradigma wird die Diskriminanzfunktion $d(x_1, x_2, \dots, x_p)$ hingegen direkt geschätzt, so dass dieser Ansatz unmittelbar zur Schätzung eines Regressionsmodells für eine nominalskalierte abhängige Variable $y_i \in \{1, 2, \dots, C\}$ als Funktion von p Regressoren $x_{i1}, x_{i2}, \dots, x_{ip}$ korrespondiert.

Küsters (2001) weist weiter daraufhin, dass jede Form der Diskriminanzanalyse lediglich eine nach statistischen Kriterien (kleinste erwartete Fehlerrate, minimale erwartete Kosten etc.) optimale Diskriminationsvorschrift liefert. Die Gründe, warum man in bestimmten Fällen zu einer konkreten Klassifikation kommt, sind aus inhaltlicher Sicht nicht immer nachvollziehbar. Dies gilt insbesondere für moderne Verfahren der nichtlinearen Diskriminanzanalyse, die bezüglich ihrer Klassifikationsgüte allerdings in vielen Fällen klassischen Verfahren überlegen sind.

Nachfolgend wird wegen der besonderen Bedeutung für diese Arbeit auf eine Sonderform der nichtlinearen Diskriminanzanalyse, die Klassifikations- und Regressionsbäume, näher eingegangen. Die theoretischen mathematischen Grundlagen hierzu gehen insbesondere zurück auf Breiman et al. (1984).

Die Grundidee dieser Methode besteht darin, dass die zugrundeliegende Stichprobe auf der Grundlage der Klassenzugehörigkeiten $y_i \in \{1, \dots, C\}$ und der Regressoren $x_{i1}, x_{i2}, \dots, x_{ip}$ sukzessiv mit Hilfe von Partitionierungsregeln wie $x_{ij} < \tau_j$ oder $x_{ij} \geq \tau_j$ mit τ_j als Schwellenwert sequentiell in immer kleiner werdende Teilsegmente zerlegt wird, bis auf der untersten Ebene weitgehend homogene Gruppen entstehen. Borgelt und Kruse (1998) bezeichnen dies auch als „Divide and conquer-Verfahren (...) mit gieriger (greedy) Auswahl der zu testenden Attribute“. Jeder Schwellenwert τ_j ist dabei eine zur Variablen x_{ij} gehörende und damit zu schätzende Größe. Bei nominalskalierten Variablen bestimmt man anstelle eines Schwellenwertes eine Partition des Wertebereichs in zwei oder mehrere disjunkte Teilmengen.

In einem nachfolgenden Stützschritt (pruning) wird anschließend der Baum auf eine sinnvolle Größe reduziert, um die mit dem „overfitting“ verbundene hohe Fehlerrate bei der Anwendung des impliziten Prognosemodells auf neue Daten weiter zu reduzieren.

Ein Klassifikationsbaum kann somit als ein diskriminanzanalytisches Verfahren mit Erklärungskomponente interpretiert werden, in dem die einzelnen Segmentationschritte auch für statistische Laien nachvollziehbar sind. Dementsprechend werden Klassifikationsbäume im Kontext des Data Mining vor allem zur automatischen Generierung von Regeln verwendet. Aufgrund dieser Tatsache werden die Schätzverfahren für Klassifikationsbäume auch als

Regellernverfahren bezeichnet. Eng verwandt sind Verfahren des maschinellen Lernens, in denen aufgrund einer Datenbasis, die alle möglichen Fälle enthält, das unterliegende Regelwerk ermittelt wird. Die theoretischen Grundlagen hierzu gehen insbesondere zurück auf Quinlan (1986).

Regressionsbäume unterscheiden sich von Klassifikationsbäumen dahingehend, dass anstelle einer nominalskalierten Variablen eine metrisch skalierte Variable y_i als Zielvariable vorliegt. Ziel des Verfahrens ist die Segmentierung der Stichprobe in Teilgruppen, die sich durch eine möglichst geringe Variation von y_i auszeichnen. Theoretische Grundlagen hierzu finden sich bei Breiman et al. (1984). Diese Verfahrensklasse findet im weiteren Verlauf der Arbeit keine Anwendung und wird daher nicht näher beschrieben.

Die im Marketing gebräuchlichsten Klassifikationsbäume sind das von Kass (1980) stammende CHAID-Verfahren (CHAID = CHI-squared Automatic Interaction Detection), das von Quinlan (1993) beschriebene C4.5-Verfahren bzw. dessen Vorläufer ID3 (Quinlan, 1986) sowie das von Breiman et al. (1984) innerhalb der Statistik entwickelte CART-Verfahren (CART = Classification and Regression Trees). Diese Algorithmen werden auf das aufgebaute Data Warehouse angewendet und daher im folgenden Kapitel detaillierter in Anhalt an Borgelt und Kruse (1998), Ferber (2003), Lusti (2002) sowie Krahl et al. (1998) beschrieben.

3.2.1.5 Beschreibung der Algorithmen CHAID, ID3, C4.5 und CART

Beim **CHAID**-Verfahren wird zur Attributsauswahl der CHI-Quadrat-Unabhängigkeitstest verwendet (Krahl et al., 1998). Die hier berechnete Kennzahl ist der sog. CHI-Quadrat-Abstand, der umso grösser ist, je grösser die Abhängigkeit der betrachteten Variablen von der Zielvariablen ist. Ein wichtiges Merkmal hier ist, dass an den inneren Knoten auch mehr als zwei Verzweigungen entstehen können. Es können also nicht nur binäre Bäume generiert werden, wenn die Trennqualität durch die Bildung n-närer Bäume erhöht werden kann.

In Pseudocodenotation folgt der CHAID-Algorithmus folgenden Schritten:

- 1) Erstelle für jede unabhängige Variable (x) eine Kontingenztabelle mit den Kategorien der unabhängigen (x) Variablen als Zeilen und den Kategorien der abhängigen (y) Variablen (n) als Spalten.
- 2) Suche aus der Kontingenztabelle die „Untertabelle“ zweier unabhängiger Variablen, die am wenigsten signifikant verschieden sind. Falls die Signifikanz

dieser Teil-Tabelle einen durch einen CHI-Quadrat-Test $\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(h_{ij}^0 - h_{ij}^e)^2}{h_{ij}^e}$

bestimmten kritischen Wert nicht überschreitet, also je unabhängiger die Variable von der Zielgröße ist, fasse das Kategorienpaar zu einer Kategorie zusammen und wiederhole rekursiv den zweiten Schritt. Die Prüfgröße folgt dabei einer CHI-Quadrat-Verteilung mit $v=(r-1)(s-1)$ Freiheitsgraden.

- 3) Suche für jede der zusammengefassten Kategorien, die aus drei oder mehr ursprünglichen Kategorien bestehen, die signifikanteste binäre Partitionierung, durch die die Zusammenfassung (der ähnlichsten Kategorien) wieder aufgelöst werden kann. Falls eine kritische Schwelle überschritten wird bzw. je unabhängiger die Variable von der Zielgröße ist, implementiere die Partitionierung und kehre zu Schritt 2 zurück.
- 4) Berechne die Signifikanz aller optimal zusammengefassten unabhängigen Variablen und isoliere die signifikanteste. Ist letztere größer als der Schwellenwert, dann partitioniere die Daten an Hand der (zusammengefassten) Kategorien der so gewählten unabhängigen Variablen.
- 5) Für jede noch nicht analysierte Partitionierung gehe zurück zu Schritt 1.

ID3 ist ein von Quinlan (1986) entwickelter Algorithmus zur Generierung von Entscheidungsbäumen. Die Abkürzung ID3 steht dabei für Iterative Dichotomiser 3. ID3 gilt als Vorgänger des ebenfalls von Quinlan entwickelten C4.5-Algorithmus. Grundlegende Funktionsbeschreibungen zu ID3 finden sich bei Lusti (2002) sowie detaillierter bei Ferber (2003).

$A_i : D \rightarrow R_i, i = 1..n$ sei eine Menge vorhersagender Attribute mit endlichen Wertebereichen R_i auf einer endlichen Menge D von Beispielen und $A_0 : D \rightarrow R_0 = \{z_1..z_m\}$ ein vorherzusagendes Attribut. Im folgenden rekursiven Algorithmus bezeichne $N = (A, K, z)$ den aktuellen Knoten, wobei $A \in \{A_0, \dots, A_n\}$ das aktuelle Attribut, $K \subset D$ die Menge der Beispiele, die dem Knoten zugeordnet sind, und $z \in R_0$ eine Kategorie bezeichnen.

1. Ordne alle Beispiele einem (Wurzel-) Knoten zu:

$$N = (A, K, z) = N_{0,0} = (A, K_{0,0}, z), K_{0,0} = D$$

2. Falls alle Tupel aus K derselben Kategorie angehören, falls also gilt $A_0(d_i) = z_k \forall d_i \in K$ und ein $z_k \in R_0$, setze $A := A_0$ und $z = z_k$. Andernfalls wähle ein Attribut $A_i : D \rightarrow R_i, i \in \{1..n\}$ das in K mindestens zwei verschiedene Werte annimmt ($d_1, d_2 \in K$) mit $A_i(d_1) \neq A_i(d_2)$ als aktuelles Attribut für den Knoten N und füge für jeden Wert $r_j \in R_i$ mit $A_i^{-1}(\{r_j\}) \cap K \neq \emptyset$ des Attributs

einen Kindknoten $N_{i,r_j} = (A, K_{i,r_j})$ mit unbestimmtem aktuellem Attribut A und $K_{i,r_j} = A^{-1}(\{r_j\}) \cap K$.

3. Falls ein unbearbeiteter Kindknoten $N_{i,r_j} =$ von N existiert, setze $N := N_{i,r_j}$. Andernfalls prüfe, ob es einen unbearbeiteten Geschwisterknoten N_{i,r_k} von N gibt und setze in diesem Fall $N := N_{i,r_k}$. Ist das auch nicht der Fall, setze N auf den Elternknoten von N .
4. Falls $N \neq N_{0,0}$ gilt, gehe zu Schritt 2, andernfalls beende den Algorithmus.

In Schritt 2 muss ein Attribut $A_i : D \rightarrow R_i$ ausgewählt werden, das zur Selektion verwendet werden soll. Für jedes Attribut, das in Frage kommt, wird der Wert

$$E(A_i, K_{j,r_m}) = \sum_{k \in R_i} \frac{|A_i^{-1}(\{k\}) \cap K_{j,r_m}|}{|K_{j,r_m}|} * I(k, K_{j,r_m}) \quad (\text{Formel 8})$$

berechnet. Es gilt hierbei:

$$I(k, K) = \sum_{r \in R_0} -q_{k,r} \ln(q_{k,r}) \quad (\text{Formel 9})$$

$$\text{mit } q_{k,r} = \frac{|A_0^{-1}(\{r\}) \cap A_i^{-1}(\{k\}) \cap K|}{|A_i^{-1}(\{k\}) \cap K|} \quad (\text{Formel 10})$$

$q_{k,r}$ gibt also den Anteil der Tupel aus der Kategorie r unter den Tupeln des Knotens an, bei denen das Attribut A_i den Wert k annimmt. Der Entropiewert $I(k, K)$ gibt damit an, wie durchmischt die Tupel des Knotens in Bezug auf das Attribut A_0 , also in Bezug auf die gesuchte Kategorisierung, sind.

Es wird das Attribut zur Selektion gewählt, bei dem der Wert $E(A_i, K_{j,r_m})$ minimal ist, bei dem die Kindknoten also bezüglich der gesuchten Kategorie möglichst wenig durchmischt sind. Der Wert $E(A_i, K_{j,r_m})$ (vgl. Formel 8) wird dabei als Information-Gain-Kriterium bezeichnet.

Die Auswahl dieser Optimierungsheuristik ist folgendermaßen motiviert: Betrachtet man die Beispiele eines Knotens als Informationsquelle über die Zugehörigkeit zu den Zielkategorien, dann gibt Formel 9 eine Abschätzung des Informationsgehalts bzw. der Entropie eines Kindknotens an, d. h. des mittleren Informationsgewinns,

den das Inspizieren eines Beispiels aus der Menge bringt. Stammen fast alle Beispiele der Knotenmenge aus einer Kategorie, ist der zu erwartende Informationsgewinn gering. Stammen sie sogar alle aus einer Kategorie, ist entweder $q_{k,r} = 0$ oder $q_{k,r} = 1$ und damit $\ln(q_{k,r}) = \ln(1) = 0$. In beiden Fällen ist das Produkt $-q_{k,r} * \ln(q_{k,r})$ und damit $I(k, K)$ gleich 0. In den anderen Fällen ist der Wert positiv, weil das Argument des Logarithmus zwischen 0 und 1 liegt und der Logarithmus daher negativ ist. Formel 9 gibt also den mittleren erwarteten Informationsgehalt der Kindknoten bei einer Aufteilung nach Attribut A_i an. Je kleiner der ist, desto größer ist der Informationsgewinn durch die Wahl von A_i .

Wie bereits angeführt, handelt es sich bei dem als **C4.5**-bezeichneten **Algorithmus** um eine Weiterentwicklung des ID3-Algorithmus. Gründe für die Weiterentwicklung sind in folgenden Schwachpunkten des ID3-Algorithmus zu sehen:

- Das Information-Gain-Kriterium bevorzugt bei der Auswahl von Merkmalsvariablen für Merkmalsknoten Variablen mit vielen Ausprägungen. Diese Eigenschaft kann die Induktion eines unnötig komplexen Entscheidungsbaumes oder Regelwerkes zur Folge haben.
- Viele metrische Merkmalsvariablen werden nicht auf diskreten sondern auf stetigen Skalen gemessen. ID3 beinhaltet keine Prozedur zur Handhabung (Diskretisierung) stetiger Merkmale.
- ID3 tendiert zur Induktion von Bäumen, die die Objekte ihrer Lernstichproben fehlerfrei klassifizieren. Dies ist insbesondere bei Vorliegen von Lernstichproben, die durch eine große Variablenzahl relativ zum Stichprobenumfang gekennzeichnet sind, der Fall. Witten und Frank (2005) weisen für letzteren Fall darauf hin, dass eine fehlerfreie oder fast fehlerfreie „In-the-sample-Klassifikation“ häufig ein Indiz für ein überspezialisiertes Regelwerk ist und dass „Out-of-the-sample“ eine hohe Fehlerrate aufweist (Overfitting-Problem).

Der C4.5 Algorithmus verwendet anstelle des Information-Gain-Kriteriums $G(S(T_i), X_i)$ das sog. Gain-Ratio-Kriterium zur Auswahl des Merkmals X_i für einen Merkmalsknoten, welches mathematisch wie folgt formuliert wird:

$$GR(S(T_i), X_i) = \frac{G(S(T_i), X_i)}{SI(S(T_i), X_i)} \quad (\text{Formel 11})$$

Hierbei ist die sog. Splitinformation

$$SI(S(T_i), X_i) = - \sum_{j=1}^{m_i} \frac{|S(T_{t_j})|}{|S(T_i)|} * \log_2 * \frac{|S(T_{t_j})|}{|S_{T_i}|} \quad (\text{Formel 12})$$

ein Strafterm, der Merkmale mit vielen Ausprägungen bestraft.

Des weiteren besitzt C4.5 einen Mechanismus zur Diskretisierung stetiger Merkmale. X_i sei eine stetige Merkmalsvariable, die Werte aus dem kontinuierlichen Intervall $R = (a, b)$ annehmen kann. C4.5 überführt X_i durch die Bildung von m_i Messwertklassen

$$R_1 = (a, \tau_1), R_2 = (\tau_1, \tau_2), \dots, R_{m_i} = (\tau_{m_i-1}, b),$$

mit $R = R_1 \cup R_2 \cup \dots \cup R_{m_i}$

in eine diskrete Variable X_i^D , wobei die Meßwertklassen R_j mit $j=1, \dots, m_i$ mit den Ausprägungen von X_i^D gleichgesetzt werden. Hierzu müssen die Klassengrenzen t_j (Schwellenwerte, thresholds) geeignet festgelegt werden. Die Schwellenwerte werden dabei mit Hilfe des Gain-Kriteriums oder mit Hilfe des Gain-Ratio-Kriteriums (vgl. Formel 8 bzw. Formel 11) ermittelt. Die Schwellenwerte und ebenfalls die Anzahl der Schwellenwerte werden in einer Suchprozedur systematisch variiert; abschließend wird der Satz von Schwellenwerten selektiert, für den der Gain-Ratio-Wert den größten Wert annimmt. Die resultierende diskrete Variable X_i^D ist somit für die Lösung der Klassifikationsaufgabe „optimiert“.

Zur Vermeidung des Overfittings wendet C4.5 Pruning-Techniken an. Kenngröße zur Einschätzung der Klassifikationsperformance ist dabei die Kreuzvalidierung. Bei letzterer wird die Stichprobe S in zwei Teilstichproben zerlegt. Die erste Teilstichprobe wird als Trainingsmenge bezeichnet und dient der Induktion eines Entscheidungsbaumes, während anhand der Objekte der Validierungsmenge die Klassifikationseigenschaften des Baumes getestet werden. Kenngröße für die Güte der Klassifikation ist hierbei die sog. Fehlklassifikationsrate, die angibt wie häufig die Objekte eines Blattes korrekt bzw. inkorrekt klassifiziert wurden.

Grundidee des sog. Prunings ist die Reduktion der Komplexität der Entscheidungsbäume durch iteratives „Beschneiden“ ihrer Äste. Hierbei werden verschiedene Arten des Prunings unterschieden.

Beim Stop-Growing oder Forward-Pruning wird nach jedem Schritt des iterativen Baumkonstruktionsprozesses geprüft, ob der aktuelle Baum erweitert werden sollte oder nicht.

Beim Postpruning, welches in der Fachliteratur bisweilen auch als Backward-Pruning bezeichnet wird, wird nach der Induktion eines Entscheidungsbaumes aus den Trainingsdaten iterativ geprüft, ob der aktuelle Baum beschnitten werden sollte oder nicht.

Das Attributsauswahlmaß wird auch beim **CART-Algorithmus** (Breiman et al., 1984) durch die Maximierung des Informationsgehaltes gesteuert, wobei der Informationsgehalt eines Attributes hoch ist, wenn aus der Attributsausprägung – im Zusammenhang mit einem zu definierenden Schwellenwert des Informationsgehaltes – mit hoher Trefferquote eine Klassifikation vorgenommen werden kann. Der Informationsbegriff basiert dabei auch hier auf dem mathematischen Ausdruck der Entropie. Dabei lässt sich aus der verallgemeinerten Entropie, die erstmals von Daroczy (1970) (zitiert nach Borgelt und Kruse, 1998) beschrieben wurde, die Shannon'sche Entropie (vgl. Formel 13 bzw. Formel 9) als Spezialfall ableiten.

$$H = \lim_{\beta \rightarrow 1} H_{gen}(\beta) = - \sum_{i=1}^n p_i * \ln_2 * p_i \quad (\text{Formel 13})$$

Die Shannon'sche Entropie ist umso kleiner, je höher der Informationsgehalt ist. Je höher der Informationsgehalt eines Attributes in Bezug auf die Zielgröße ist, desto „weiter oben“ im Entscheidungsbaum findet sich dieses Attribut. Da der Entropiewert mit einem vom Anwender vorgegebendem Schwellenwert verglichen wird, folgt daraus, dass der Wert nur kleiner oder größer als der Schwellenwert sein kann, was dazu führt, dass mit dem CART-Algorithmus berechnete Entscheidungsbäume nur binäre Bäume sein können.

3.2.2 Nichtlineare Regression

Regressionsanalytische Ansätze dienen nach Precht und Kraft (1993) primär zur Erklärung und Prognose einer abhängigen Variablen y_i durch p unabhängige Variablen $x_{i1}, x_{i2}, \dots, x_{ip}$ auf der Grundlage einer Stichprobe durch ein Modell der Form

$$y_i = f(x_{i1}, x_{i2}, \dots, x_{ip}, \beta, \varepsilon_i) \quad (\text{Formel 14})$$

mit ε_i als Fehlerterm, der die nicht modellierte bzw. modellierbare Variation erfasst. Der Parametervektor β repräsentiert dabei die Abhängigkeitsstruktur.

Die konkrete Form der Regressionsanalyse hängt in erster Linie von folgenden Merkmalen ab:

- 1) Der *Funktionsform von f* . Dabei unterscheidet man im Wesentlichen nur zwischen linearen und nichtlinearen Fällen; bei Vorliegen von ausschließlich metrisch skalierten Variablen handelt es sich dann um das klassische lineare Regressionsmodell und die klassische nichtlineare Regression.
- 2) Der *Skalierung der abhängigen Variablen y_i* , wobei üblicherweise nach metrischen, ordinalen und nominalen Merkmalen unterschieden wird. Für ordinale abhängige Variablen verwendet man typischerweise ordinale Logit- und Probitmodelle, für nominalskalierte Merkmale überwiegend multinominale Logitmodelle sowie – in einer speziellen Parametrisierung – loglineare Modelle.
- 3) Der *Skalierung der unabhängigen Variablen x_{ip}* , wobei meist nur zwischen metrischen und nominalskalierten Merkmalen unterschieden wird; bei letzteren verwendet man varianzanalytische Designmatrizen als Regressormatrizen. Der Standardfall ist die ein- und mehrfaktorielle Varianzanalyse mit und ohne Interaktionseffekte.
- 4) Der *Datengrundlage*, wobei im Wesentlichen zwischen kontemporären Querschnittsdaten und dynamischen Längsschnittdaten (Zeitreihen) sowie Kombinationen (Paneldaten) unterschieden wird.
- 5) Der *Verteilung (Form, Parametrisierung) des Fehlerterms ε_{ij}* , so dass man gegebenenfalls Streuungsvariationen (Heteroskedasdität) zwischen den Beobachtungen und Abhängigkeiten im Zeitverlauf (Autokorrelation) adäquat berücksichtigen kann.
- 6) Der *Schätzung des Parameters β i. V. m. der Funktionsform f sowie der Verteilungsfunktion des Fehlerterms ε_i* . Standardverfahren sind die lineare und nichtlineare kleinste-Quadrat-Schätzung und die Maximum-Likelihood-Schätzung, obwohl man für umfangreichere Datensätze auch nichtparametrische Methoden (robuste Regression, Kernschätzer) verwenden kann.

In der multiplen linearen Regression werden Funktionen f betrachtet, die linear sind in den Parametern β

$$y_i = f(x_{i1}, x_{i2}, \dots, x_{ip}, \beta_1, \beta_2, \dots, \beta_p) = \beta_1 \tilde{x}_1 + \beta_2 \tilde{x}_2 + \dots + \beta_p x_p \quad (\text{Formel 15})$$

wobei \tilde{x}_j beliebige Funktionen der ursprünglichen erklärenden Variablen x_j sein können.

In der nichtlinearen Regression werden hingegen Funktionen von f untersucht, die sich nicht als lineare Funktion in den Parametern schreiben lassen. Oft wird eine solche Funktion aus der Theorie abgeleitet. Es bestehen im Prinzip unbeschränkte Möglichkeiten, den deterministischen Teil des Modells anzusetzen.

Einige nichtlineare Regressionsfunktionen lassen sich durch Transformation der Zielgröße und der erklärenden Variablen linearisieren. So lässt sich beispielsweise die Potenzfunktion $f(\chi; \underline{\theta})$ zu einer in den Parametern linearen Funktion

$$\text{transformieren } \ln\langle f(\chi; \underline{\theta}) \rangle = \ln\langle \theta_1 \rangle + \theta_2 * \ln\langle \chi \rangle = \beta_0 + \beta_1 * \tilde{\chi} = \tilde{f}(\tilde{\chi}; \beta)$$

wobei $\beta_0 = \ln(\theta_1)$, $\beta_1 = \theta_2$ und $\tilde{\chi} = \ln(\chi)$ ist. Die Regressionsfunktion f ist demnach linearisierbar, wenn sie durch Transformationen der Argumente und eine monotone Transformation des Resultats in eine in den Parametern lineare Funktion verwandelt werden kann. Eine Linearisierung ist allerdings nur dann angebracht, wenn damit auch die Annahmen über die Zufallsabweichungen besser befriedigt werden können. Diese Annahmen werden dabei mit Hilfe der Residuen-Analyse geprüft.

Sind die Funktionen jedoch komplexer, so kann häufig durch Transformation keine Lösung mehr erreicht werden. Funktionen, die sich nicht transformieren lassen, sind zum Beispiel trigonometrische Funktionen und auch Exponentialfunktionen. Sollen gemessene Punktwolken an diese Art von Funktionen angepasst werden, muss eine der Methoden zur nichtlinearen Ausgleichsrechnung genutzt werden. Nach Precht und Kraft (1993) ist die Methodik der nichtlinearen Regression die folgende: Ausgehend von einer vorliegenden Schätzfunktion f werden aufgrund von a priori Informationen Startwerte für die Koeffizienten der Schätzfunktion vorgegeben. Mit bestimmten Iterationsverfahren, wie z. B. Gauss-Newton, Levenberg-Marquardt, Gradientenverfahren, Sekantenverfahren u. a., werden diese Anfangswerte der Koeffizienten so verändert, dass die Abweichungsquadratsumme der Fehler minimiert wird. Detailliert werden die Iterationsverfahren beschrieben bei Draper und Smith (1981). In jedem Iterationsschritt fordert man:

$$SQ_{\text{Res}}(b^{(0)} + k \cdot \Delta) < SQ_{\text{Res}}(b^{(0)}) \quad (\text{Formel 16})$$

Dabei ist $b^{(0)} = (b_0, b_1, \dots)'$ der Startvektor der Koeffizienten. Die Veränderung Δ des Koeffizientenvektors differiert je nach Methode. Ein Iterationsverfahren ist beendet, falls die Fehlerabweichungsquadratsumme nicht mehr verkleinert werden kann.

3.2.3 Anwendung von Data-Mining-Methoden zum Auffinden standörtlicher Muster in Forstinventurdaten

Im weiteren Verlauf der Arbeit sollen die in Abschnitt 3.2.1 beschriebenen Data-Mining-Methoden sowie die in Abschnitt 3.2.2 kurz beschriebene nichtlineare Regression angewendet werden, um innerhalb der Inventurdaten eines Wuchsbezirkes standörtlich möglichst homogene Gruppen zu isolieren. Diese Gruppen sollen eine regionale, standortbezogene Feinjustierung des Wachstumsmodells SILVA ermöglichen.

Aus der Vielzahl der Methodenklassen des Data-Mining sollen im weiteren Verlauf der Arbeit Klassifikationsmethoden zum Einsatz kommen. Die Datengrundlage stellt das aufgebaute Data Warehouse (vgl. Abschnitt 3.1.6) dar. Grundlage für die Klassifikation soll hierbei die in Abschnitt 3.1.4 eingeführte Klasseneinteilung sein. Mit Hilfe der vorgestellten Klassifikationsalgorithmen werden die Datensätze der Forstinventur im Bayerischen Staatswald gemäß der angegebenen Alters-Höhenbeziehung nach den Attributen aufgeteilt, die nach den in Abschnitt 3.2.1 vorgestellten Algorithmen jeweils den besten „Trenneffekt“ aufweisen. Ergebnis dieses Klassifikationsschrittes sind Gruppen, deren Elemente (Datensätze) ähnliche Häufigkeitsverteilungen in Bezug auf die Lage der klasseneinteilenden Regressionslinie bzw. zu anderen isolierten Gruppen aufweisen. Am Ende des Klassifikationsschrittes liegen Entscheidungsbäume vor, deren Endknoten (Blätter) die standörtlich isolierten Gruppen gemäß der Höhenentwicklung der Forstinventurbäume auf diesen Standorten repräsentieren. Für jede isolierte Gruppe bzw. für jedes Blatt des Entscheidungsbaumes dient die ermittelte Fehlklassifikationsrate als Indikator für die Beurteilung der Höhenentwicklung in den isolierten Gruppen.

Abbildung 17 zeigt exemplarisch einen Entscheidungsbaum für die Höhenmessbäume der Baumart Fichte im Wuchsbezirk 4.2 („Südliche Fränkische Platte“). Obwohl die Baumart Fichte in dieser Region aufgrund der relativen Trockenheit und Niederschlagsarmut keine optimalen Wachstumsbedingungen vorfindet, wurde der Datensatz ausgewählt, weil die relativ geringe Anzahl von 604 Datensätzen eine anschauliche Ergebnisdarstellung ermöglicht. Zur Generierung des Entscheidungsbaumes wurde der CHAID-Algorithmus (vgl. Abschnitt 3.2.1) angewendet. Für das Trennen als auch für das Verbinden von Datensätzen wurde ein p-Wert von 0,05 voreingestellt. Als Abbruchregel wurde eine Baumtiefe von maximal 5 Ebenen und eine Mindestanzahl von 50 Datensätzen pro Blatt festgelegt.

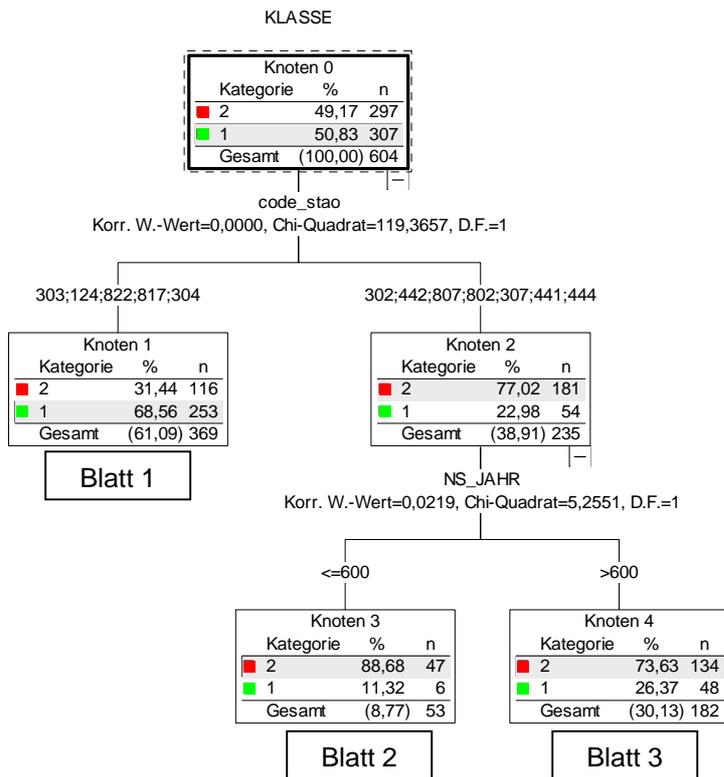


Abb. 17: Mit dem CHAID-Algorithmus generierter Entscheidungsbaum für die Baumart Fichte im Wuchsbezirk 4.2 („Südliche Fränkische Platte“).

Wuchsbezirk 4.2 (Südliche Fränkische Platte) - Fichte - CHAID

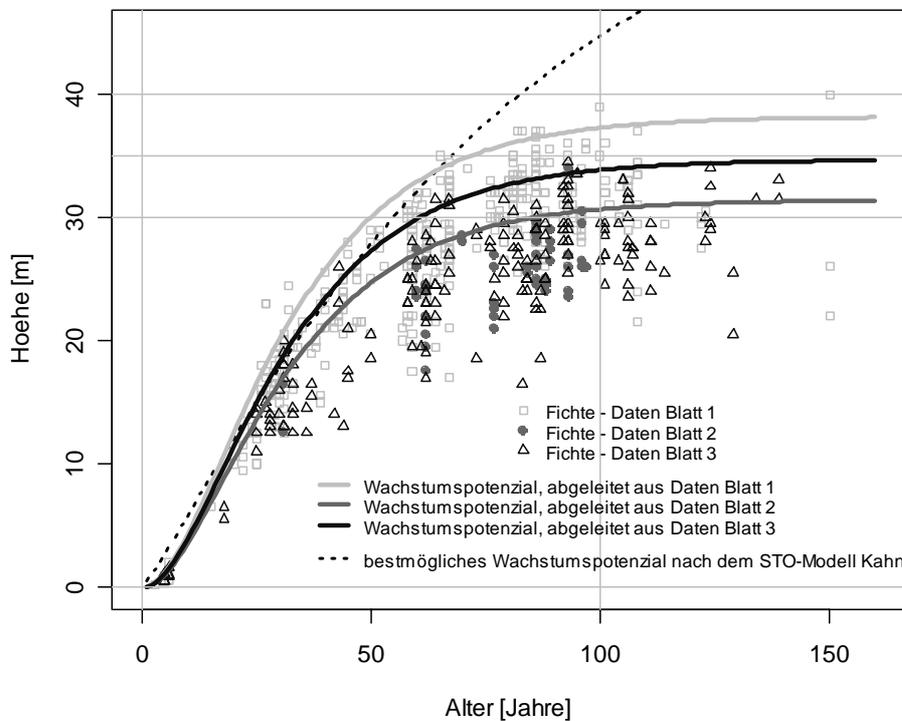


Abb. 18: Potenzielle Verläufe der Wachstumskurven für die Daten in den „Blättern“ gemäß Abb. 17. Zum Vergleich enthält Abb. 18 das bestmögliche Wachstumspotenzial des SILVA-Standort-Leistungsmodells nach Kahn (1994).

Gemäß Abbildung 17 konnten in der Zentralen Forsteinrichtungsdatenbank 604 Höhenmessbäume der Baumart Fichte für diesen Wuchsbezirk gefunden werden (vgl. Knoten 0). Mit Hilfe des CHAID-Algorithmus werden diese Daten nach der Standorteinheit (Attribut: code_stao) in zwei weitere Gruppen unterteilt. Die in Knoten 1 eingeordneten Datensätze befinden sich auf Standorten der Standorteinheiten 303, 124, 822, 817 und 304. In diesen Knoten wurden 369 Datensätze sortiert, wobei von diesen 253 Datensätze als Bäume mit überdurchschnittlichem Höhenwachstum (Klasse 1) und 116 Bäume als Bäume mit unterdurchschnittlichem Höhenwachstum (Klasse 2) gekennzeichnet waren. In Knoten 2 hingegen wurden 235 Fälle zugeordnet, von denen lediglich 54 Fälle ein überdurchschnittliches Höhenwachstum aufwiesen und 181 Fälle bezüglich des Höhenwachstums der Klasse 2 zugeordnet wurden. Demnach wiesen die Fichten auf den Standorteinheiten 302, 442, 807, 802, 307, 441 und 444 ein unterdurchschnittliches Höhenwachstum auf. Der hohe Chi-Quadrat-Wert von 119,3657 weist auf höchstsignifikant unterschiedliche Häufigkeitsverteilungen zwischen beiden Knoten hin.

Betrachtet man die zusammengefassten Standorteinheiten, so stellt man fest, dass in Knoten 1 Höhenmessbäume der Baumart Fichte auf Standorten zusammengefasst werden, die einen vergleichsweise guten Wasserhaushalt aufweisen während in Knoten 2 Baumdatensätze auf Standorten zusammengefasst sind, die einen schlechten oder unausgeglichene Wasserhaushalt aufweisen. Eine Ausnahme hierbei stellen die Fichten-Höhenmessbäume auf den Standorteinheiten 817 und 444 dar, die nach forstlichem Sachverstand zwischen den beiden Knoten ausgetauscht werden sollten. Grund für diese maschinell vorgenommene Zuordnung ist in beiden Fällen die mangelnde Anzahl an Fichten-Höhenmessbäumen auf den genannten Standorteinheiten (817: 2 Höhenmessbäume; 444:3 Höhenmessbäume; vgl. Abb. 83). Nach dieser grundlegenden Einteilung ist in Knoten 2 das voreingestellte Abbruchkriterium noch nicht erreicht. Mit Hilfe des CHAID-Algorithmus kann eine weitere Verbesserung des Klassifikationsergebnisses dadurch erreicht werden, dass die Fälle in Blatt 2 (schlechter bzw. unausgeglichener Wasserhaushalt) weiter nach der Niederschlagsmenge im Jahr (NS_Jahr) unterteilt werden. Demnach kommt es zu einer weiteren Verschlechterung des Verhältnisses von über- und unterdurchschnittlichen hohen Fichten, wenn die Niederschlagsmenge im Jahr kleiner oder gleich 600 mm ist (Knoten 3) wohingegen das schlechte Verhältnis von Knoten 2 etwas verbessert werden kann, wenn die Niederschlagsmenge im Jahr über 600 mm liegt (Knoten 4).

Zur Durchführung der Klassifikationsanalysen wurde für diese Arbeit zunächst die Software SPSS-AnswerTree 3.0 verwendet. Hiermit wurden die Algorithmen CHAID und CART angewendet. Mit der frei zugänglichen Data-Mining-Software WEKA wurden die Berechnungen mit dem C4.5-Algorithmus durchgeführt. Wie erste Berechnungsläufe mit WEKA gezeigt haben, führte die Anwendung des ID3-

Algorithmus zu unbefriedigenden Ergebnissen. Der Grund hierfür waren in der Regel zu komplexe Baumaufbauformen bzw. an die Trainingsdatensätze überangepaßte Entscheidungsbäume. Aus diesem Grund wurden nach den ersten Probeläufen keine weiteren Berechnungen mit diesem Algorithmus durchgeführt. In der Ergebnisdarstellung dieser Arbeit wird ebenfalls auf die Darstellung der ID3-Ergebnisse verzichtet.

Im Anschluss an den Klassifikationsschritt werden für alle Blätter des Entscheidungsbaumes bzw. für alle darin enthaltenen Baumdatensätze die Parameterwerte A , k und P der Wachstumsfunktion nach Chapman-Richards ermittelt. Die Auswahl fiel deshalb auf diese Wachstumsfunktion, da sie maßgebend für die Ermittlung des potenziellen Höhenzuwachses in SILVA ist (vgl. Abschnitt 2.2). Um ein in sich konsistentes Kurvensystem zu generieren, wird nach folgenden Schritten vorgegangen:

- 1) Ermittle für die gesamten verfügbaren Höhenmessbäume einer Baumart in einem Wuchsbezirk regressionsanalytisch die Werte für die Parameter A , k und P der Wachstumsfunktion nach Chapman-Richards gemäß der (Inventur)-Alters-Höhenbeziehung. Das Ergebnis dieses Rechenschrittes spiegelt den Wachstumsgang einer Baumart zum Auswertungszeitpunkt bei mittleren standörtlichen Verhältnissen innerhalb des Wuchsbezirkes wieder. Gleichzeitig stellt die ermittelte Kurve die Grenze für die Klasseneinteilung zur Zuordnung einer Klasse für jeden Baum dar (vgl. Abb. 16).
- 2) Hebe oder senke für jedes Blatt des Entscheidungsbaumes den nach Schritt 1 errechneten Wert des Parameters A solange, bis eine festgelegte Anzahl an Höhenmessbäumen dieses Blattes unterhalb der neuen Kurve zum Liegen kommt. Die so neu ermittelte Kurve stellt demnach eine Art obere Umhüllende um die Datenpunkte eines Blattes dar und wird als Ausdruck des maximal möglichen Wachstumsganges auf den zusammengefassten Standorten gesehen.

Die nichtlinearen Regressionsrechnungen wurden für diese Arbeit zuerst mit der Statistiksoftware SPSS 12.0, danach mit der frei verfügbaren Statistiksoftware R (R Development Core Team, 2006) durchgeführt.

Abbildung 18 zeigt für die Blätter des nach Abbildung 17 generierten Entscheidungsbaumes den Verlauf der errechneten und „angehobenen“ Chapman-Richards-Kurven. Die Kurven wurden dabei im Parameter A solange angehoben, bis 95% aller Datenpunkte unterhalb des neuen Kurvenverlaufes zu liegen kamen. Zum Vergleich ist in Abbildung 18 zusätzlich die Wachstumskurve mit den bestmöglichen

Standorteinstellungen nach dem Standort-Leistungsmodell von Kahn (1994) eingezeichnet.

Ein Ziel dieser Arbeit besteht darin eine Softwarelösung zu entwickeln, die sowohl die Entscheidungsbaumbaumgenerierung als auch die Durchführung der regressionsanalytischen Rechnungen übernimmt. Damit die Ergebnisse dieser Berechnungsschritte für das Wachstumsmodell SILVA zugänglich gemacht werden, müssen sie diesem in geeigneter Form übergeben werden. Hierzu soll die Softwarelösung die Ergebnisse der beiden Berechnungsschritte in Form von maschinellen Regeln darstellen, die in Datenbanktabellen oder Dateien geschrieben werden und dort von SILVA aufgenommen werden können. Die Regeln weisen hierbei folgendes Schema auf:

Beispielregel für Knoten 1 bzw. Blatt 1 des in Abbildung 17 dargestellten Entscheidungsbaumes:

```
WENN      code_stoe = 303 ODER
          code_stoe = 124 ODER
          code_stoe = 822 ODER
          code_stoe = 817 ODER
          code_stoe = 304
DANN  A = 38,18 UND k=-0,045 UND P=2,1566
```

Beispielregel für Knoten 3 bzw. Blatt 2 des in Abbildung 17 dargestellten Entscheidungsbaumes:

```
WENN code_stoe = 302 UND ns_jahr<=600 ODER
     code_stoe = 442 UND ns_jahr<=600 ODER
     code_stoe = 807 UND ns_jahr<=600 ODER
     code_stoe = 802 UND ns_jahr<=600 ODER
     code_stoe = 307 UND ns_jahr<=600 ODER
     code_stoe = 441 UND ns_jahr<=600 ODER
     code_stoe = 444 UND ns_jahr<=600
DANN  A = 31,18 UND k=-0,045 UND P=2,1566
```

Beispielregel für Knoten 4 bzw. Blatt 3 des in Abbildung 17 dargestellten Entscheidungsbaumes:

```
WENN code_stoe = 302 UND ns_jahr>600 ODER
     code_stoe = 442 UND ns_jahr>600 ODER
     code_stoe = 807 UND ns_jahr>600 ODER
     code_stoe = 802 UND ns_jahr>600 ODER
     code_stoe = 307 UND ns_jahr>600 ODER
     code_stoe = 441 UND ns_jahr>600 ODER
     code_stoe = 444 UND ns_jahr>600
DANN  A = 34,52 UND k=-0,045 UND P=2,1566
```

Zahlreiche Autoren weisen in der Fachliteratur auf die Bedeutung des Experten- oder Domänenwissen für die Interpretation und Evaluation der Ergebnisse von Data-Mining-Prozessen hin (vgl. z. B. Witten und Frank, 2005; Rudolph, 2006).

Die nach Abschnitt 3.2.3 dargestellten Ergebnisse des Knowledge-Discovery-in-Databases-Prozesses sollen daher forstlichen Experten so zur Verfügung gestellt werden, dass diese durch die graphische Aufbereitung frühzeitig falsche Klassifikationen oder falsche regressionsanalytisch hergeleitete Werte erkennen können. Zudem sollen die Experten, die letztendlich eine regionale, standortbezogene Feinjustierung mit Hilfe der erarbeiteten Techniken anhand neuer Daten durchführen, zu jeder Zeit in die Lage versetzt werden, diese Fehler nach eigener Einschätzung zu korrigieren. Dem Expertenwissen wird also auch in dieser Arbeit eine große Bedeutung zugemessen.

3.2.4 Online-Befragung zum Vergleich von Forsteinrichtungsdatenquellen in Deutschland

Ein Teilziel der vorliegenden Arbeit besteht darin eine Softwarelösung zu entwickeln, die eine regionale, standortbezogene Feinjustierung des Wachstumsmodells SILVA ermöglicht. Im ersten Entwicklungsschritt soll diese Software dabei auf die Inventurdaten aus Bayern angewendet werden. Die Softwarelösung soll allerdings so konzeptioniert und konstruiert werden, dass sie leicht auf Forstinventurdaten anderer Regionen oder Bundesländer in Deutschland angewendet werden kann. Aus diesem Grund ist es notwendig, eine aktuelle Übersicht über die Datenhaltung von Inventurdaten in Deutschland zu erstellen. Zu diesem Zweck wurde im August bzw. September 2005 eine Online-Befragung bei verschiedenen Forstbetrieben in Deutschland durchgeführt.

Eine Befragung ist eine wissenschaftliche Methode, um systematisch Informationen über Eigenschaften von Menschen oder Organisationen zu gewinnen. Befragungen als Basis für die Analyse- und Dokumentationsarbeit sind ein wichtiges Instrument der empirischen Sozialwissenschaften. Unterscheidungskriterium für unterschiedliche Formen der Befragung (im Rahmen der Erhebung (Empirie)) ist vor allem das Kommunikationsmedium, während der Grad der Standardisierung und Strukturiertheit sich neben der Abgrenzung zwischen qualitativer und quantitativer Forschung vor allem auf Art und Aufbau des Fragebogens bezieht. Nach der Art der Kommunikation mit den Befragten unterscheidet man persönliche, telefonische, schriftliche und Online-Befragungen. Jede Form hat ihre spezifischen Vor- und Nachteile (WIKIPEDIA 2006e).

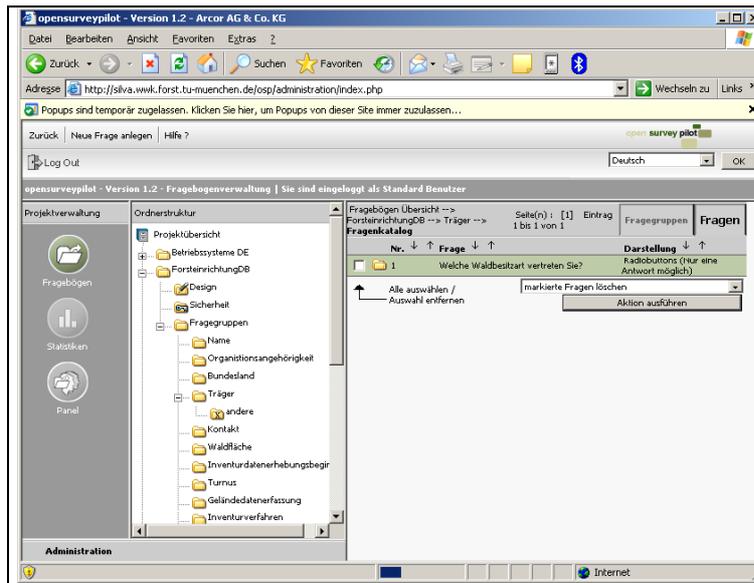


Abb. 19: Bildschirmabzug der browserbasierten Entwicklungsoberfläche der Software „opensurvey-pilot“.

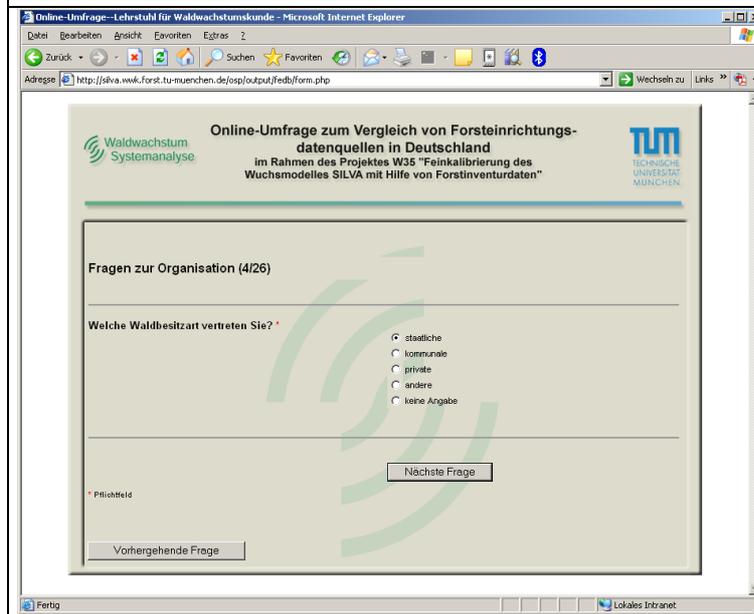


Abb. 20: Bildschirmabzug einer Frage (hier Frage 4, vgl. Tabelle 5), wie sie den Befragungsteilnehmern präsentiert wurde.

In vorliegender Arbeit wurde eine Online-Befragung durchgeführt, da diese Form besonders geeignet ist, wenn Mitglieder einer speziellen Grundgesamtheit befragt werden sollen, die mit der nötigen Technik ausgerüstet sind und diese auch nutzen (WIKIPEDIA 2006e). Um eine gewisse Repräsentativität der Umfrage zu gewährleisten, wurden im Vorfeld der Umfrage in allen Bundesländern der Bundesrepublik Deutschland Waldbesitzer der Besitzarten „Staatlicher Besitz“, „Kommunaler Waldbesitz“ und „Privater Waldbesitz“ angeschrieben, von denen ermittelt werden konnte, dass sie mindestens über eine Waldfläche von 1250 ha verfügen. Diese Grenze wurde für diese Arbeit gutachtlich auf Basis der durchschnittlichen Reviergrößen für den Staatswald in Bayern (Bayerische Staatsforstverwaltung, 2004) unter der Annahme festgelegt, dass diese Flächengröße den kostendeckenden Einsatz von eigenem Personal ermöglicht, das ausschließlich mit der Waldbewirtschaftung betraut ist. Da für vorliegende Arbeit

keine Liste der Waldbesitzer in Deutschland zur Verfügung stand, wurden im Rahmen einer Internet-Recherche die Adressen von Waldbesitzern bei den beiden großen Zertifizierungsorganisationen PEFC (<http://www.pefc.de>) und FSC (<http://www.fsc-deutschland.de>) ermittelt, die die Mindestwaldfläche von 1250 ha aufwiesen. Über die benannten Quellen konnten 242 Adressen ermittelt werden, von denen 60 das genannte Flächenkriterium erfüllten. Diese 60 potenziellen Umfrageteilnehmer wurden im Vorfeld der Umfrage angeschrieben und um Teilnahme gebeten.

Die Online-Befragung wurde mit Hilfe der als Open-Source erhältlichen PHP-basierten Software „opensurveypilot“ durchgeführt (<http://osp.homedns.org/>), die von van de Weyer (2004) kurz vorgestellt wird. Abbildung 19 zeigt einen Bildschirmabzug der Entwicklungsoberfläche dieser Softwarelösung. Abbildung 20 zeigt exemplarisch eine Frage, wie sie den Befragungsteilnehmern präsentiert wurde.

Tabelle 5 auf der folgenden Seite listet die gestellten Fragen der Online-Befragung auf.

Tab. 5: Fragen der Online-Befragung zur vergleichenden Darstellung der Datenhaltung von Forstinventurdaten in Deutschland.

Nr.	Frage
1	Bitte geben Sie Ihren Namen ein
2	Welche Organisation vertreten Sie?
3	In welchem Bundesland in der Bundesrepublik Deutschland ist Ihre Organisation ansässig?
4	Welche Waldbesitzart vertreten Sie? (staatlich, privat, kommunal)
5	Bitte geben Sie eine gültige eMail/ Telefonnummer Ihrer Organisation für mögliche Rückfragen an.
6	Wie groß ist die von Ihrer Organisation verwaltete Waldfläche? [ha]
7	Seit wann werden in der von Ihrer Organisation verwalteten Waldfläche Inventurdaten erhoben?
8	In welchem Turnus (Jahre) werden auf den gleichen Waldflächen in Ihrer Organisation Inventurdaten erhoben?
9	Welche Hilfsmittel verwenden Sie derzeit zur Datenaufnahme im Gelände? (Bsp.: Latschbacher TimbaTec)
10	Welches Inventurverfahren nutzen Sie?
11	Werden die Daten im Wald auf permanenten oder temporären Aufnahmeflächen aufgenommen?
12	Nutzt Ihre Organisation forstliche Standardsoftware oder individuell angepasste Softwarelösungen (Erläuterung: zur Verarbeitung der Inventurdaten)
13	Seit wann werden in Ihrer Organisation mit einem Datenbankmanagementsystem (Inventur-)Daten verwaltet?
14	Wie viele Datensätze (Baumdaten) wurden zum Stichtag 01.08.2005 von Ihrem Datenbankmanagementsystem verwaltet?
15	Wie erfolgt die Datenhaltung? (zentral/ dezentral)
16	Bitte benennen Sie die Architektur Ihres Datenbankmanagementsystems (Fernverarbeitung/ Client-Server-Lösung/ File-Server-Lösung etc.)
17	Nach welchem Datenmodell halten Sie ihre Forsteinrichtungsdaten? (hierarchisch/ relational/ objektorientiert etc.)
18	Bitte geben Sie den Hersteller Ihres Datenbankmanagementsystems an.
19	Wie viele Benutzer haben insgesamt Zugriff auf die Forsteinrichtungsdatenbanken?
20	Wie viele Benutzer haben im Durchschnitt gleichzeitig Zugriff auf die Forsteinrichtungsdaten?
21	Wie viele Personen sind mit der Administration Ihres Datenbankmanagementsystems beschäftigt? (Vollzeit)
22	Welche Techniken nutzen Sie zur Auswertung der Daten? (Klassische Abfragesprachen/ OLAP/ Data-Mining/ andere)
23	Welche Datenbanksprache benutzen Sie? (SQL/ QBE/ andere)
24	Wer nutzt in Ihrer Organisation die Forsteinrichtungsdaten am häufigsten? (Betriebsleitung/ Forsteinrichtung/ GIS/ andere)
25	Werden die erhobenen Daten standardmäßig mit einem GIS-System weiterverarbeitet?
26	Anmerkungen (freies Textfeld, max. 255 Zeichen).

4. Ergebnisse

4.1 Regressionsanalysen zur Einteilung der Forstinventurdaten einer Region

Nachfolgend werden nach Wuchsgebieten und Baumarten getrennt die Ergebnisse der nichtlinearen Regressionsanalysen dargestellt. Aufgrund der flächenmäßigen Bedeutung des Staatswaldes in den Teilwuchsbezirken der Wuchsbezirke 2.2, 12.9 und 14.4 in Verbindung mit den hohen Zahlen an Höhenmessbäumen in allen Teilwuchsbezirken werden für diese Teilwuchsbezirke zusätzlich zu den Wuchsbezirksangaben die Koeffizienten für die Parameter A, k und P der Wachstumsfunktion nach Chapman-Richards angegeben.

Die ermittelten Koeffizienten für die Wachstumskurven stellen, wie in Abschnitt 3.1.4 (Abbildung 16) beschrieben, die Grundlage für die Klasseneinteilung dar.

Tab. 6: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume der benannten Baumarten des Wuchsgebietes 2 „Spessart“ mit seinen Wuchsbezirken. Für den Wuchsbezirk 2.2 („Buntsandsteinspessart“) erfolgt zusätzlich die Darstellung der Werte für die Teilwuchsbezirke 2.2.1 („Hochspessart“), 2.2.2 („Nordspessart“) und 2.2.3 („Mainspessart“).

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
2	1		Buche	30,027	-0,0298	1,541	0,723
2	1		Kiefer	29,086	-0,0396	1,785	0,771
2	1		Lärche	36,606	-0,0179	0,907	0,780
2	2		Buche	32,981	-0,0225	1,221	0,665
2	2	1	Buche	33,061	-0,0225	1,234	0,660
2	2	2	Buche	29,855	-0,0320	1,548	0,688
2	2	3	Buche	35,536	-0,0226	1,195	0,798
2	2		Douglasie	41,681	-0,0242	1,286	0,773
2	2	1	Douglasie	39,973	-0,0264	1,325	0,761
2	2		Eiche	31,712	-0,0203	1,149	0,876
2	2	1	Eiche	31,816	-0,0202	1,146	0,877
2	2	2	Eiche	28,477	-0,0467	2,956	0,893
2	2		Fichte	33,867	-0,0374	1,699	0,779
2	2	1	Fichte	33,850	-0,0383	1,747	0,789
2	2	2	Fichte	35,156	-0,0271	1,243	0,769
2	2		Kiefer	29,699	-0,0423	1,916	0,811
2	2	1	Kiefer	29,320	-0,0413	1,726	0,741
2	2	2	Kiefer	28,637	-0,0340	1,305	0,799
2	2		Lärche	34,238	-0,0324	1,293	0,742
2	2	1	Lärche	34,443	-0,0317	1,263	0,738
2	2	2	Lärche	33,397	-0,0334	1,399	0,773

Aus Tabelle 6 wird ersichtlich, dass für die einzelnen Wuchsbezirke bzw. Teilwuchsbezirke nicht mehr für alle Baumarten die im Vorfeld gesetzte Zahl von mindestens 100 Höhenmessbäumen gefunden wurde und daher nicht mehr für alle Baumarten die Werte der Parameter A, k und P ermittelt werden konnten. Die korrigierten Bestimmtheitsmaße liegen über alle Baumarten des Wuchsgebietes zwischen 0,66 und 0,876 und deuten bereits auf eine sehr gute und sehr stabile Anpassung der Kurvenverläufe an die gegebenen Datenpunkte hin.

Tabelle 7 enthält die Werte für das Wuchsgebiet 3 „Rhön“.

Tab. 7: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume der benannten Baumarten des Wuchsgebietes 3 „Rhön“ mit seinen Wuchsbezirken.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
3	1		Buche	28,263	-0,0261	1,369	0,694
3	1		Douglasie	37,855	-0,0246	1,279	0,794
3	1		Eiche	26,361	-0,0235	1,228	0,756
3	1		Fichte	29,085	-0,0356	1,609	0,682
3	1		Kiefer	28,022	-0,0316	1,359	0,768
3	1		Lärche	31,271	-0,0296	1,254	0,773
3	1		Tanne	28,908	-0,0387	1,913	0,691
3	2		Buche	30,127	-0,0295	1,796	0,689
3	2		Fichte	32,254	-0,0215	1,208	0,750
3	2		Lärche	35,599	-0,0225	1,466	0,717

Für das Wuchsgebiet 3 zeigt Tabelle 7, dass für den Wuchsbezirk 3.1 („Vorrhön“) für alle in die Auswertung einbezogenen Baumarten die Werte für die Parameter A, k und P ermittelt werden konnten, wohingegen für den Wuchsbezirk 3.2 („Hohe Rhön“) lediglich für die Baumarten Buche, Fichte und Lärche die entsprechenden Werte berechnet werden konnten. Die korrigierten Bestimmtheitsmaße liegen über alle Baumarten des Wuchsgebietes zwischen 0,689 und 0,794. Damit liegt die obere Grenze im Vergleich zu den Werten aus dem Wuchsgebiet 2 („Spessart“) etwas niedriger.

In Tabelle 8 werden für das Wuchsgebiet 4 „Fränkische Platte“ die Werte der drei Parameter der Chapman-Richards-Funktion dargestellt.

Tab. 8: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume der benannten Baumarten des Wuchsgebietes 4 „Fränkische Platte“ mit seinen Wuchsbezirken.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
4	1		Buche	24,495	-0,0507	2,801	0,727
4	1		Eiche	23,931	-0,0331	1,620	0,837
4	1		Fichte	27,993	-0,0680	3,765	0,833
4	1		Kiefer	24,015	-0,0465	1,828	0,658
4	1		Lärche	28,213	-0,0406	1,653	0,660
4	2		Buche	34,793	-0,0211	1,248	0,750
4	2		Douglasie	38,718	-0,0289	1,587	0,828
4	2		Eiche	28,795	-0,0292	1,503	0,832
4	2		Fichte	30,683	-0,0450	2,156	0,784
4	2		Kiefer	29,619	-0,0283	1,228	0,780
4	2		Lärche	34,175	-0,0254	1,106	0,629

An den Parameterwerten in Tabelle 8 fällt auf, dass die Werte des Parameters A für alle Baumarten des Wuchsbezirkes 4.1 („Nördliche Fränkische Platte“) deutlich unter den vergleichbaren Werten des Wuchsbezirkes 4.2 („Südliche Fränkische Platte“) liegen. Diese Aussage gilt auch für den entsprechenden Vergleich mit den anderen Wuchsgebieten Unterfrankens. Nachdem der Parameter A für die Lage der Asymptote verantwortlich ist, an die sich die ermittelte Wachstumskurve annähert, könnten die niedrigen A-Werte im Wuchsbezirk 4.1 auf folgende zwei Phänomene zurückzuführen sein: Das Wuchsgebiet 4 „Fränkische Platte“ gilt seit langem als waldarmes Bauernland. Nirgendwo anders in Unterfranken ist Ackerbau so ertragreich wie in den Gäulandschaften um Ochsenfurt, zwischen Würzburg und Schweinfurt oder im Grabfeld. Seit einem Jahrhundert wird hier intensive Landwirtschaft betrieben. Gezielt rodeten die Bauern den Wald auf allen Böden, die gute Ernte versprachen. Dem Wald blieben in diesen Altsiedlungslandschaften fast nur die kargen Muschelkalkhochflächen wie um Münnerstadt oder zwischen Bad Kissingen und Haßfurt. (Bayerisches Staatsministerium für Ernährung, Landwirtschaft und Forsten, 1986). Weiterhin gilt die Fränkische Platte als eine der niederschlagsärmsten Regionen Bayerns (BayFORKLIM, 1996), was zusätzlich einen entscheidenden Einfluss auf das Höhenwachstum der Bäume ausübt.

Tabelle 9 enthält die berechneten Werte der Parameter A, k und P für das Wuchsgebiet 5 („Fränkischer Keuper und Albvorland“).

Tab. 9: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume der benannten Baumarten des Wuchsgebietes 5 „Fränkischer Keuper und Albvorland“ mit seinen Wuchsbezirken.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
5	1		Buche	32,939	-0,0169	0,894	0,573
5	1		Eiche	28,927	-0,0174	0,730	0,566
5	1		Fichte	34,243	-0,0270	1,141	0,669
5	1		Kiefer	30,919	-0,0261	1,020	0,692
5	1		Lärche	32,078	-0,0330	1,365	0,749
5	2		Buche	32,662	-0,0182	1,004	0,585
5	2		Eiche	28,281	-0,0234	0,986	0,479
5	2		Fichte	32,673	-0,0198	0,838	0,725
5	2		Kiefer	29,611	-0,0250	0,917	0,736
5	2		Lärche	36,823	-0,0168	0,767	0,744
5	3		Fichte	28,941	-0,0277	1,365	0,714
5	3		Kiefer	29,976	-0,0323	1,679	0,783
5	4		Buche	29,374	-0,0340	1,793	0,768
5	4		Eiche	27,031	-0,0373	1,794	0,808
5	4		Fichte	30,133	-0,0444	2,143	0,792
5	4		Kiefer	30,086	-0,0275	1,082	0,765
5	5		Eiche	28,174	-0,0369	2,753	0,923
5	5		Fichte	34,372	-0,0250	1,804	0,836
5	5		Kiefer	28,543	-0,0466	3,380	0,669
5	6		Buche	29,068	-0,0355	2,297	0,862
5	6		Eiche	26,795	-0,0352	2,292	0,903
5	6		Fichte	30,050	-0,0323	1,953	0,775
5	6		Kiefer	25,583	-0,0353	1,722	0,594
5	6		Lärche	30,111	-0,0260	1,103	0,654
5	7		Buche	31,737	-0,0331	1,899	0,818
5	7		Eiche	30,088	-0,0324	1,623	0,869
5	7		Fichte	32,125	-0,0531	3,623	0,926
5	7		Kiefer	32,969	-0,0152	0,618	0,667
5	7		Lärche	33,423	-0,0176	0,681	0,679
5	8		Buche	30,539	-0,0261	1,725	0,755
5	8		Eiche	30,341	-0,0267	1,630	0,847
5	8		Fichte	33,680	-0,0224	1,200	0,730
5	8		Kiefer	30,109	-0,0263	1,447	0,718
5	8		Lärche	33,140	-0,0225	1,056	0,825
5	8		Tanne	31,301	-0,0368	2,492	0,661

Betrachtet man die korrigierten Bestimmtheitsmaße so zeigt sich, dass die Werte zwischen 0,479 und 0,923 über alle Baumarten und Wuchsbezirke variieren. Auffällig sind die unterdurchschnittlichen Werte für alle Baumarten in den Wuchsbezirken 5.1 und 5.2. Bei einer genaueren Betrachtung der zugrundeliegenden Daten fällt hierbei auf, dass in diesen Wuchsbezirken nur vergleichsweise alte Forsteinrichtungen als Datengrundlage zur Verfügung standen. Bei diesen wurden nach den vorliegenden Daten in den Wuchsbezirken 5.1 und 5.2 lediglich Bäume als Bäume der Oberschicht gekennzeichnet, die sich im äußersten Messkreis der konzentrisch angeordneten Probekreise befanden. Eine Überprüfung dieser Annahme konnte allerdings für diese Altinventuren nicht vorgenommen werden, da keine Inventuranweisungen vorlagen.

Tabelle 10 zeigt die ermittelten Werte für das Wuchsgebiet 6 „Frankenalb und Oberpfälzer Jura“. Aufgrund der hohen Zahl an Höhenmessbäumen in diesem Wuchsgebiet konnten für nahezu alle in die Auswertung dieser Arbeit einbezogenen Baumarten die entsprechenden Werte berechnet werden. Auffällig ist, dass die Werte des Parameters A für die Baumart Douglasie im Vergleich zu den anderen Baumarten, wie in den übrigen Wuchsgebieten Bayerns auch, am höchsten liegen.

Tab. 10: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume der benannten Baumarten des Wuchsgebietes 6 „Frankenalb und Oberpfälzer Jura“ mit seinen Wuchsbezirken.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
6	1		Buche	27,894	-0,0387	2,186	0,654
6	1		Douglasie	43,063	-0,0224	1,509	0,856
6	1		Eiche	24,932	-0,0483	2,405	0,862
6	1		Fichte	30,076	-0,0406	2,132	0,686
6	1		Kiefer	26,961	-0,0424	2,139	0,448
6	1		Lärche	32,615	-0,0298	1,510	0,690
6	1		Tanne	29,653	-0,0438	2,939	0,754
6	2		Buche	30,194	-0,0326	1,969	0,808
6	2		Douglasie	42,651	-0,0245	1,536	0,846
6	2		Eiche	30,506	-0,0316	1,788	0,885
6	2		Fichte	32,265	-0,0372	1,996	0,801
6	2		Kiefer	27,870	-0,0404	2,114	0,702
6	2		Lärche	33,705	-0,0297	1,594	0,808
6	2		Tanne	30,630	-0,0338	2,100	0,792
6	3		Buche	30,057	-0,0309	2,173	0,755
6	3		Eiche	28,209	-0,0382	2,344	0,897
6	3		Fichte	33,975	-0,0400	2,661	0,867
6	3		Kiefer	30,808	-0,0366	2,233	0,857

Tab. 10 (Fortsetzung)

WUG	WBZ	TWBZ	Baumart	A	k	P	R ²
6	3		Lärche	34,352	-0,0378	2,540	0,831
6	4		Buche	24,405	-0,0695	3,894	0,827
6	4		Eiche	24,477	-0,0551	3,117	0,916
6	4		Fichte	27,757	-0,0584	3,071	0,798
6	4		Kiefer	25,824	-0,0446	2,064	0,691
6	4		Lärche	26,414	-0,0546	2,196	0,643
6	5		Buche	26,783	-0,0426	2,450	0,817
6	5		Eiche	26,799	-0,0330	1,730	0,861
6	5		Fichte	28,733	-0,0396	1,890	0,640
6	5		Kiefer	27,750	-0,0372	1,774	0,684
6	5		Lärche	30,583	-0,0328	1,505	0,701

Tabelle 11 enthält die Werte für das Wuchsgebiet 8 „Frankenwald, Fichtelgebirge und Steinwald“.

Tab. 11: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R² für die Höhenmessbäume der benannten Baumarten des Wuchsgebietes 8 „Frankenwald, Fichtelgebirge und Steinwald“.

WUG	WBZ	TWBZ	Baumart	A	k	P	R ²
8	1		Buche	27,336	-0,0345	1,933	0,773
8	1		Douglasie	37,725	-0,0223	1,215	0,743
8	1		Fichte	33,495	-0,0262	1,332	0,782
8	1		Kiefer	25,379	-0,0328	1,307	0,524
8	1		Tanne	35,433	-0,0185	1,158	0,800

Auffällig bei der Betrachtung von Tabelle 11 ist, dass diese Tabelle lediglich Werte für den Wuchsbezirk 8.1 „Frankenwald“ enthält. Die Gründe liegen in der restriktiven Vorauswahl des Datenmaterials. Sie wurden in Abschnitt 3.1.4 dieser Arbeit erläutert. Betrachtet man die Zahlenwerte genauer, so fällt das unterdurchschnittliche korrigierte Bestimmtheitsmaß für die Baumart Kiefer auf. Die Gründe hierfür sind aus dem Datenmaterial nicht direkt ersichtlich. Lässt man die Baumart Kiefer außer Acht, so liegen die korrigierten Bestimmtheitsmaße zwischen 0,743 und 0,8.

Tabelle 12 zeigt die errechneten Werte für das Wuchsgebiet 9 „Oberpfälzer Becken- und Hügelland“.

Tab. 12: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume der Baumart Kiefer des Wuchsgebietes 9 „Oberpfälzer Becken- und Hügelland“.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
9	1		Kiefer	27,106	-0,0396	2,979	0,521

Für den Wuchsbezirk 9.1 konnten, aufgrund der restriktiven Vorauswahl der Daten, nur wenige Daten genutzt werden, die zudem nur dem Randbereich des Wuchsgebietes entstammten. Die Baumart Kiefer war die einzige Baumart, bei der die gesetzte Mindestzahl an Höhenmessbäumen erreicht wurde. Die für dieses Datenkollektiv ermittelten Werte der Parameter der verwendeten Wachstumsfunktion sind in Tabelle 12 angegeben. Aufgrund dieser Einschränkung in Verbindung mit den niedrigen korrigierten Bestimmtheitsmaßen der Parameteranpassung sollte dieses Berechnungsergebnis vor einer Verwendung in SILVA bzw. vor einer breiten Anwendung noch einmal anhand eines größeren Datensatzes aus dieser Wuchsregion überprüft werden.

Tabelle 13 enthält die Werte für das Wuchsgebiet 11 („Bayerischer Wald“).

Tab. 13: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume der Baumarten des Wuchsgebietes 11 „Bayerischer Wald“.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
11	1		Kiefer	31,540	-0,0196	1,408	0,787
11	2		Buche	30,909	-0,0285	1,536	0,804
11	2		Eiche	26,655	-0,0433	2,192	0,821
11	2		Fichte	35,039	-0,0319	1,826	0,795
11	2		Kiefer	27,485	-0,0247	0,949	0,724
11	2		Tanne	31,913	-0,0296	1,532	0,790
11	3		Buche	26,908	-0,0247	1,614	0,542
11	3		Fichte	30,060	-0,0354	2,309	0,618
11	3		Tanne	30,910	-0,0257	1,837	0,641

Auffällig an den Werten in Tabelle 13 sind die unterdurchschnittlichen korrigierten Bestimmtheitsmaße für die Baumarten Buche, Fichte und Tanne im Wuchsbezirk 11.3 „Innerer Bayerischer Wald“. Die vergleichsweise schlechte Anpassung der Ausgleichskurven an das Datenmaterial könnte eventuell bedingt sein durch eine erhöhte Variation der Höhen der Höhenmessbäume über dem angegebenen Inventuralter. Diese größere Variation könnte bedingt sein z. B. durch erhöhte Schneebruchanteile in den Hochlagen des Bayerischen Waldes. Die Daten des

Wuchsbezirkes 11.3 enthalten neben Baumdaten aus dem Wirtschaftswald auch Daten aus dem Nationalpark Bayerischer Wald.

Tabelle 14 führt die Parameterwerte für das Wuchsgebiet 12 („Tertiäres Hügelland“) auf.

Tab. 14: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume der Baumarten des Wuchsgebietes 12 „Tertiäres Hügelland“. Zudem sind für die Teilwuchsbezirke 12.9/1 („Westliches Niederbayerisches Tertiärhügelland“) und 12.9/2 („Östliches Niederbayerisches Tertiärhügelland“) die Werte getrennt dargestellt.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
12	1		Eiche	27,079	-0,0294	1,548	0,906
12	1		Fichte	29,229	-0,0392	2,153	0,715
12	1		Kiefer	26,707	-0,0573	4,621	0,684
12	7		Buche	36,091	-0,0255	1,795	0,858
12	7		Douglasie	60,226	-0,0128	1,123	0,843
12	7		Eiche	30,802	-0,0340	2,043	0,921
12	7		Fichte	36,917	-0,0354	2,224	0,908
12	7		Kiefer	33,248	-0,0362	2,070	0,794
12	7		Lärche	38,964	-0,0293	1,524	0,803
12	7		Tanne	35,488	-0,0333	2,268	0,867
12	8		Buche	34,209	-0,0253	1,667	0,748
12	8		Douglasie	45,677	-0,0314	2,211	0,858
12	8		Eiche	30,954	-0,0277	1,597	0,866
12	8		Fichte	36,283	-0,0386	2,612	0,819
12	8		Kiefer	31,300	-0,0478	3,454	0,764
12	8		Lärche	34,882	-0,0411	2,294	0,745
12	8		Tanne	31,859	-0,0497	3,498	0,860
12	9		Buche	40,838	-0,0135	1,018	0,847
12	9	2	Buche	39,234	-0,0335	1,795	0,808
12	9		Douglasie	43,845	-0,0171	0,884	0,841
12	9		Eiche	33,491	-0,0171	0,981	0,805
12	9		Fichte	39,717	-0,0165	1,054	0,815
12	9	1	Fichte	36,238	-0,0236	1,180	0,617
12	9	2	Fichte	30,510	-0,0329	1,593	0,723
12	9		Kiefer	26,520	-0,0327	1,284	0,479
12	9	1	Kiefer	27,135	-0,0560	2,702	0,428
12	9	2	Kiefer	24,899	-0,0332	1,276	0,526
12	9		Lärche	36,512	-0,0184	0,973	0,686
12	9		Tanne	37,213	-0,0172	1,180	0,840

Auffällig an den Werten in Tabelle 14 sind relativ hohe korrigierte Bestimmtheitsmaße über alle Baumarten in allen Wuchsbezirken. Eine Ausnahme stellen hier die Kiefern (R^2 : 0,479) im Wuchsbezirk 12.9 bzw. in den Teilwuchsbezirken 12.9/1 und 2 dar (R^2 : 0,428 bzw. 0,526) dar. Die Gründe für dieses abweichende Verhalten liegen vermutlich darin begründet, dass bei allen verwendeten Forsteinrichtungen alle höhengemessenen Kiefern mit der Schichtenkennzeichnung „0“ (ungeschichtet) gekennzeichnet waren. Hierdurch ergibt sich eine größere Varianz der Streuung der Daten und dementsprechend eine schlechtere Anpassungsmöglichkeit mit der verwendeten Wachstumsfunktion.

In Tabelle 15 sind die Werte für das Wuchsgebiet 13 („Schwäbisch-Bayerische-Schotterplatten und Altmoränenlandschaft“) eingetragen.

Tab. 15: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume der Baumarten des Wuchsgebietes 13 „Schwäbisch Bayerische Schotterplatten und Altmoränenlandschaft“.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
13	2		Buche	30,005	-0,0367	2,365	0,936
13	2		Douglasie	49,672	-0,0184	1,306	0,834
13	2		Eiche	24,683	-0,0481	2,594	0,922
13	2		Fichte	36,064	-0,0356	2,329	0,937
13	2		Kiefer	30,472	-0,0393	2,361	0,884
13	2		Lärche	35,754	-0,0373	2,164	0,904
13	2		Tanne	32,317	-0,0319	1,905	0,870
13	3		Fichte	33,282	-0,0451	3,054	0,905
13	3		Kiefer	30,438	-0,0386	2,179	0,867
13	4		Buche	34,107	-0,0338	2,201	0,897
13	4		Eiche	38,652	-0,0211	1,547	0,944
13	4		Fichte	36,891	-0,0327	2,011	0,930
13	4		Tanne	35,089	-0,0328	2,059	0,918
13	5		Fichte	41,701	-0,0267	1,696	0,929
13	6		Buche	30,651	-0,0374	2,490	0,949
13	6		Eiche	28,275	-0,0400	2,244	0,951
13	6		Fichte	34,970	-0,0372	2,424	0,949
13	6		Kiefer	28,106	-0,0466	2,317	0,959
13	6		Tanne	32,315	-0,0346	2,380	0,956

Wie aus dieser Tabelle ersichtlich wird, konnten über alle Baumarten und Wuchsbezirke sehr hohe korrigierte Bestimmtheitsmaße ermittelt werden. Der kleinste errechnete Wert lag in dieser Region bei 0,834, der größte ermittelte Wert

lag bei 0,959. Dies deutet auf eine relativ geringe Streuung des Datenmaterials um den errechneten Verlauf der Wachstumskurve hin.

In Tabelle 16 sind die Werte für das Wuchsgebiet 14 eingetragen. Aufgrund der Größe des Wuchsbezirkes 14.4 und der räumlichen Verteilung des Staatswaldes auf die Teilwuchsbezirke 14.4/1 („Westliche kalkalpine Jungmoräne“) und 14.4/2 („Inn-Jungmoräne“) werden die Werte zusätzlich getrennt für diese beiden Teilwuchsbezirke dargestellt.

Tab. 16: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume des Wuchsgebietes 14 „Schwäbisch Bayerische Jungmoräne und Molassevorberge“ mit seinen Wuchs- bzw. Teilwuchsbezirken.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
14	3		Buche	31,128	-0,0334	2,337	0,890
14	3		Fichte	33,005	-0,0335	2,202	0,838
14	3		Kiefer	27,256	-0,0426	3,124	0,663
14	3		Tanne	32,997	-0,0307	2,334	0,908
14	4		Buche	30,713	-0,0378	2,548	0,908
14	4	1	Buche	30,219	-0,0400	2,681	0,896
14	4	2	Buche	31,666	-0,0342	2,335	0,933
14	4		Eiche	26,834	-0,0496	2,913	0,922
14	4	1	Eiche	24,711	-0,0654	4,101	0,908
14	4	2	Eiche	27,787	-0,0420	2,434	0,938
14	4		Fichte	32,034	-0,0450	3,125	0,832
14	4	1	Fichte	29,698	-0,0496	3,383	0,821
14	4	2	Fichte	35,098	-0,0389	2,700	0,880
14	4		Kiefer	20,034	-0,0620	4,963	0,310
14	4	1	Kiefer	18,902	-0,0163	0,995	0,299
14	4	2	Kiefer	28,614	-0,0467	3,380	0,690
14	4		Lärche	34,493	-0,0394	2,331	0,886
14	4	1	Lärche	32,057	-0,0459	2,776	0,890
14	4	2	Lärche	39,176	-0,0259	1,473	0,911
14	4		Tanne	31,815	-0,0450	3,374	0,952
14	4	1	Tanne	31,896	-0,0409	3,168	0,954
14	4	2	Tanne	31,955	-0,0458	3,325	0,945

Über alle Baumarten und Wuchsbezirke konnten wiederum sehr gute Anpassungen, ausgedrückt durch hohe korrigierte Bestimmtheitsmaße, ermittelt werden. Eine Ausnahme stellt in diesem Zusammenhang die Baumart Kiefer im Wuchsbezirk 14.4 dar, für die lediglich ein Bestimmtheitsmaß von 0,31 errechnet werden konnte. Betrachtet man für diese Baumart die Werte für die einzelnen Teilwuchsbezirke so

zeigt sich, dass im Teilwuchsbezirk 14.4/1 lediglich ein korrigiertes Bestimmtheitsmaß von 0,299 ermittelt werden konnte. Der Grund hierfür liegt ebenso wie für die Baumart Kiefer im Wuchsbezirk 12.9 darin begründet, dass keine weitergehende Schichtenzuordnung für die Höhenmessbäume bei den verwendeten Forsteinrichtungen vorgenommen wurde. Der R^2 -Wert für die Kiefern im Teilwuchsbezirk 14.4/2 liegt hingegen mit 0,69 wieder etwas höher.

Abschließend sind in Tabelle 17 die Werte der Parameter der der Chapman-Richards-Wachstumsfunktion für das Wuchsgebiet 15 „Bayerische Alpen“ eingetragen, wobei lediglich, wie in Abschnitt 3.1.4 erläutert, Werte für den Wuchsbezirk 15.1 „Kürnacher Molassebergland“ enthalten sind.

Tab. 17: Werte der Parameter A, k und P sowie korrigiertes Bestimmtheitsmaß R^2 für die Höhenmessbäume des Wuchsgebietes 15 „Bayerische Alpen“.

WUG	WBZ	TWBZ	Baumart	A	k	P	R^2
15	1		Buche	31,635	-0,0286	2,348	0,866
15	1		Fichte	40,404	-0,0256	1,961	0,896

An den Werten in Tabelle 17 fallen wiederum die vergleichsweise hohen Bestimmtheitsmaße für die Baumarten Buche und Fichte auf. Zudem deutet der A-Wert von ca. 40,4 für die Baumart Fichte auf ein hohes Höhenwachstumspotenzial bei mittleren standörtlichen Verhältnissen in dieser Region Bayerns hin.

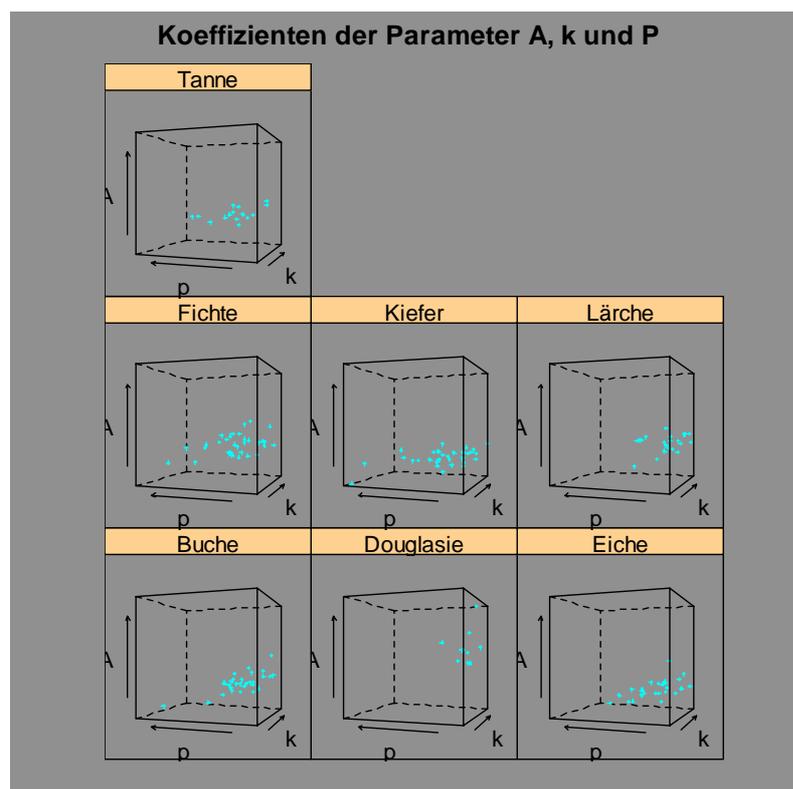


Abb. 21: Werte der Parameter A, k und P für die aus den Inventurdaten für mittlere Standortverhältnisse einer Wuchsregion abgeleiteten Wachstumskurven nach Chapman-Richards. Die Werte wurden den Tabellen 6 bis 17 entnommen. Der dargestellte Wertebereich für die Koeffizienten des Parameters A variiert dabei zwischen 18 und 68 (von unten nach oben), für k zwischen $-0,06$ und $-0,013$ (von vorne nach hinten) sowie für P zwischen 0,61 und 4,1 (von rechts nach links).

Abschließend sind in Abbildung 21 die in den Tabellen 6 bis 17 aufgeführten Werte der Parameter A, k und P nach Baumarten getrennt dargestellt. Auffällig sind die vergleichsweise hohen Werte des Parameters A (Asymptote) bei der Baumart Douglasie sowie die vergleichsweise geringen Werte für den Parameter P bei der Baumart Lärche. Beide Aussagen müssen allerdings unter dem Hintergrund gesehen werden, dass es sich um Parameter von Wachstumskurven handelt, die den Wachstumsgang einer Baumart bei mittleren standörtlichen Verhältnissen in einer Wuchsregion charakterisieren.

4.2 Beispielhafte Darstellung der Entscheidungsbäume für Wuchsbezirk 6.2

Nachfolgend werden die mit Hilfe des CHAID-, CART- und C4.5 (resp. J4.8)-Algorithmus ermittelten Entscheidungsbäume dargestellt. Die Ergebnisdarstellung erfolgt dabei aufgrund des erheblichen Umfangs nicht für alle Wuchsbezirke sondern exemplarisch für den Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Diese Region wurde wegen der flächenmäßigen Bedeutung des Staatswaldes (vgl. hierzu Anhang 9.1 bzw. Abb. 9) ausgewählt. Für die übrigen Wuchsbezirke können die berechneten Entscheidungsbäume mit Hilfe der Softwarelösung Silva-PEM (vgl. Abschnitt 5), die Bestandteil der schriftlichen Version dieser Arbeit ist, betrachtet werden.

Für den Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“ konnten mit Hilfe der Algorithmen CHAID, CART und C4.5 für die Baumarten Buche, Douglasie, Eiche, Fichte, Kiefer, Lärche und Tanne Entscheidungsbäume generiert werden. Eine Ausnahme stellt die Baumart Douglasie dar, für die die Anwendung des CHAID-Algorithmus keine Generierung eines Entscheidungsbaumes ermöglicht hat. Auf den folgenden Seiten werden die generierten Entscheidungsbäume dargestellt. Die Darstellung erfolgt dabei, für die mit Hilfe des CHAID und CART-Algorithmus ermittelten Entscheidungsbäume, anhand der SPSS-AnswerTree-Grafiken. Die, mit dem C4.5-Algorithmus ermittelten, Entscheidungsbäume werden nach der alphanumerischen Ausgabe von WEKA dargestellt (Witten und Frank, 2005). Da die generierten Entscheidungsbäume z. T. eine erhebliche Größe aufweisen, ist nur eine sehr kleine Darstellung in der schriftlichen Version dieser Arbeit möglich. Aus diesem Grund befindet sich vorliegende Arbeit im PDF-Format auf beiliegendem Datenträger. Die darin eingebetteten EMF-Grafiken für die Abbildungen der Abschnitte 4.2.1 bis 4.2.7 können verlustfrei skaliert und vergrößert werden.

4.2.1 Entscheidungsbäume für die Baumart Buche

Mit Hilfe des CHAID-Algorithmus (Abb. 22) erfolgt eine Zusammenfassung der Buchen-Höhenmessbäume auf der obersten Ebene nach Standorteinheiten (code_stao). Als Standorte mit besonders ungünstigen Verhältnissen für die Höhenwuchsleistung der Buche wurden mäßig trockene Sande, z. T. in Hanglagen (081, 481) oder mit hoch anstehendem Kalk (041) zusammengefasst. Zu dieser Gruppe wurden auch Buchen auf nährstoffarmen Feinlehmen (313) zugeordnet. Da lediglich zwei Buchen auf dieser Standorteinheit gefunden wurden, weist dies auf die Notwendigkeit der Berücksichtigung der absoluten Zahl der Bäume in den zusammengefassten Gruppen bei der Interpretation der Entscheidungsbäume hin.

Als Gruppe mit der günstigsten Verteilung der Höhenmessbäume in Bezug auf die eingeführte Klassengrenze wurden Buchen zusammengefasst, die auf überwiegend gut wasserversorgten Standorten zu finden waren (Knoten 1), wobei eine zusätzliche Verbesserung der Verteilung in Bezug auf die Höhenwuchsleistung erreicht werden konnte, wenn diese Gruppe nach der Dauer der Vegetationszeit weiter unterteilt wurde. Die meisten Bäume mit einer überdurchschnittlichen Höhenwuchsleistung wurden demnach auf gut wasserversorgten Standorten isoliert, die zudem in einer Gegend zu finden waren, wo die Dauer der Vegetationszeit besonders lang war (Knoten 9).

Nach dem CART-Algorithmus (Abb. 23) erfolgt die Unterteilung der Buchen auf der obersten Ebene des Entscheidungsbaumes nach der Wasserhaushaltsstufe, wobei einerseits besonders gut wasserversorgte Standorte (Knoten 2) und auf der anderen Seite Standorte mit ungünstiger oder unausgeglichener Wasserversorgung (Knoten 1) abgetrennt wurden. Das ungünstigste Verhältnis der Verteilung der Buchen wurde für trockene Sande, Tonlehme und Tone ermittelt (Knoten 4). Das beste Verhältnis hingegen konnte mit Hilfe des CART-Algorithmus für die Baumart Buche im Wuchsbezirk 6.2 für gut wasserversorgte Standorte mit normaler Trophieausstattung, sowie in Hanglagen oder bei tiefanstehendem, aber wurzelerreichbarem Kalk im Boden ermittelt werden (Knoten 6).

Nach dem C4.5-Algorithmus, dessen Anwendungsergebnisse in Textnotation in Anschluss an Abbildung 23 aufgeführt sind, wurden besonders gute Verteilungen der Höhenwuchsleistung auf Standorten der Geologie 418 (= tertiäre Rückstandsbildungen, solidifluidal oder fluviatil; ~ „lehmige Albüberdeckung“) mit überdurchschnittlich hoher Jahresdurchschnittstemperatur und guter Wasserversorgung ermittelt. Unterdurchschnittliche Höhenwuchsleistungen wurden für Buchen auf nährstoffarmen (code_stoe_2 = 1) Standorten, auf Sand-, Kies- und Geröllstandorten (code_stoe_1 = 0) oder auf trockenen (code_stoe_3 = 1) oder wechselfeuchten Standorten (code_stoe_3 = 8) ermittelt.

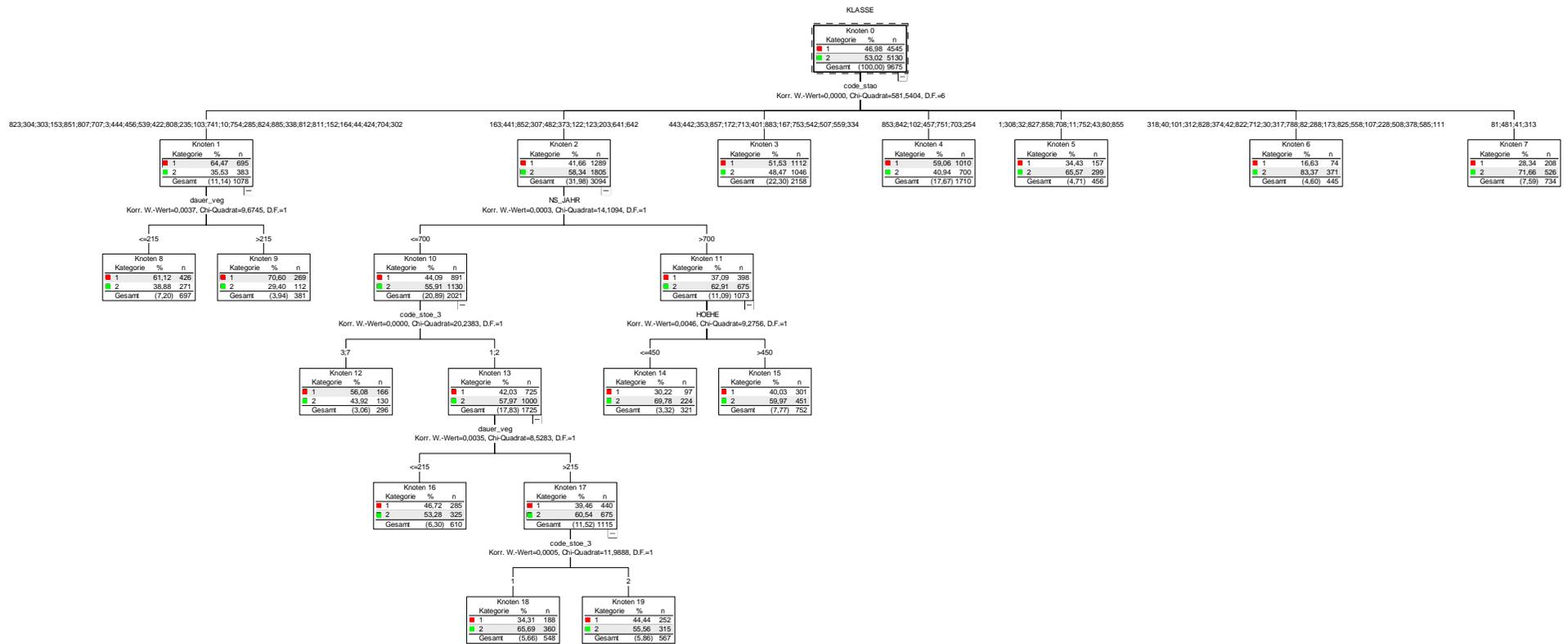


Abb. 22: Entscheidungsbaum für die Baumart Buche im Wuchsgebiet 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CHAID-Algorithmus generiert.

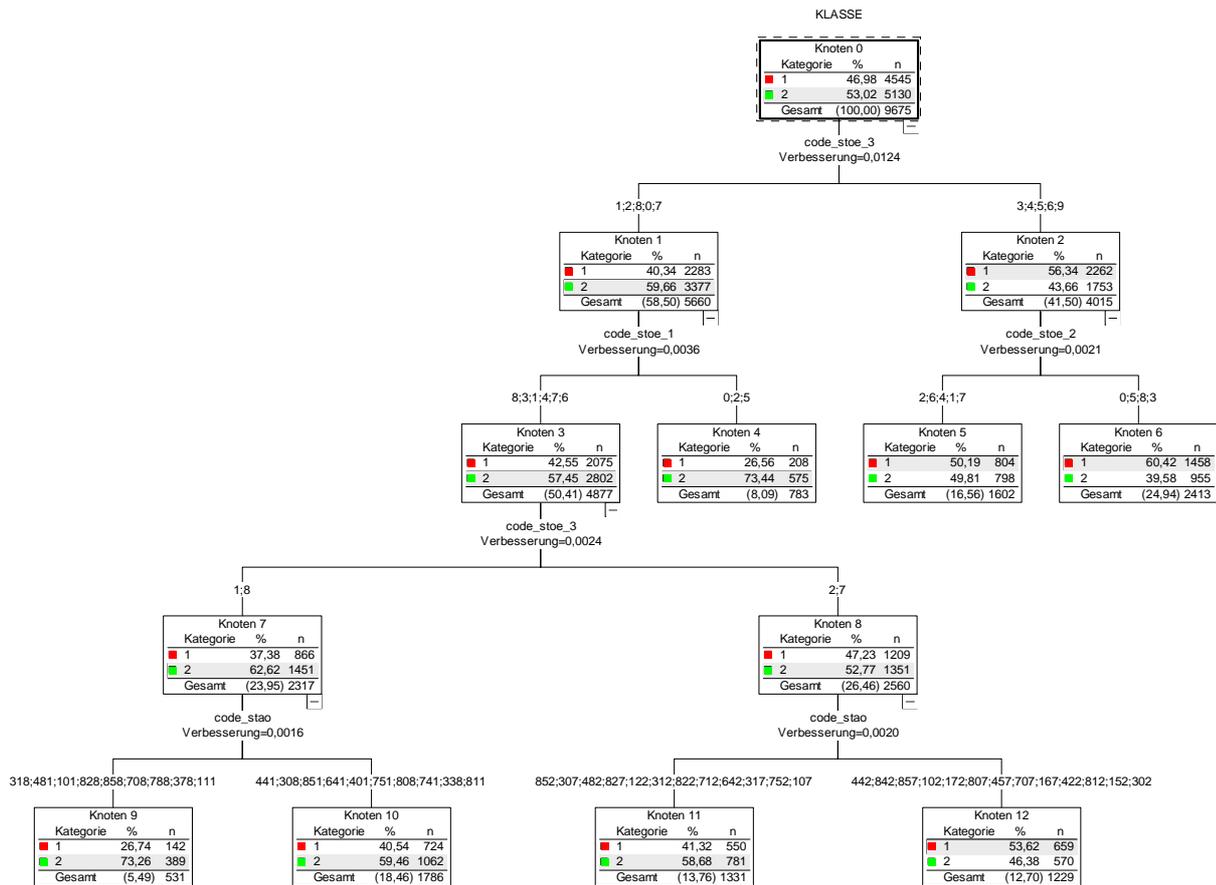


Abb. 23: Entscheidungsbaum für die Baumart Buche im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CART-Algorithmus generiert.

J48 pruned tree (Buche – WBZ 6.2)

```

code_stoe_2 = 1: 2 (117.0/34.0)
code_stoe_2 != 1
| code_stoe_1 = 0: 2 (786.0/212.0)
| code_stoe_1 != 0
| | code_stoe_3 = 8: 2 (213.0/63.0)
| | code_stoe_3 != 8
| | | code_stoe_3 = 1: 2 (2100.0/797.0)
| | | code_stoe_3 != 1
| | | | code_stao = 482: 2 (733.0/309.0)
| | | | code_stao != 482
| | | | geologie = 418
| | | | | temp_jahr = 6.5: 2 (264.0/101.0)
| | | | | temp_jahr != 6.5
| | | | | code_stoe_3 = 3: 1 (162.0/73.0)
| | | | | code_stoe_3 != 3: 2 (179.0/78.0)
| | | | | geologie != 418
| | | | | code_stoe_2 = 2: 2 (215.0/94.0)
| | | | | code_stoe_2 != 2
| | | | | code_stao = 852: 2 (159.0/69.0)
| | | | | code_stao != 852: 1 (4747.0/2048.0)

```

4.2.2 Entscheidungsbäume für die Baumart Douglasie

Unter Anwendung des CHAID-Algorithmus war keine Entscheidungsbaumgenerierung mit einer ausreichenden Anzahl von Datensätzen pro Blattelement für die Baumart Douglasie im Wuchsbezirk 6.2 möglich.

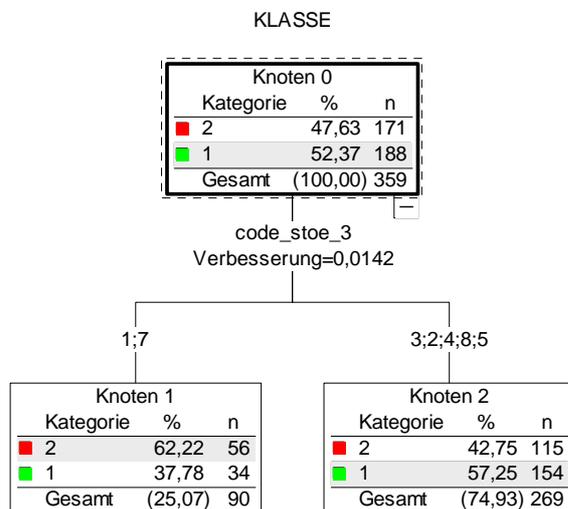


Abb. 24: Entscheidungsbaum für die Baumart Douglasie im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CART-Algorithmus generiert.

J48 pruned tree (Douglasie WBZ 6.2)

code_stoe_3 = 1: 2 (78.0/31.0)

code_stoe_3 != 1: 1 (281.0/124.0)

Sowohl mit CART (Abb. 24) als auch mit C4.5 werden Entscheidungsbäume aufgebaut, die im Wurzelknotenelement nach der Wasserhaushaltsstufe unterscheiden. Bei der Anwendung beider Algorithmen wurde eine unterdurchschnittliche Verteilung der Datensätze für Douglasien auf mäßig trockenen Standorten (code_stoe3 = 1) gefunden werden. Mit Hilfe des CART-Algorithmus werden dieser Gruppe noch die Datensätze für Douglasien auf schwach bis mäßig wechselfeuchten Standorten zugeordnet.

4.2.3 Entscheidungsbäume für die Baumart Eiche

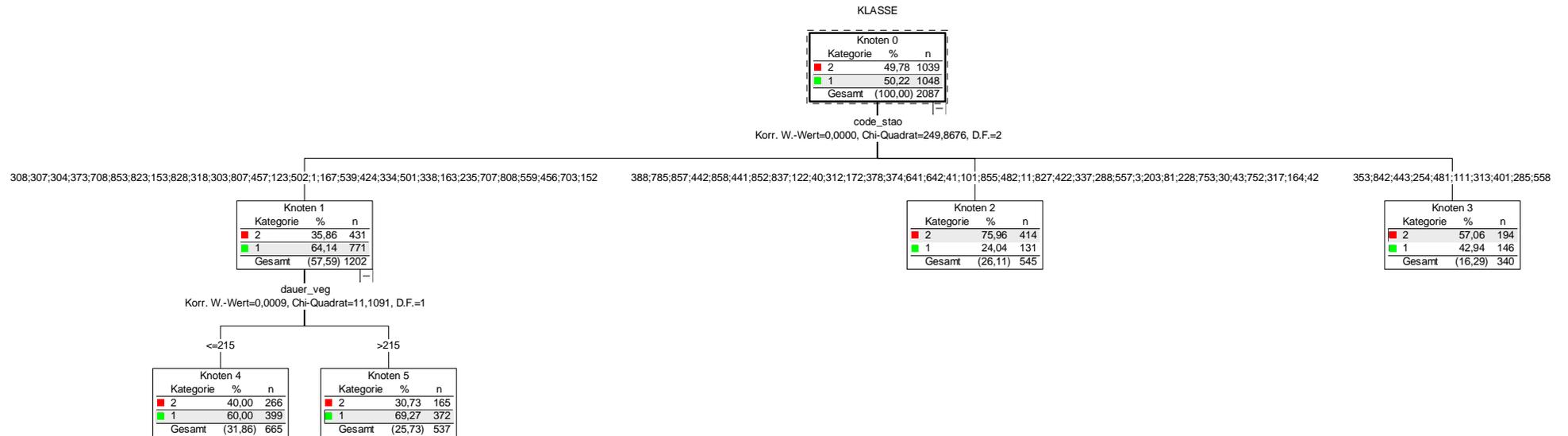


Abb. 25: Entscheidungsbaum für die Baumart Eiche im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CHAID-Algorithmus generiert.

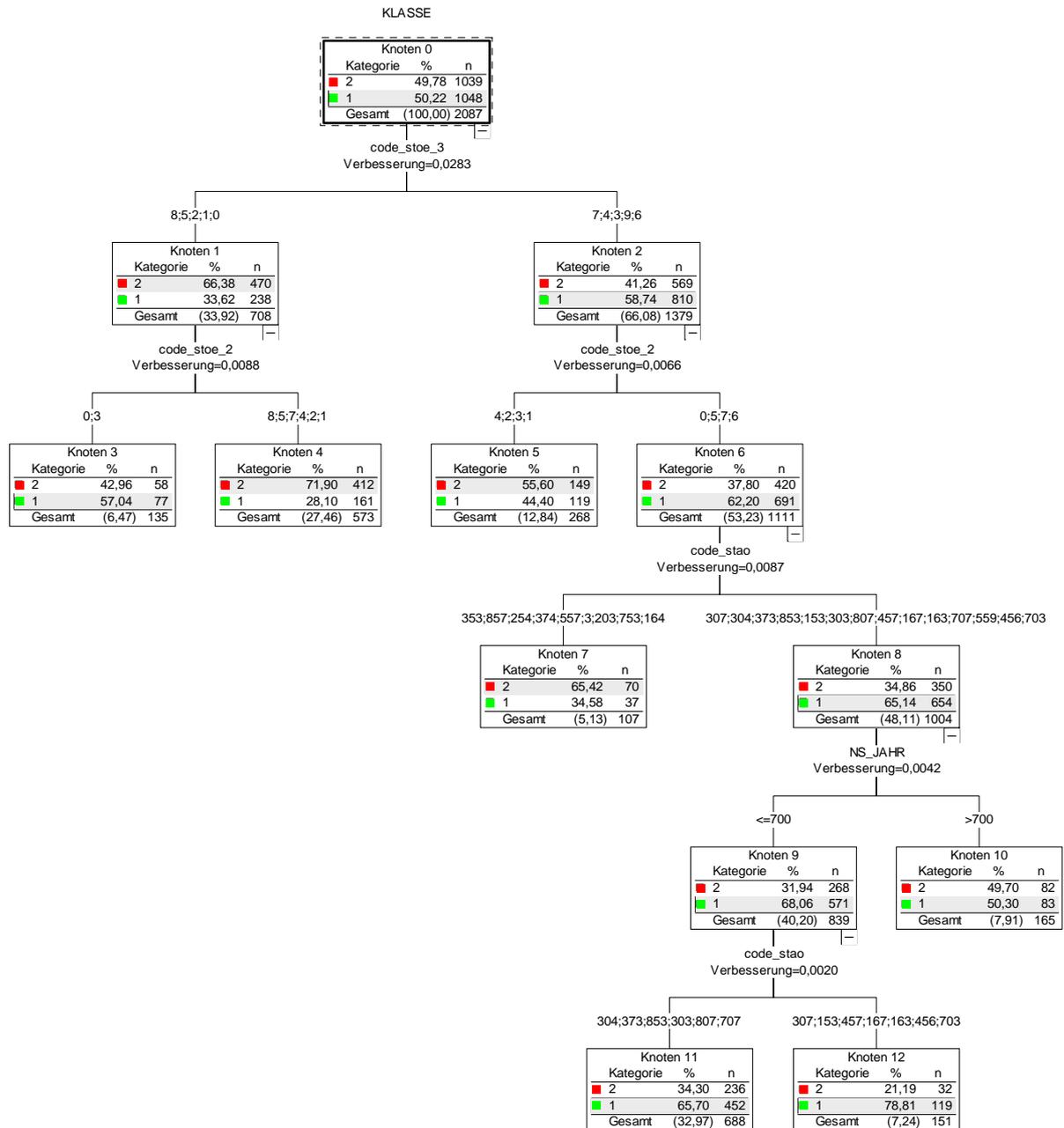


Abb. 26: Entscheidungsbaum für die Baumart Eiche im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CART-Algorithmus generiert.

J48 pruned tree (Eiche WBZ 6.2)

```

-----
code_stoe_3 = 2: 2 (263.0/73.0)
code_stoe_3 != 2
| code_stoe_3 = 1: 2 (220.0/68.0)
| code_stoe_3 != 1
| | hoehe = 550: 2 (321.0/132.0)
| | hoehe != 550
| | | code_stoe_2 = 4: 2 (171.0/72.0)
| | | code_stoe_2 != 4: 1 (1112.0/409.0)
    
```

Aufgrund der CHAID-Aufteilung (Abb. 25) wurden für die Baumart Eiche in Knoten 2 Datensätze auf eher trockenen Standorten sowie wiederum auf Standorten mit einer ausgeprägten unausgeglichener Wasserversorgung als Bäume mit unterdurchschnittlichem Höhenwachstum zusammengefasst. Besonders günstig war hingegen die Verteilung der Höhenmessbäume auf gut wasserversorgten Standorten, die zudem mit einer besonders langen Dauer der Vegetationszeit in Verbindung gebracht werden konnten (Knoten 5). Insgesamt erscheint die Zusammenfassung in diesem Fall sehr grob. Eine feinere Unterteilung wäre bei der Wahl einer geringeren Mindestzahl von Datensätzen pro Knoten- bzw. Blattelement vermutlich möglich.

CART (Abb. 26) unterteilt die Eichendatensätze im Wurzelknoten nach den Wasserhaushaltsstufen (`code_stoe_3`). Insgesamt konnten Datensätze mit einer besonders ungünstigen Verteilung der Höhenmessbäume auf trockenen oder wechselfeuchten Standorten, die zudem eine Besonderheit in der Trophie (Knoten 4) aufweisen, zusammengefasst werden. Dagegen konnte die ungünstige Wasserversorgung auf Standorten mit „normaler“ Trophieausstattung bzw. auf Feinlehmstandorten (Knoten 3) z. T. kompensiert werden. Die besten Verteilungen für die Höhenmessbäume der Baumart Eiche wurden für die Knotenelemente 9 bzw. 11 und 12 gefunden. Zusammengefasst wurden Eichen auf Standorten mit einer für diesen Wuchsbezirk mittleren Niederschlagsausstattung. Auf den gleichen Standorten brachte eine verbesserte Niederschlagsausstattung dagegen keine Verbesserung der Verteilung der Höhenmessbäume.

4.2.4 Entscheidungsbäume für die Baumart Fichte

Mit Hilfe des CHAID-Algorithmus (Abb. 27) wurde wiederum bei der sehr großen Zahl an Datensätzen ein sehr breiter Baum generiert. Besonders ungünstige Verteilungen der Höhenwuchsleistung bei der Baumart Fichte fanden sich hier für eher trockene Standorte, die zudem unterdurchschnittlich niederschlagsversorgt waren und eine besonders lange Dauer der Vegetationszeit hatten (Knoten 24). Besonders günstige Verhältnisse fanden sich dagegen wiederum für gut wasserversorgte Standorte oder Standorte mit einer guten Wasserhaltekapazität bei gleichzeitig lediglich durchschnittlicher Niederschlagsversorgung im Jahr (Blatt 27). Bei der Betrachtung aller einzelnen Knoten fällt auf, dass jeweils die Verteilung in der Höhenwuchsleistung der einzelnen Höhenmessbäume eines Knotens bei überdurchschnittlicher Niederschlagsausstattung schlechter wurde (Knoten 10, 11, 12, 13, 18, 19, 27, 28).

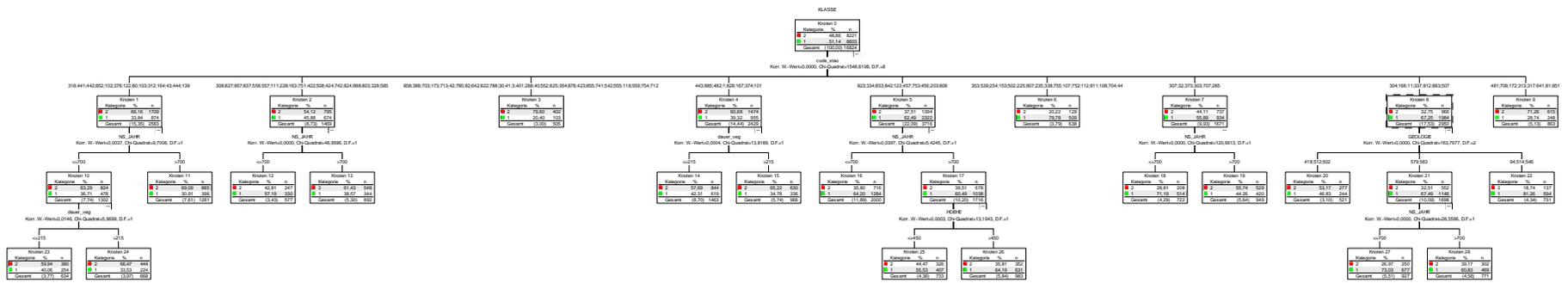


Abb. 27: Entscheidungsbaum für die Baumart Fichte im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CHAID-Algorithmus generiert.

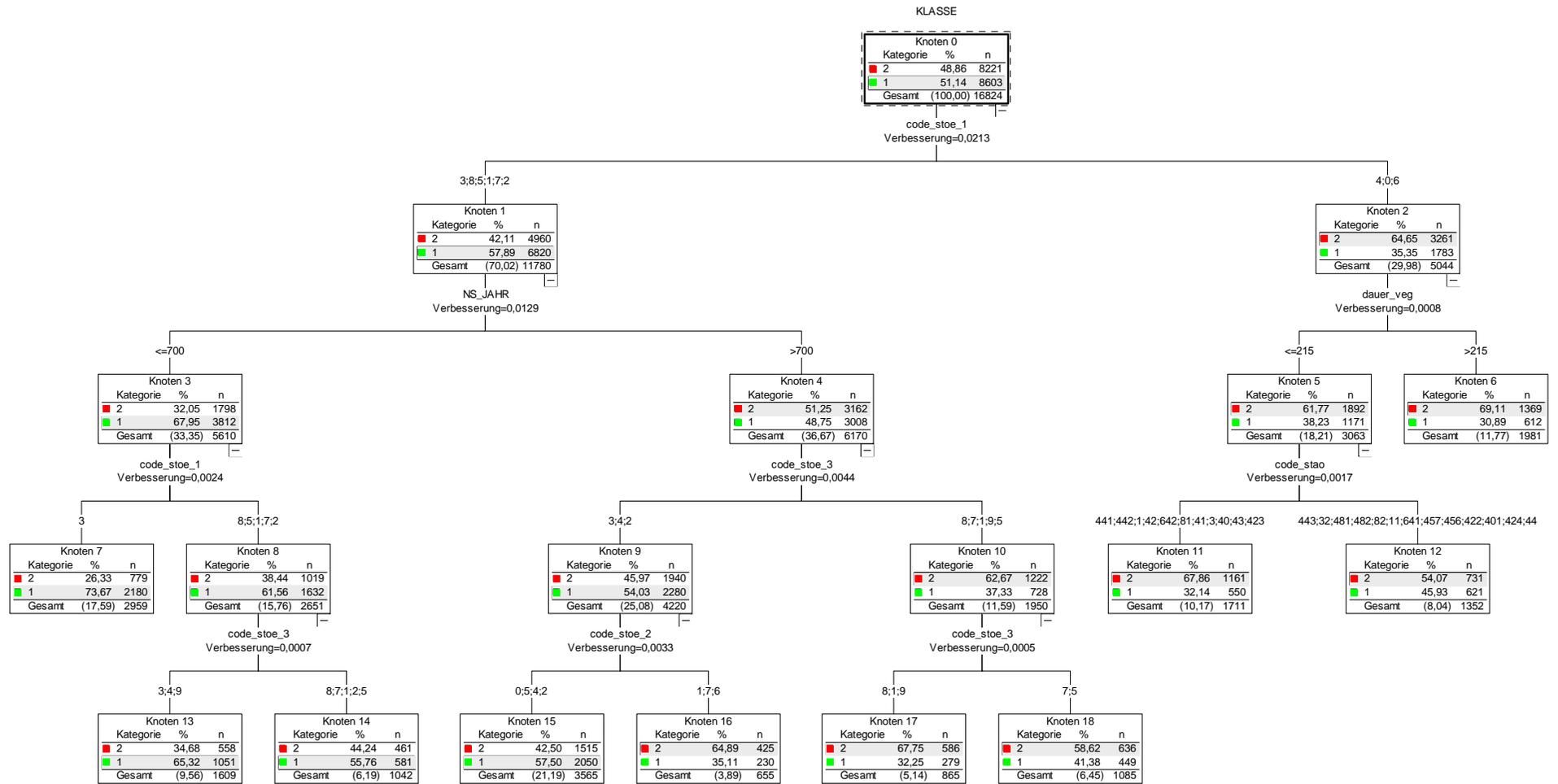


Abb. 28: Entscheidungsbaum für die Baumart Fichte im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CART-Algorithmus generiert.

J48 pruned tree (Fichte WBZ 6.2)

```

-----
geologie = 94: 1 (1015.0/249.0)
geologie != 94
| code_stoe_1 = 4: 2 (4465.0/1598.0)
| code_stoe_1 != 4
| | code_stoe_3 = 8: 2 (1064.0/408.0)
| | code_stoe_3 != 8
| | | ns_jahr = 800
| | | | code_stoe_2 = 5: 1 (1865.0/784.0)
| | | | code_stoe_2 != 5
| | | | | code_stao = 304: 1 (1222.0/548.0)
| | | | | code_stao != 304: 2 (2383.0/1004.0)
| | | ns_jahr != 800
| | | | code_stoe_1 = 3: 1 (2126.0/570.0)
| | | | code_stoe_1 != 3
| | | | | code_stoe_3 = 3: 1 (1446.0/525.0)
| | | | | code_stoe_3 != 3: 2 (1238.0/595.0)

```

Nach CART (Abb. 28) konnte eine besonders günstige Verteilung der Höhenmessbäume auf Feinlehmstandorten im Wuchsbezirk 6.2 ermittelt werden, wobei die Niederschlagsausstattung nur den durchschnittlichen Verhältnissen dieser Region entsprechen musste (Knoten 7). Besonders ungünstige Verhältnisse fanden sich dagegen für Standorte mit hoch anstehendem Kalk, insbesondere auf tonigen Substraten sowie auf trockenen Standorten (Knoten 11). Betrachtet man in diesem Entscheidungsbaum das Knotenelement 11 genauer, sieht man die Zuordnung des Standortes mit der Standorteinheit 423 (= frischer nährstoffreicher Feinlehm). Bei dieser Zuordnung gilt es wiederum zu beachten, dass lediglich zwei Höhenmessbäume in diesem Knoten dieser Standorteinheit zugeordnet werden konnten.

C4.5 fasst die Datensätze dagegen dahingehend zusammen, dass auf frischen Feinlehmstandorten im Wuchsbezirk 6.2 die Fichten-Höhenmessbäume eine überdurchschnittliche Verteilung der Höhenmessbäume aufwiesen, auf allen anderen Standorten bei überdurchschnittlicher Jahresniederschlagsmenge dagegen die Verteilung unterdurchschnittlich war. Bei für die Region durchschnittlichen Niederschlagsmengen wurden auf Feinlehmstandorten ($\text{code_stoe_1} = 3$) sowie auf frischen Standorten ($\text{code_stoe_3} = 3$) besonders günstige Verteilungen gefunden, alle übrigen Standorte erwiesen sich als unterdurchschnittlich bezüglich der Verteilung.

4.2.5 Entscheidungsbäume für die Baumart Kiefer

Auch für die Baumart Kiefer im Wuchsbezirk 6.2 untergliedert CHAID (Abb. 29) im Wurzelknoten nach Standorteinheiten. Besonders ungünstige Verhältnisse im Höhenwuchs der Kiefer fanden sich für die dem Knoten 6 zugeordneten Standorte. Für die Standorteinheiten, die dem Knoten 3 zugeordnet waren, fand sich eine besonders gute Verteilung bei Vorliegen einer hohen Jahresdurchschnittstemperatur (Knoten 12).

Nach CART (Abb. 30) wurden besonders gute Verhältnisse für Standorte mit einer hohen Jahresdurchschnittstemperatur bei gleichzeitigem Zugang zu Ton oder Kalk im Unterboden ermittelt (Knoten 6). Besonders ungünstige Verhältnisse dagegen wurden für Kiefern auf trockenen oder Standorten mit humosen Böden gefunden (Knoten 8). Die Zuordnung der Standorteinheit 824 (nährstoffreiche, sehr frische Schichtlehme) ist wiederum eine Folge geringer Anzahlen an Kiefern-Höhenmessbäumen auf dieser Standorteinheit.

Mit C4.5 konnten ähnliche standortkundliche Muster für das Höhenwachstum der Kiefer ermittelt werden, wie mit dem CART-Algorithmus. Besonders gute Verhältnisse fanden sich für Kiefern-Höhenmessbäume sowohl auf frischen Standorten als auch auf Standorten, die nicht als „trocken“ gekennzeichnet waren und die gleichzeitig lange Dauern der Vegetationsperiode sowie hohe Jahresdurchschnittstemperaturen aufwiesen. Für Standorte hingegen, auf denen Kalk wurzelerreichbar vorhanden war, wurde eine eigene Gruppe mit überdurchschnittlicher Höhenwuchsleistung isoliert.

J4.8 pruned tree (Kiefer, WBZ 6.2)

```

-----
code_stoe_2 = 5: 1 (640.0/198.0)
code_stoe_2 != 5
| code_stoe_3 = 4: 1 (422.0/148.0)
| code_stoe_3 != 4
| | code_stoe_2 = 8: 2 (247.0/76.0)
| | code_stoe_2 != 8
| | | temp_jahr = 6.5: 2 (521.0/188.0)
| | | temp_jahr != 6.5
| | | | code_stoe_3 = 1: 2 (441.0/175.0)
| | | | code_stoe_3 != 1
| | | | | dauer_veg = 215: 1 (881.0/403.0)
| | | | | dauer_veg != 215: 2 (257.0/112.0)

```

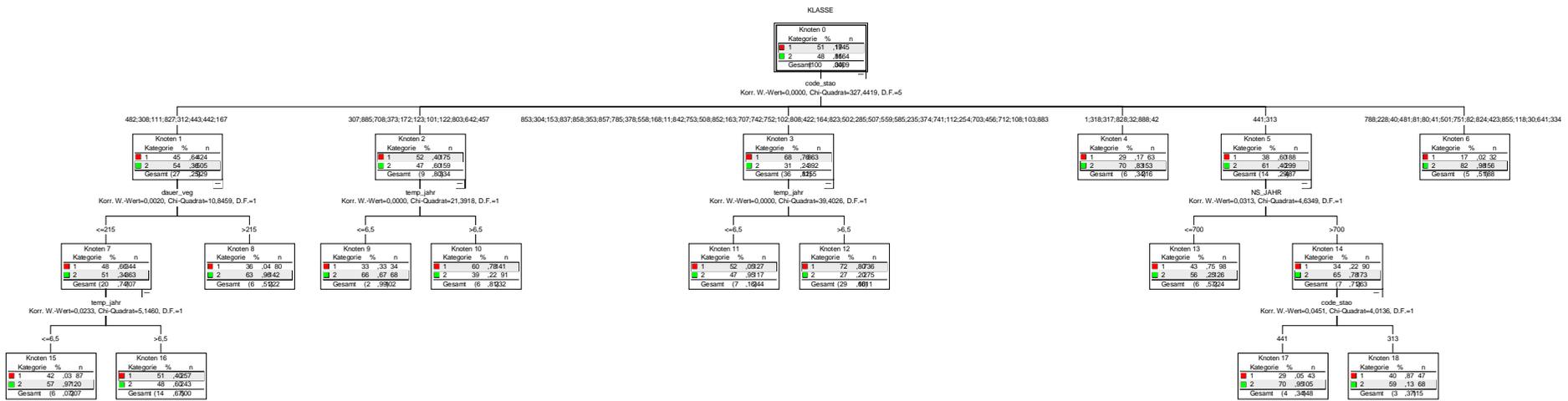


Abb. 29: Entscheidungsbaum für die Baumart Kiefer im Wuchsgebiet 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CHAID-Algorithmus generiert.

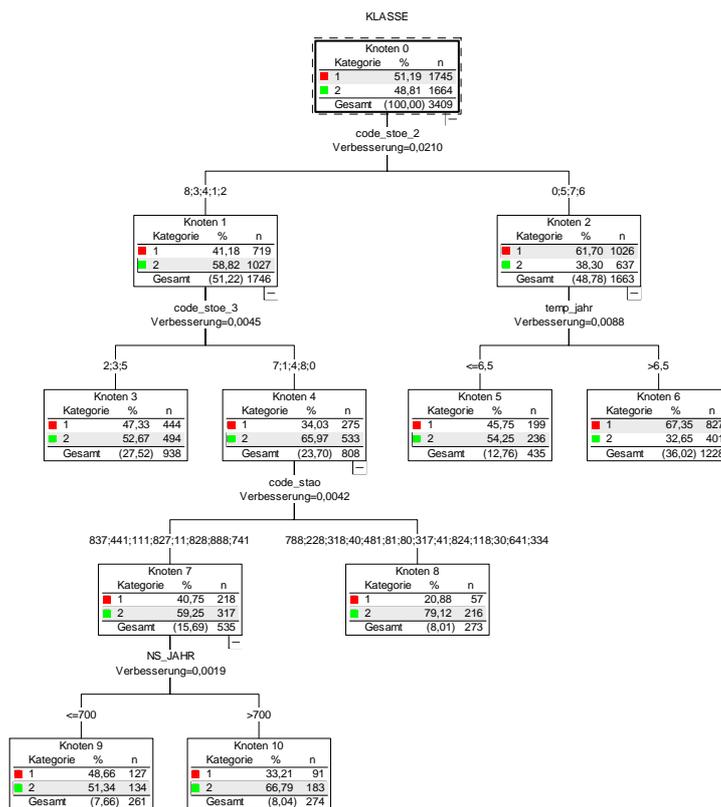


Abb. 30: Entscheidungsbaum für die Baumart Kiefer im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CART-Algorithmus generiert.

4.2.6 Entscheidungsbäume für die Baumart Lärche

Für die Baumart Lärche (Abb. 31) wurden nach CHAID, ähnlich wie für die bisher mit diesem Algorithmus erzeugten Entscheidungsbäume, bereits auf der obersten Ebene des Entscheidungsbaumes verschiedene Standorteinheiten zu Gruppen zusammengefasst. Interessant in diesem Zusammenhang ist, dass die Gruppe mit dem schlechtesten Verhältnis an über- und unterdurchschnittlichen Lärchen-Höhenmessbäumen sehr der für die Baumart Kiefer (Abb. 29) ähnelt. Das beste Verhältnis fand sich aufgeteilt nach der Geologie für diverse Standorteinheiten auf („Lößstandorten“) bzw. auf Standorten, die aus tertiärem tonigen Ausgangsmaterial hervorgegangen sind.

Auch nach CART (Abb. 32) wurde eine Gruppe, die letzterem Muster folgt, als Gruppe mit dem besten Verhältnis zwischen über- und unterdurchschnittlich hohen Bäumen, ausgeschieden (Knoten 6). Das ungünstigste Verhältnis fand sich dagegen auf Standorteinheiten, die in Knoten 12 zusammengefasst sind.

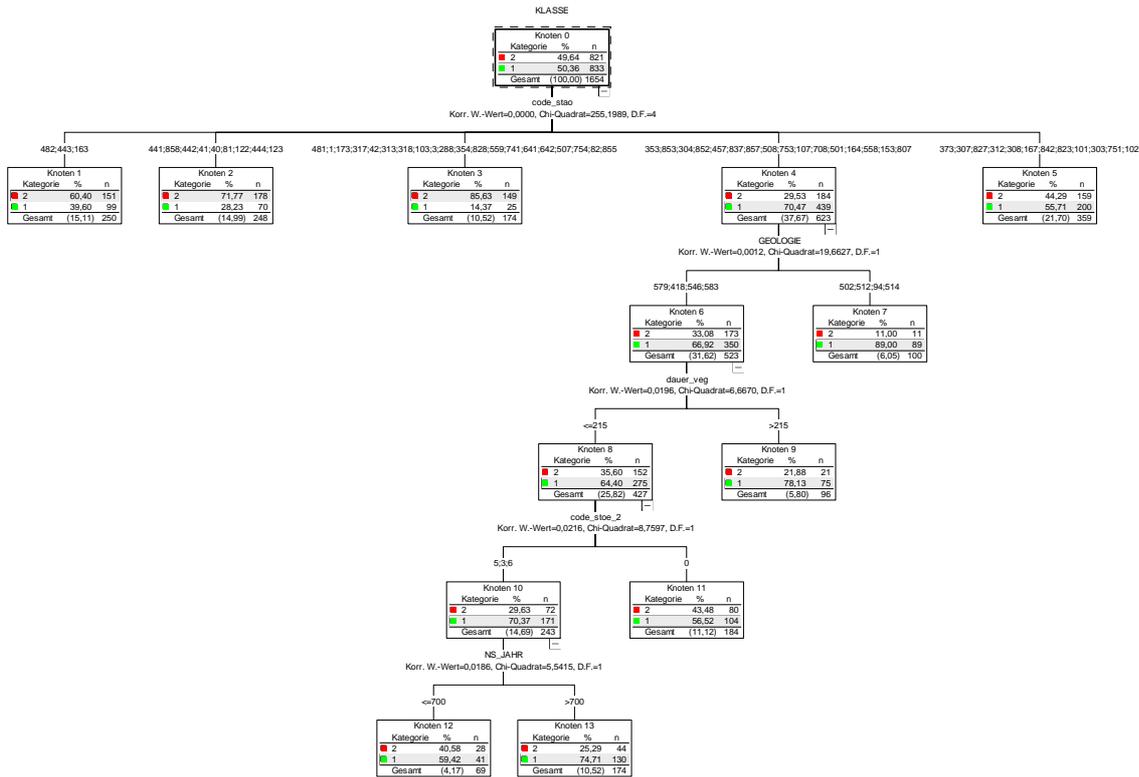


Abb. 31: Entscheidungsbaum für die Baumart Lärche im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CHAID-Algorithmus generiert.

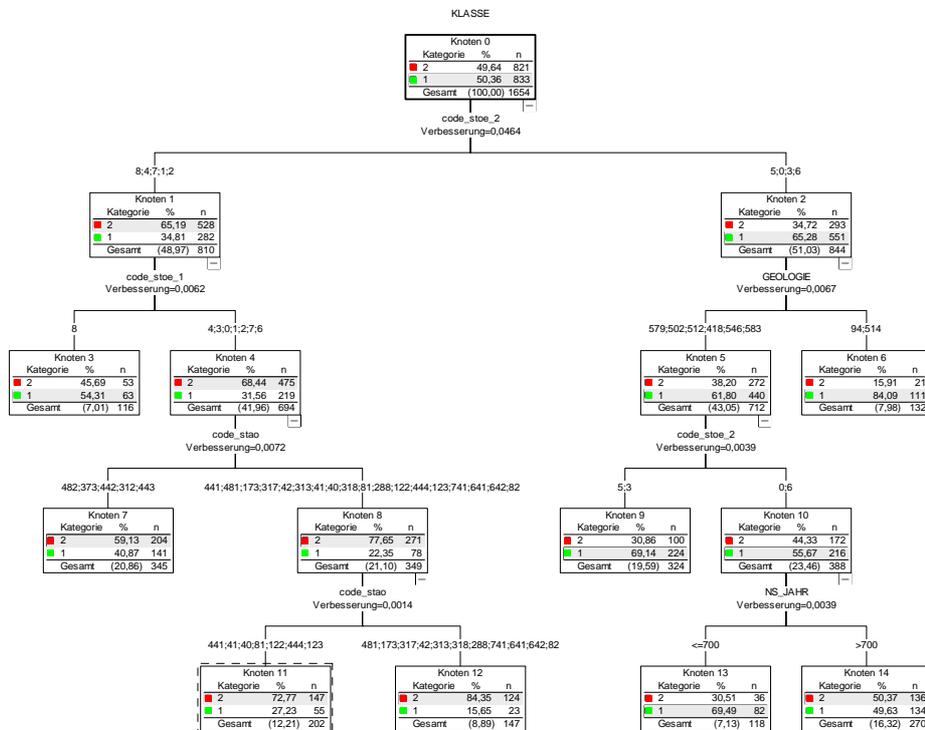


Abb. 32: Entscheidungsbaum für die Baumart Lärche im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CART-Algorithmus generiert.

J48 pruned tree (Lärche, WBZ 6.2)

```

-----
geologie = 94: 1 (132.0/25.0)
geologie != 94
| code_stoe_2 = 5: 1 (323.0/100.0)
| code_stoe_2 != 5
| | code_stoe_3 = 4: 1 (231.0/87.0)
| | code_stoe_3 != 4
| | | code_stoe_1 = 8: 1 (112.0/50.0)
| | | code_stoe_1 != 8: 2 (856.0/297.0)
    
```

Auch C4.5 nutzt das Attribut „Geologie“ zur Unterteilung der Datenmenge. Besonders günstige Verhältnisse zwischen überdurchschnittlich hohen und unterdurchschnittlich hohen Lärchen fanden sich auf Lößstandorten sowie auf „Nicht-Lößstandorten, auf denen aber Kalk wurzelerreichbar vorhanden ist, sowie auf sehr frischen Standorten oder auf Schichtlehmen mit normaler Wasserversorgung der Böden. Alle übrigen Standorte wurden nach den gewählten Voreinstellungen für den Baumaufbau (,confidenceFactor' = 0.05, Mindestanzahl der Fälle pro Blatt: 100, maximale Baumtiefe: 5) zu einer Gruppe mit unterdurchschnittlichem Höhenwachstum zusammengefasst.

4.2.7 Entscheidungsbäume für die Baumart Tanne

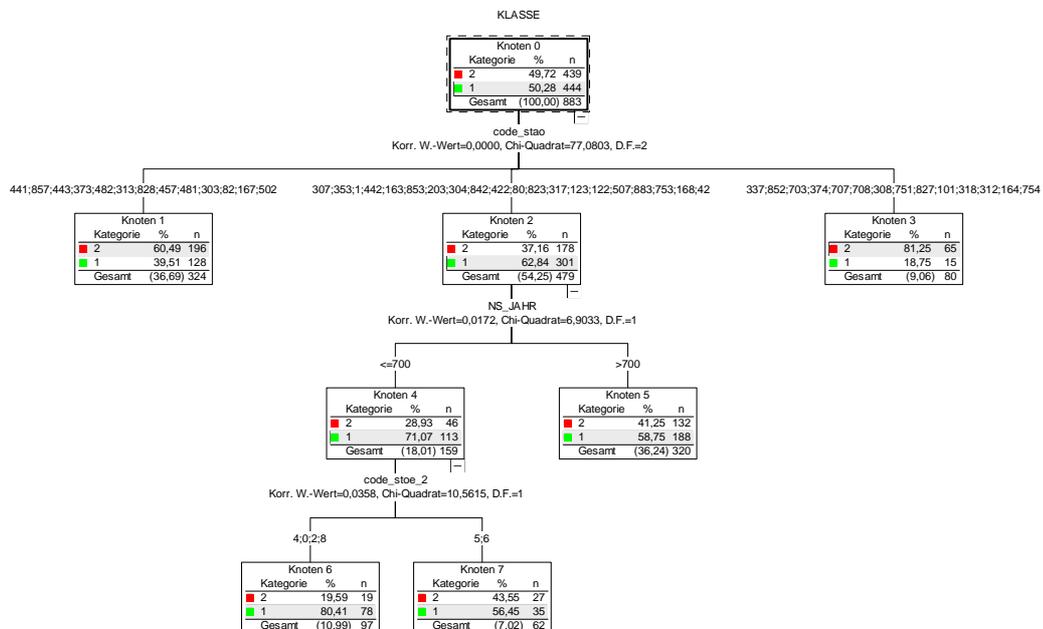


Abb. 33: Entscheidungsbaum für die Baumart Tanne im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CHAID-Algorithmus generiert.

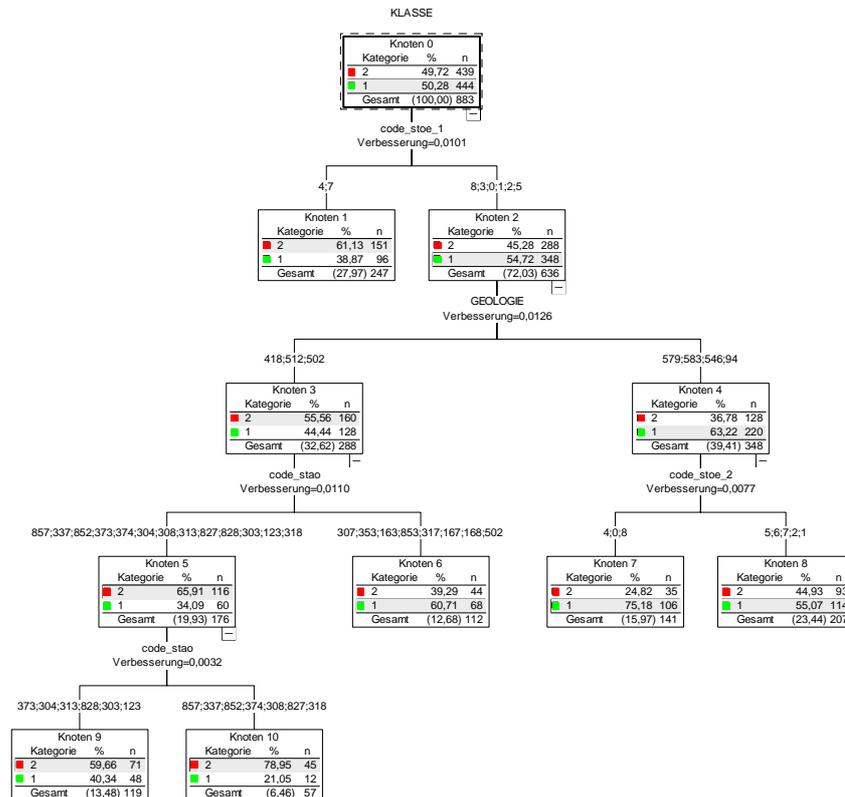


Abb. 34: Entscheidungsbaum für die Baumart Tanne im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“. Der Entscheidungsbaum wurde mit Hilfe des CART-Algorithmus generiert.

J4.8 pruned tree (Tanne, WBZ 6.2)

```
code_stoe_1 = 4: 2 (236.0/95.0)
code_stoe_1 != 4
| geologie = 418
| | code_stoe_3 = 3: 1 (114.0/51.0)
| | code_stoe_3 != 3: 2 (169.0/65.0)
| geologie != 418: 1 (364.0/143.0)
```

Nach CHAID (Abb. 33) wurden auch für die Baumart Tanne im Wuchsbezirk 6.2 auf oberster Baumbene die Standorteinheiten als Grundlage für die Klasseneinteilung benutzt. Das beste Verhältnis von über- zu unterdurchschnittlich hohen Tannen fand sich dabei für die in Knoten 2 zusammengefassten Standorte, wenn diese entweder als nährstoffreich gekennzeichnet waren oder Kalk wurzelerreichbar war oder die Standorte in Hanglagen zu finden waren und gleichzeitig die Niederschlagssituation mittleren Verhältnissen im Wuchsbezirk entsprach (Knoten 6). Die schlechtesten Verhältnisse fanden sich dagegen für die in Knoten 3 zusammengefassten Standorte.

Mit Hilfe des CART-Algorithmus (Abb. 34) wurden besonders ungünstige standörtliche Verhältnisse für das Höhenwachstum der Tanne auf den Standorteinheiten 857, 337, 852, 374, 308, 827 sowie 318 (Knoten 10), die aus

vorwiegend limnischen fluviatilen Sanden oder Kiesen hervorgegangen sind, festgestellt (Knoten 6). Besonders günstige Verhältnisse fanden sich hingegen für Standorte in Hanglage oder Standorte mit wurzelerreichbarem Kalk, die aus geologischem Ausgangsmaterial der Einheiten 579 (Malmmergel), 583 (Tonmergel des Dogger), 546 (Kreidetone) oder 94 (Löß) hervorgegangen sind.

Ein ähnliches Bild liefert die Ermittlung mit Hilfe des C4.5-Algorithmus für die Baumart Tanne. Auf Standorten, die aus solifluidalen Sanden oder Schutten hervorgegangen sind und die nicht ausreichend gut wasserversorgt waren, fand sich ein schlechtes Verhältnis zwischen überdurchschnittlich und unterdurchschnittlich hohen Tannen. Dagegen war das Verhältnis für Standorte, die nicht aus dem eben genannten Ausgangsmaterial hervorgegangen sind und deren Substrattyp nicht „Tonlehm“ war, überdurchschnittlich gut.

4.2.8 Zusammenfassung der Ergebnisse zur Entscheidungsbaumgenerierung

Zusammenfassend lässt sich festhalten, dass bei geeigneter Einstellung der Grenzwerte zur Steuerung des Baumaufbaus (Baumtiefe, Mindestanzahl der Fälle pro Blatt oder Knotenelement, Grenzwerte für die Aufteilung von Klassen etc.) mit allen angewendeten Algorithmen ähnliche Bäume mit einer ähnlichen Auswahl von Prädiktoren zur Auftrennung der Daten erstellt werden konnten.

Für die Beurteilung der Güte der Aufteilung kommt dem Expertenwissen des Nutzers erhebliche Bedeutung zu. Auch gilt es bei der Betrachtung der einem Blatt zugeordneten Fälle die absolute Anzahl der Fälle im gesamten Datensatz zu beachten, da dies insbesondere bei geringer Anzahl von Fällen häufig zu einer Zuordnung zu einer Klasse führt, die mit dem lokalen Erfahrungswissen eines Forstpraktikers nur bedingt in Einklang zu bringen ist. Diesem Aspekt ist insbesondere bei der Erstellung eines Programmsystems zur automatischen Generierung von Entscheidungsbäumen aus Forstinventurdaten ein besonderes Augenmerk zuzumessen.

4.3 Potenzielle Höhenwachstumsgänge dargestellt am Beispiel der Hauptbaumarten im Wuchsbezirk 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“

In den nachfolgenden Abschnitten werden die nach dem in Abschnitt 3.2.3 beschriebenen Verfahren ermittelten Koeffizienten der Parameter A, k und P für die in Abschnitt 4.2 dargestellten Entscheidungsbäume grafisch und numerisch dargestellt.

Die in den Tabellen aufgeführten Werte der Parameter A, k und P können direkt in SILVA zur Einstellung des Standort-Leistungsbezuges einer Baumart in einer Region verwendet werden. Einleitend wird darauf hingewiesen, dass lediglich Werte für die Endblätter der Bäume aus Abschnitt 4.2 aufgeführt sind. Diese werden in den nachfolgenden Darstellungen mit der Bezeichnung „Blatt xx“ gekennzeichnet. Die Blätter entsprechen den Endknoten der jeweiligen Entscheidungsbäume bei einer von links nach rechts beginnenden und von oben nach unten fortgeführten, laufenden Nummerierung der Knotenelemente der Abbildungen in Abschnitt 4.2 (vgl. hierzu auch Abb. 17 und 18, S. 75).

4.3.1 Potenzielle Höhenwachstumsgänge der Baumart Buche

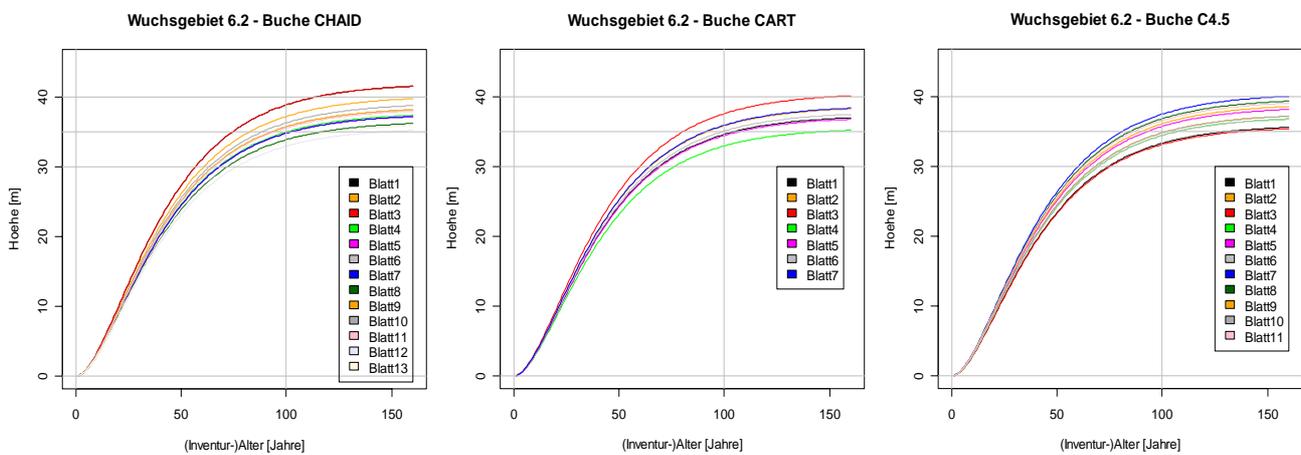


Abb. 35 – 37 (v. l. n. r.): Verläufe der Wachstumskurven für die Baumart Buche nach den in Abschnitt 4.2.1 vorgestellten Entscheidungsbäumen.

Tab. 18: Werte der Parameter A, k und P der Wachstumsfunktion nach Chapman-Richards für die in Abschnitt 4.2.1 dargestellten Entscheidungsbäume für die Baumart Buche im Wuchsbezirk 6.2. In der linken Spalte finden sich die Werte für die Entscheidungsbaumgenerierung nach CHAID, in der Mitte nach CART sowie rechts nach C4.5. Die Verläufe der Wachstumskurven finden sich in den Abbildungen 35-37.

A-Werte: Blatt1: 41,9944946 Blatt2: 40,1944946 Blatt3: 41,7944946 Blatt4: 37,7944946 Blatt5: 37,5944946 Blatt6: 38,5944946 Blatt7: 37,4944946 Blatt8: 36,5944946 Blatt9: 38,5944946 Blatt10: 39,1944946 Blatt11: 38,3944946 Blatt12: 35,5944946 Blatt13: 36,2994495	A-Werte: Blatt1: 37,3944946 Blatt2: 38,6944946 Blatt3: 40,5944946 Blatt4: 35,5944946 Blatt5: 37,1944946 Blatt6: 37,8944946 Blatt7: 38,7944946	A-Werte: Blatt1: 35,9944946 Blatt2: 37,5944946 Blatt3: 35,7944946 Blatt4: 37,1944946 Blatt5: 38,5944946 Blatt6: 37,1944946 Blatt7: 40,4944946 Blatt8: 39,7944946 Blatt9: 38,9944946 Blatt10: 37,5944946 Blatt11: 39,4944946
K-Wert: -0,03267488	K-Wert: -0,03267488	K-Wert: -0,03267488
P-Wert: 1,96944785	P-Wert: 1,96944785	P-Wert: 1,96944785

Die Wachstumskurven für die Blätter der nach den Algorithmen CHAID, CART und C4.5 generierten Entscheidungsbäume zeigen einen ähnlichen Verlauf (Abb. 35-37). Nach dem CHAID-Algorithmus (Abb. 35 bzw. Tab. 18, linke Spalte) erreichen die Buchen bei bestmöglichen standörtlichen Verhältnissen (Blatt 1) potenzielle Höhen im Alter von 50 Jahren von 27,3 m bzw. von 23,9 m bei schlechtesten standörtlichen Verhältnissen (Blatt 12). Im Alter von 100 Jahren werden potenzielle Höhen von 38,9 m für beste standörtliche Verhältnisse bzw. 33,8 m bei schlechtesten standörtlichen Verhältnissen erreicht. Demzufolge ergeben sich nach dem CHAID-Entscheidungsbaum für die Altersphase von 0 bis 50 Jahren mittlere potenzielle jährliche Höhenzuwächse zwischen 54,6 cm/ Jahr und 47,8 cm/ Jahr. Für die Altersphase zwischen 50 und 100 Jahren ergeben sich mittlere potenzielle jährliche Höhenzuwächse zwischen 23,2 cm/ Jahr und 19,8 cm/ Jahr.

Nach dem CART-Algorithmus (Abb. 36 bzw. Tab. 18, mittig) ergeben sich für die Altersphase von 0 bis 50 Jahren mittlere potenzielle jährliche Höhenzuwächse von 52,9 cm/ Jahr für beste standörtliche Verhältnisse (Blatt 3) und 46,4 cm/ Jahr für schlechteste Verhältnisse (Blatt 4). Für die Altersphase zwischen 50 und 100 Jahren ergeben sich mittlere potenzielle jährliche Höhenzuwächse zwischen 22,6 cm/ Jahr und 19,5 cm/ Jahr.

Wendet man zur Entscheidungsbaumgenerierung den C4.5-Algorithmus (Abb. 37 bzw. Tab. 18, rechte Spalte) an, so ergeben sich für die Altersphase zwischen 0 und 50 Jahren mittlere jährliche potenzielle Höhenzuwächse von 52,8 cm/ Jahr für beste standörtliche Verhältnisse (Blatt 7) und 46,6 cm/ Jahr für schlechteste standörtliche Verhältnisse (Blatt 3)

4.3.2 Potenzielle Höhenwachstumsgänge der Baumart Douglasie

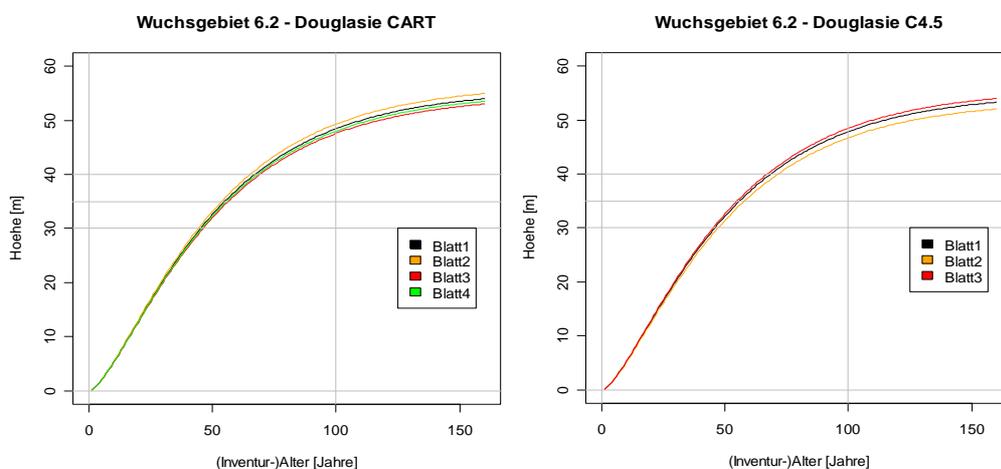


Abb. 38 – 39 (v. l. n. r.): Verläufe der Wachstumskurven für die Baumart Douglasie nach den in Abschnitt 4.2.2 vorgestellten Entscheidungsbäumen.

Tab. 19: Werte der Parameter A, k und P der Wachstumsfunktion nach Chapman-Richards für die in Abschnitt 4.2.2 dargestellten Entscheidungsbäume für die Baumart Douglasie im Wuchsbezirk 6.2. In der mittleren Spalte finden sich die Werte für die Entscheidungsbaumgenerierung nach CART sowie rechts nach C4.5. Die Verläufe der Wachstumskurven dazu finden sich in den Abbildungen 38 und 39.

Mit Hilfe des CHAID-Algorithmus konnte für die Baumart Douglasie im Wuchsbezirk 6.2 kein Entscheidungsbaum generiert werden	A-Werte: Blatt1: 55,6519337 Blatt2: 56,6519337 Blatt3: 54,6519337 Blatt4: 55,1519337	A-Werte: Blatt1: 54,9519337 Blatt2: 53,6519337 Blatt3: 55,6519337
---	K-Wert: -0,024512	K-Wert: -0,024512
---	P-Wert: 1,53644932	P-Wert: 1,53644932

Auch für die Baumart Douglasie zeigen die in Abb. 38 und 39 dargestellten Wachstumskurven einen ähnlichen Verlauf.

Nach CART (Abb. 38 bzw. Tab. 19, mittlere Spalte) erreichen die Douglasien im Wuchsbezirk 6.2 bei bestmöglichen standörtlichen Verhältnissen (Blatt 2) potenzielle Höhen von 33,2 m im Alter 50 und 49,3 m im Alter von 100 Jahren. Bei schlechtesten standörtlichen Verhältnissen (Blatt 3) können im Alter von 50 Jahren potenzielle Höhen von 32,0 m und im Alter von 100 Jahren von 47,6 m erreicht werden. Demzufolge ergeben sich potenzielle jährliche Höhenzuwächse für die Altersphase zwischen 0 und 50 Jahren von 66,4 cm/ Jahr auf besten Standorten und 64,0 cm/ Jahr auf schlechtesten Standorten. Für die Altersphase zwischen 50 Jahren und 100 Jahren ergeben sich potenzielle jährliche Höhenzuwächse zwischen 32,2 cm/ Jahr und 31,1 cm/ Jahr.

Nach C4.5 ergeben sich für die Altersphase zwischen 0 und 50 Jahren mittlere potenzielle jährliche Höhenzuwächse von 64,4 cm/ Jahr für beste standörtliche Verhältnisse (Blatt 3) und 62,9 cm/ Jahr für schlechteste standörtliche Verhältnisse (Blatt 2). Für die Altersphase zwischen 50 Jahren und 100 Jahren ergeben sich mittlere potenzielle jährliche Höhenzuwächse zwischen 31,2 cm/ Jahr und 30,5 cm/ Jahr.

4.3.3 Potenzielle Höhenwachstumsgänge der Baumart Eiche

In Abbildung 40 bis 42 dargestellt sind die potenziellen Höhenwachstumsgänge für die Baumart Eiche im Wuchsbezirk 6.2.

Auch hier ähneln die Kurvenverläufe einander sehr stark. Auffällig ist der Verlauf der Wachstumskurve für die in Blatt 4 nach CHAID zusammengefassten Eichen. Dieser Verlauf ist auf den vergleichsweise hohen Werten des Parameters A zurückzuführen. Berechnet man nach dem gleichen Muster wie für die Baumarten Buche und Douglasie durchgeführt die mittleren jährlichen potenziellen Höhenzuwächse so ergeben sich nach CHAID (Abb. 40 bzw. Tab. 20, linke Spalte) für die Altersphase

zwischen 0 und 50 Jahren für beste standörtliche Verhältnisse (Blatt 4) 56,0 cm/ Jahr bzw. 47,3 cm/ Jahr für schlechteste standörtliche Verhältnisse (Blatt 1). Für die Altersspanne zwischen 50 und 100 Jahren errechnen sich mittlere potenzielle jährliche Höhenzuwächse von 22,2 cm/ Jahr für beste Verhältnisse und 18,8 cm/ Jahr für schlechteste Verhältnisse.

Nach CART (Abb. 41 bzw. Tab. 20, mittlere Spalte) errechnen sich mittlere potenzielle jährliche Höhenzuwächse für die Altersspanne von 0 bis 50 Jahren zwischen 54,2 cm/ Jahr und 48,5 cm/ Jahr bzw. für die Altersspanne zwischen 50 Jahren und 100 Jahren zwischen 21,5 cm/ Jahr und 19,3 cm/ Jahr.

Bei Anwendung des C4.5-Algorithmus zur Entscheidungsbaumgenerierung (Abb. 42 bzw. Tab. 20, rechte Spalte) ergeben sich mittlere potenzielle jährliche Höhenzuwächse für die Altersspanne von 0 bis 50 Jahren zwischen 53,2 cm/ Jahr und 47,4 cm/ Jahr bzw. für die Altersspanne von 50 bis 100 Jahren zwischen 21,2 cm/ Jahr und 18,8 cm/ Jahr.

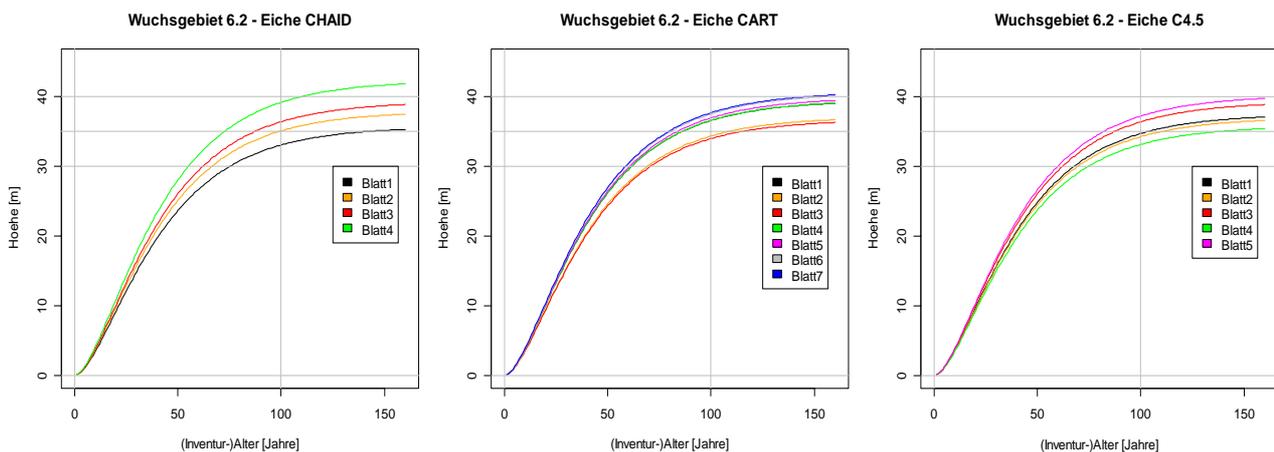


Abb. 40– 42 (v. l. n. r.): Verläufe der Wachstumskurven für die Baumart Eiche nach den in Abschnitt 4.2.3 vorgestellten Entscheidungsbäumen.

Tab. 20: Werte der Parameter A, k und P der Wachstumsfunktion nach Chapman-Richards für die in Abschnitt 4.2.3 dargestellten Entscheidungsbäume für die Baumart Eiche im Wuchsbezirk 6.2. In der linken Spalte finden sich die Werte für die Entscheidungsbaumgenerierung nach CHAID, in der Mitte nach CART sowie rechts nach C4.5. Die Verläufe der Wachstumskurven dazu finden sich in den Abbildungen 40-42.

A-Werte: Blatt1: 35,7066665 Blatt2: 37,9066665 Blatt3: 39,3066665 Blatt4: 42,3066665	A-Werte: Blatt1: 39,5066665 Blatt2: 37,1066665 Blatt3: 36,7066665 Blatt4: 39,5066665 Blatt5: 39,9066665 Blatt6: 40,5066665 Blatt7: 40,9066665	A-Werte: Blatt1: 37,5066665 Blatt2: 37,0066665 Blatt3: 39,3066665 Blatt4: 35,8066665 Blatt5: 40,2066665
K-Wert: -0,0316884	K-Wert: -0,0316884	K-Wert: -0,0316884
P-Wert: 1,78891325	P-Wert: 1,78891325	P-Wert: 1,78891325

4.3.4 Potenzielle Höhenwachstumsgänge der Baumart Fichte

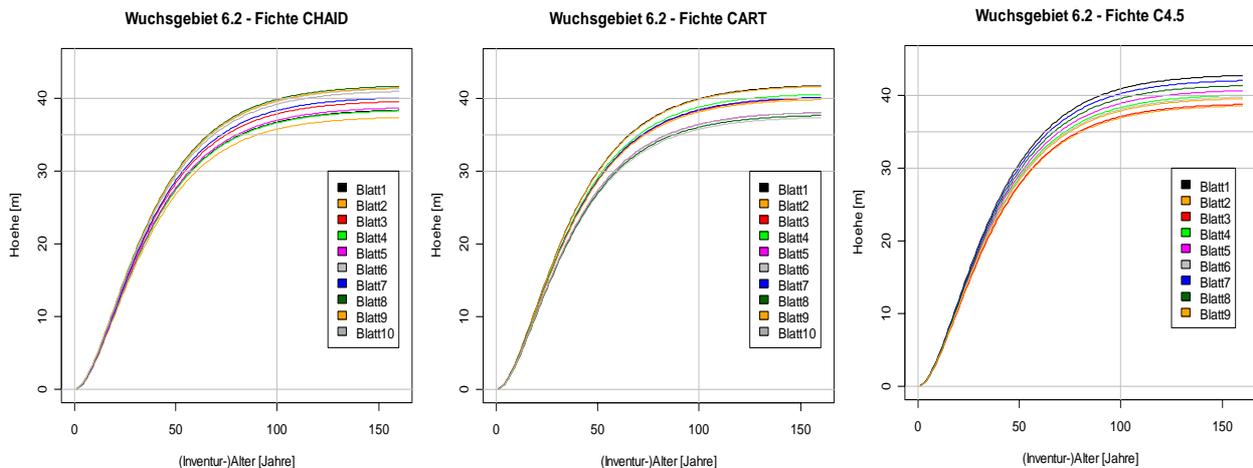


Abb. 43 – 45 (v. l. n. r.): Verläufe der Wachstumskurven für die Baumart Fichte nach den in Abschnitt 4.2.4 vorgestellten Entscheidungsbäumen.

Tab. 21: Werte der Parameter A, k und P der Wachstumsfunktion nach Chapman-Richards für die in Abschnitt 4.2.4 dargestellten Entscheidungsbäume für die Baumart Fichte im Wuchsbezirk 6.2. In der linken Spalte finden sich die Werte für die Entscheidungsbaumgenerierung nach CHAID, in der Mitte nach CART sowie rechts nach C4.5. Die Verläufe der Wachstumskurven dazu finden sich in den Abbildungen 43-45.

A-Werte:	A-Werte:	A-Werte:
Blatt1: 38,5652013	Blatt1: 41,9652013	Blatt1: 42,9652013
Blatt2: 37,5652013	Blatt2: 41,8652013	Blatt2: 38,7652013
Blatt3: 39,7652013	Blatt3: 40,2652013	Blatt3: 38,9652013
Blatt4: 38,4652013	Blatt4: 40,7652013	Blatt4: 40,2652013
Blatt5: 38,8652013	Blatt5: 38,2652013	Blatt5: 40,8652013
Blatt6: 41,5652013	Blatt6: 37,5652013	Blatt6: 39,9652013
Blatt7: 40,2652013	Blatt7: 40,3652013	Blatt7: 42,2652013
Blatt8: 41,8652013	Blatt8: 37,8651013	Blatt8: 41,5652013
Blatt9: 41,6652103	Blatt9: 40,0652013	Blatt9: 39,7652013
Blatt10: 41,1652013	Blatt10: 38,2652013	
K-Wert: -0,03721807	K-Wert: -0,03721807	K-Wert: -0,03721807
P-Wert: 1,99678795	P-Wert: 1,99678795	P-Wert: 1,99678795

In den Abbildungen 43 bis 45 dargestellt sind die Verläufe der potenziellen Wachstumskurven für die Baumart Fichte im Wuchsbezirk 6.2. In Tabelle 21 befinden sich die dazugehörigen Werte der Parameter der Wachstumsfunktion nach Chapman-Richards. Bei der Betrachtung der Werte des Parameters A in dieser Tabelle fällt auf, dass dieser für die Differenzen zwischen größtem und kleinstem Wert lediglich zwischen 4,2 und 4,4 zwischen den drei Algorithmen variiert. Dies deutet wiederum auf einen sehr ähnlichen Verlauf der Wachstumskurven hin, da, wie in Abschnitt 3.2.3 erläutert, lediglich der Parameter A zur Ermittlung der Parametersätze für die Blätter des zugrundeliegenden Entscheidungsbaumes, variiert wurde.

Nach dem CHAID-Algorithmus (Abb. 43 bzw. Tab. 21, linke Spalte) werden für die Altersspanne zwischen 0 und 50 Jahren für beste standörtliche Verhältnisse potenzielle Höhenzuwächse von 59,7 cm/ Jahr bzw. von 53,6 cm/ Jahr für schlechteste standörtliche Verhältnisse ermittelt. In der Altersphase zwischen 50 und 100 Jahren werden mittlere jährliche potenzielle Höhenzuwächse zwischen 20,0 cm/ Jahr und 17,9 cm/ Jahr erreicht.

Wendet man den CART-Algorithmus zur Entscheidungsbaumgenerierung an, so errechnen sich für die besten standörtlichen Verhältnisse mittlere jährliche potenzielle Höhenzuwächse für die Altersphase zwischen 0 und 50 Jahren zwischen 59,8 cm/ Jahr und 53,6 cm/ Jahr bzw. ebenfalls zwischen 20,0 cm/ Jahr und 17,9 cm/ Jahr für die Altersphase zwischen 50 und 100 Jahren.

Die Anwendung des C4.5-Algorithmus führt zu mittleren potenziellen jährlichen Höhenzuwächsen in der Altersphase zwischen 0 und 50 Jahren zwischen 61,3 cm/ Jahr und 55,2 cm/ Jahr bzw. in der Altersphase zwischen 50 und 100 Jahren zwischen 20,5 cm/ Jahr und 18,5 cm/ Jahr.

4.3.5 Potenzielle Höhenwachstumsgänge der Baumart Kiefer

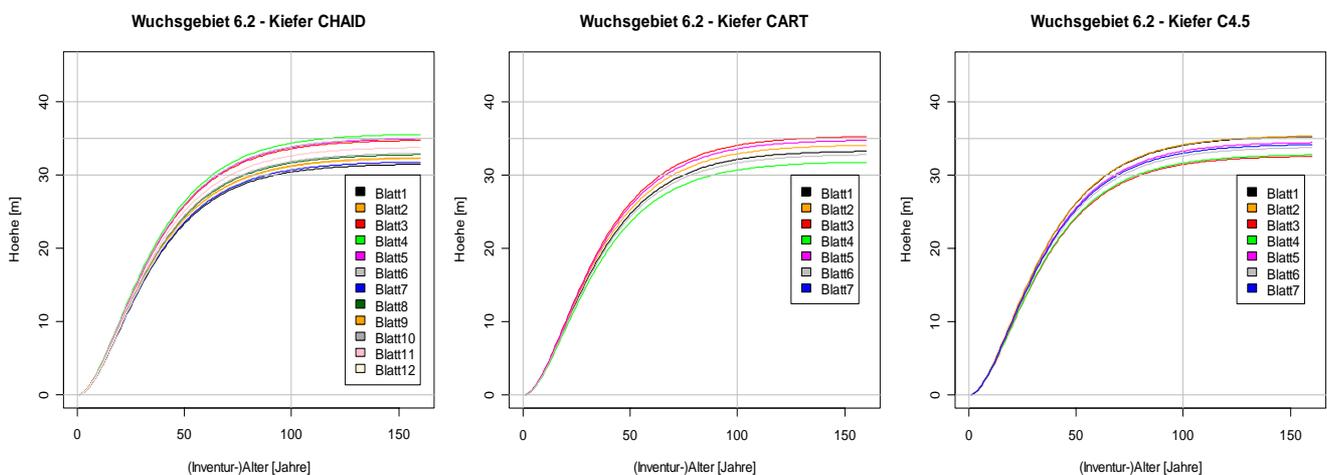


Abb. 46 – 48 (v. l. n. r.): Verläufe der Wachstumskurven für die Baumart Kiefer nach den in Abschnitt 4.2.5 vorgestellten Entscheidungsbäumen.

Auch für die Baumart Kiefer verlaufen die Kurven in den Abbildungen 46 bis 48 aufgrund der in Tabelle 22 dargestellten Parameterwerte sehr ähnlich.

Nach dem CHAID-Entscheidungsbaum (Abb. 46 bzw. Tab. 22, linke Spalte) ergeben sich für die Altersphase von 0 bis 50 Jahren mittlere potenzielle jährliche Höhenzuwächse zwischen 53,0 cm/ Jahr und 46,7 cm/ Jahr bzw. für die Altersphase zwischen 50 und 100 Jahren zwischen 15,9 cm/ Jahr und 14,1 cm/ Jahr.

Wendet man den CART-Algorithmus an, so errechnen sich für beste bzw. schlechteste Kiefernstandorte in der Altersphase zwischen 0 und 50 Jahren mittlere

potenzielle jährliche Höhenzuwächse zwischen 52,0 cm/ Jahr und 47,2 cm/ Jahr bzw. für die Altersphase zwischen 50 und 100 Jahren zwischen 15,8 cm/ Jahr und 14,2 cm/ Jahr. Mit Werten für die Altersphase von 0 bis 50 Jahren zwischen 52,5 cm/ Jahr und 48,3 cm/ Jahr bzw. für die Altersphase zwischen 50 und 100 Jahren zwischen 15,8 cm/ Jahr und 14,6 cm/ Jahr konnten, nach Anwendung des C4.5-Algorithmus, ähnliche potenzielle Höhenzuwächse ermittelt werden.

Tab. 22: Werte der Parameter A, k und P der Wachstumsfunktion nach Chapman-Richards für die in Abschnitt 4.2.5 dargestellten Entscheidungsbäume für die Baumart Kiefer im Wuchsbezirk 6.2. In der linken Spalte finden sich die Werte für die Entscheidungsbaumgenerierung nach CHAID, in der Mitte nach CART sowie rechts nach C4.5. Die Verläufe der Wachstumskurven dazu finden sich in den Abbildungen 46-48.

A-Werte: Blatt1: 31,5702707 Blatt2: 32,3702707 Blatt3: 34,8702707 Blatt4: 35,6702707 Blatt5: 35,0702707 Blatt6: 35,1702707 Blatt7: 31,8702707 Blatt8: 32,8702707 Blatt9: 32,4702707 Blatt10: 33,0702707 Blatt11: 33,8702707 Blatt12: 32,1702707	A-Werte: Blatt1: 33,3702707 Blatt2: 34,1702707 Blatt3: 35,3702707 Blatt4: 31,8702707 Blatt5: 34,8702707 Blatt6: 32,8702707	A-Werte: Blatt1: 35,3702707 Blatt2: 35,4702707 Blatt3: 32,6702707 Blatt4: 32,8702707 Blatt5: 34,5702707 Blatt6: 33,8702707 Blatt7: 34,2702707
K-Wert: -0,04045919	K-Wert: -0,04045919	K-Wert: -0,04045919
P-Wert: 2,11476695	P-Wert: 2,11476695	P-Wert: 2,11476695

4.3.6 Potenzielle Höhenwachstumsgänge der Baumart Lärche

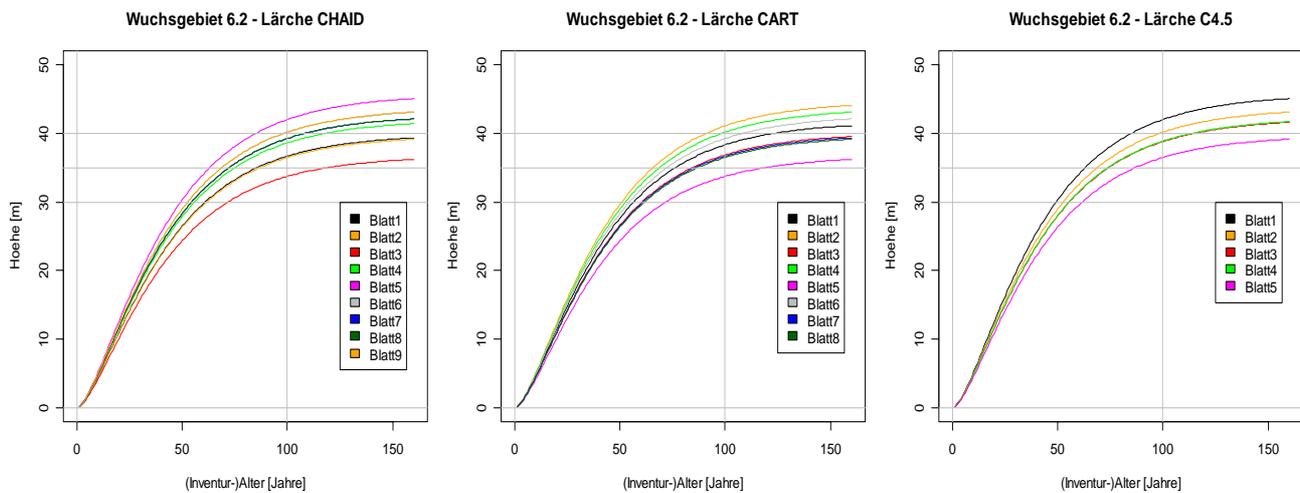


Abb. 49 –51 (v. l. n. r.): Verläufe der Wachstumskurven für die Baumart Lärche nach den in Abschnitt 4.2.6 vorgestellten Entscheidungsbäumen.

Tab. 23: Werte der Parameter A, k und P der Wachstumsfunktion nach Chapman-Richards für die in Abschnitt 4.2.6 dargestellten Entscheidungsbäume für die Baumart Lärche im Wuchsbezirk 6.2. In der linken Spalte finden sich die Werte für die Entscheidungsbaumgenerierung nach CHAID, in der Mitte nach CART sowie rechts nach C4.5. Die Verläufe der Wachstumskurven dazu finden sich in den Abbildung 49-51.

A-Werte: Blatt1: 39,9054384 Blatt2: 39,7054384 Blatt3: 36,7054384 Blatt4: 42,0054384 Blatt5: 45,7054384 Blatt6: 43,7054384 Blatt7: 42,7054384 Blatt8: 42,7054384 Blatt9: 43,7054384	A-Werte: Blatt1: 41,7054384 Blatt2: 44,7054384 Blatt3: 40,1054384 Blatt4: 43,7054384 Blatt5: 36,7054384 Blatt6: 42,7054384 Blatt7: 39,9054384 Blatt8: 39,7054384	A-Werte: Blatt1: 45,7054384 Blatt2: 43,7054384 Blatt3: 42,2054384 Blatt4: 42,3054384 Blatt5: 39,7054384
K-Wert: -0,02979388	K-Wert: -0,02979388	K-Wert: -0,02979388
P-Wert: 1,59438227	P-Wert: 1,59438227	P-Wert: 1,59438227

In den Abbildungen 49 bis 51 sind die Kurvenverläufe der potenziellen Höhenwachstumsgänge für die Baumart Lärche im Wuchsbezirk 6.2 dargestellt. Auffällig ist der vergleichsweise hohe Verlauf des potenziellen Wachstumsgangs für die schlechtesten standörtlichen Verhältnisse, die nach der Anwendung des C4.5-Algorithmus isoliert werden konnten.

Nach der Anwendung des CHAID-Algorithmus (Abb. 49 bzw. Tab. 23, linke Spalte) ergeben sich für die Altersphase zwischen 0 und 50 Jahren mittlere potenzielle Höhenzuwächse zwischen 60,8 cm/ Jahr (beste Standortsgruppe) und 49,2 cm/ Jahr (schlechteste Standortsgruppe) bzw. in der Altersphase zwischen 50 und 100 Jahren zwischen 23,0 cm/ Jahr und 18,7 cm/ Jahr.

Bei Anwendung des CART-Algorithmus (Abb. 50 bzw. Tab. 23, mittlere Spalte) errechnen sich für die Altersphase zwischen 0 und 50 Jahren mittlere potenzielle Höhenzuwächse zwischen 59,8 cm/ Jahr und 49,0 cm/ Jahr bzw. für die Altersphase zwischen 50 und 100 Jahren zwischen 22,8 cm/ Jahr und 18,7 cm/ Jahr.

Legt man den mit Hilfe des C4.5-Algorithmus generierten Entscheidungsbaum der Berechnung zugrunde, so errechnen sich mittlere potenzielle jährliche Höhenzuwächse für die Altersphase zwischen 0 und 50 Jahren zwischen 60,8 cm / Jahr und 51,8 cm/ Jahr bzw. zwischen 23,4 cm/ Jahr und 20,2 cm/ Jahr.

4.3.7 Potenzielle Höhenwachstumsgänge der Baumart Tanne

Abschließend sind in den Abbildungen 52 bis 54 die potenziellen Höhenwachstumsgänge für die Baumart Tanne im Wuchsbezirk 6.2 dargestellt. Die dazugehörigen Parameterwerte finden sich in Tabelle 24.

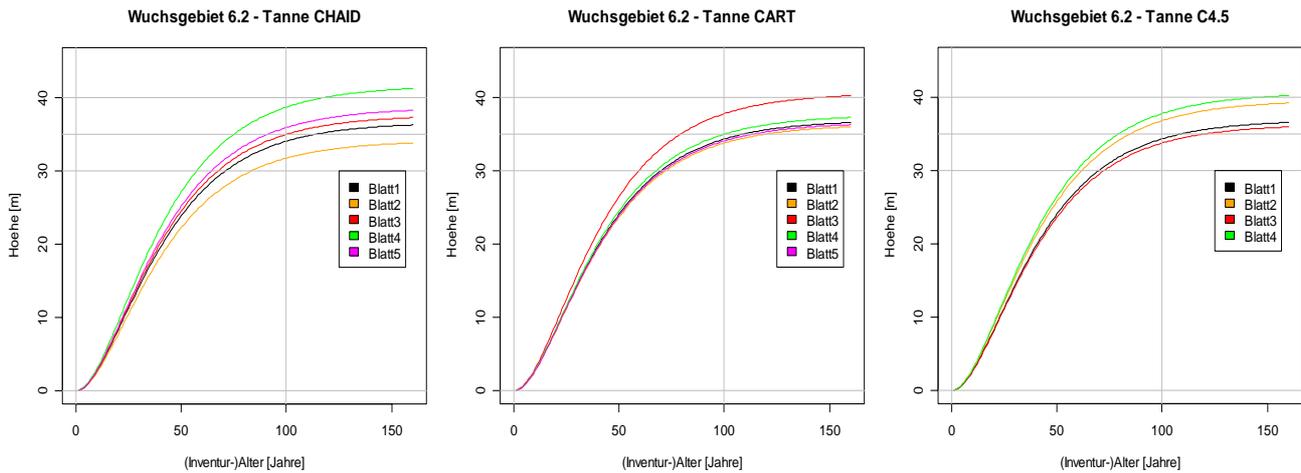


Abb. 52 – 54 (v. l. n. r.): Verläufe der Wachstumskurven für die Baumart Tanne nach den in Abschnitt 4.2.7 vorgestellten Entscheidungsbäumen.

Tab. 24: Werte der Parameter A, k und P der Wachstumsfunktion nach Chapman-Richards für die in Abschnitt 4.2.7 dargestellten Entscheidungsbäume für die Baumart Tanne im Wuchsbezirk 6.2. In der linken Spalte finden sich die Werte für die Entscheidungsbaumgenerierung nach CHAID, in der Mitte nach CART sowie rechts nach C4.5. Die Verläufe der Wachstumskurven dazu finden sich in den Abbildung 52 - 54.

A-Werte: Blatt1: 36,6302808 Blatt2: 34,1302808 Blatt3: 37,6302808 Blatt4: 41,6302808 Blatt5: 38,6302808	A-Werte: Blatt1: 36,9301808 Blatt2: 36,3301808 Blatt3: 40,6301808 Blatt4: 37,6301808 Blatt5: 36,6301808	A-Werte: Blatt1: 36,3302808 Blatt2: 39,6302808 Blatt3: 36,3302808 Blatt4: 40,6302808
K-Wert: -0,03384367	K-Wert: -0,03384367	K-Wert: -0,03384367
P-Wert: 2,10003721	P-Wert: 2,10003721	P-Wert: 2,10003721

Errechnet man auf Basis der in den Abbildungen 52 bis 54 dargestellten Kurven die potenziellen jährlichen Höhenzuwächse so ergibt sich nach der Anwendung des CHAID-Algorithmus (Abb. 52 bzw. Tab. 24, linke Spalte) für die Altersphase zwischen 0 und 50 Jahren eine Wertespanne zwischen 54,2 cm/ Jahr und 44,4 cm/ Jahr. Für die Altersphase zwischen 50 und 100 Jahren ergeben sich potenzielle mittlere Höhenzuwachswerte zwischen 23,2 cm/ Jahr und 19,0 cm/ Jahr.

Aufgrund identischer Werte für beste und schlechteste standörtliche Verhältnisse (vgl. Tab. 24, mittlere und rechte Spalte) ergeben sich sowohl nach der Anwendung des CART als auch des C4.5-Algorithmus zur Entscheidungsbaumgenerierung potenzielle jährliche Höhenzuwächse zwischen 52,9 cm/ Jahr und 47,3 cm/ Jahr für die Altersphase zwischen 0 und 50 Jahren bzw. zwischen 22,6 cm/ Jahr und 20,2 cm/ Jahr für die Altersphase zwischen 50 und 100 Jahren.

4.3.8 Zusammenfassung der Ergebnisse zu den dargestellten potenziellen Höhenwachstumsverläufen

Zusammenfassend lässt sich für den Vergleich der in den Abschnitten 4.3.1 bis 4.3.8 dargestellten Wachstumskurven feststellen, dass über das beschriebene Verfahren in sich schlüssige Kurvenscharen für alle in die Untersuchung einbezogenen Baumarten im Wuchsbezirk 6.2 erzeugt werden konnten.

Von besonderer Bedeutung erscheint die Betrachtung der Werte für den Parameter A, da dieser die maximal erreichten Baumhöhen auf den isolierten Standortgruppen in dieser Region beschreibt. Bei einer Einstellung des Standort-Leistungsbezuges eines Baumart in SILVA stellt der Wert des Parameters A gleichzeitig die größte erreichbare Baumhöhe in Simulationsläufen dar. Die höchsten Werte für den Parameter A (Asymptote) fanden sich für die Baumart Douglasie (vgl. Tab. 19), gefolgt von Lärche (vgl. Tab. 23) und Fichte (vgl. Tab. 21). Die niedrigsten Werte für den Parameter A fanden sich für die Baumart Kiefer (vgl. Tab. 22). Bei der gleichen Baumart variierten die Werte des Parameters A zwischen höchstem und niedrigstem Wert in Abhängigkeit von dem verwendeten Entscheidungsbaum-Algorithmus zwischen 3,8 m und 5,2 m für die Baumart Buche, zwischen 3,8 und 6,6 m für die Baumart Eiche, zwischen 4,2 und 4,4 m für die Baumart Fichte, zwischen 2,8 und 4,1 m für die Baumart Kiefer, zwischen 6 und 9 m für die Baumart Lärche sowie zwischen 4,3 und 7,5 m für die Baumart Tanne. Für die Baumart Douglasie variierten die Werte des Parameters A sowohl bei der Verwendung des CART-Algorithmus als auch des C4.5-Algorithmus jeweils um 2 m zwischen der besten und schlechtesten Standortgruppe.

Betrachtet man die potenziellen mittleren jährlichen Zuwächse, so findet man über alle Baumarten hinweg nur geringe Unterschiede bei der Anwendung der unterschiedlichen Algorithmen.

4.4 Validierung der Berechnungsergebnisse durch Vergleich von Höhenzuwächsen aus SILVA-Simulationen mit realen Höhenzuwächsen auf Versuchsfeldern

4.4.1 Flächenauswahl und Vorgehensweise

In Anhalt an die Arbeiten von Dursky (1999) sowie Hasenauer (1999) sollen die Berechnungsergebnisse durch Vergleich von Simulation und Wirklichkeit validiert werden. Hierzu werden die zur Einstellung des Standort-Leistungsbezuges

errechneten Parameterwerte A, k und P nach der Wachstumsfunktion von Chapman-Richards (AkP-Werte) in eine entsprechende AkP-Datei der automatisierten Version von SILVA (Seifert et. al, 1999) eingesteuert. Die ermittelten Höhenzuwächse der Simulation werden verglichen mit den auf Flächen des langfristigen ertragskundlichen Versuchswesens in Bayern ermittelten Höhenzuwächsen. Weiterhin wird neben dem Simulationslauf mit den neuen AkP-Werten eine SILVA-Simulation mit den AkP-Werten gemäß dem Standort-Leistungsmodell Kahn (Kahn, 1994) durchgeführt. Zu diesem Vergleichszweck wurden aus den Flächen des langfristigen ertragskundlichen Versuchswesens in Bayern Reinbestandsparzellen der Baumarten Fichte, Kiefer, Lärche, Eiche und Buche ausgewählt. Bei der Auswahl von Reinbestandsflächen wurde unterstellt, dass der innerartliche Konkurrenzeinfluss bereits in der Version SILVA 2.2 ausreichend gut in der aktuellen Parametrisierung abgebildet werden kann. Mischbestandsflächen wurden nicht verwendet, da bei diesen Flächen der zwischenartliche Konkurrenzeinfluss eine zusätzliche potenzielle Verzerrung des simulierten Höhenzuwachses bewirken könnte. Bei der Auswahl der Versuchsflächen wurde weiterhin darauf geachtet, dass junge, mittelalte und alte Flächen vorhanden waren. Simuliert wurde jeweils für einen kurzen Zeitraum von 15 bis 20 Jahren, um für den späteren Einsatzzweck die Übereinstimmung zwischen Simulation und Wirklichkeit in kurzen Zeiträumen zu überprüfen. Der simulierte Zeitraum hing dabei ab von den Aufnahmezeitpunkten der Versuchsflächen. Um die Ergebnisse besser vergleichen zu können, wurde innerhalb der Aufnahme- bzw. Simulationsperioden auf Einzeljahre normiert. Tabelle 25 zeigt eine Übersicht über die zu Validierungszwecken verwendeten Versuche des langfristigen ertragskundlichen Versuchswesens in Bayern.

Tab. 25: Zusammenstellung der Versuchsflächen, die für die Validierung der SILVA-Simulationsläufe genutzt wurden.

	jung	mittelalt	alt
Buche	Rothenbuch/ Ebrach 640 (P 21/ P 22) (ROT/ EBR 640)	Rothenbuch/ Ebrach 640 (P 1/ P 2) (ROT/ EBR 640)	Rothenbuch/ Ebrach 640 (P 19/ P 20) (ROT/ EBR 640)
Eiche	---	Rohrbrunn 620 (ROB 620)	Illertissen 39 (ILL 39), Lohr 59/ 60 (LOH 59/ 60)
Fichte	Zusmarshausen 603 (ZUS 603)	Weißenburg 613 (WBU 613)	Zwiesel 111 (P5) (ZWI 111)
Kiefer	Bodenwöhr 610 (P1,2) (BOD 610)	---	Schnaittenbach 58 (SNA 58)
Lärche	---	---	Mittelsinn 47/ 66 (MIS 47/ 66)

Aus Übersichtlichkeitsgründen wird in den nachfolgenden Abschnitten jeweils der Vergleich exemplarisch für einzelne Parzellen der benannten Versuche grafisch

dargestellt. Die Berechnungen wurden für alle Parzellen eines Versuches durchgeführt. Wie aus Tabelle 25 ersichtlich wird, konnten keine Vergleiche für junge Eichenreinbestände sowie mittelalte Kiefernreinbestände sowie junge und mittelalte Lärchenreinbestände durchgeführt werden. Für die Baumart Eiche findet sich zwar im langfristigen ertragskundlichen Versuchswesen in Bayern mit dem Standraum- und Durchforstungsversuch Geisenfeld 649 ein junger Reinbestandsversuch, allerdings liegt für diesen zum Zeitpunkt der Erstellung der Arbeit lediglich eine Versuchsflächenaufnahme vor, so dass eine Ermittlung von Höhenzuwächsen derzeit noch nicht möglich ist.

Einführend wird angemerkt, dass die Abbildungen 55 bis 77 den gleichen Aufbau besitzen. Die linken Boxplots (hellblau, dunkelblau) stellen die in der Realität auf den angegebenen Versuchspartellen gemessenen Höhenzuwächse dar, wobei die hellblaue Säule jeweils die ältere Aufnahmeperiode und die dunkelblaue Säule die jüngere Aufnahmeperiode wiedergibt. Da für die verschiedenen Aufnahmeperioden die Aufnahmejahre unterschiedlich waren, müssen diese der Legende entnommen werden. Es wurde über alle Versuchspartellen bzw. Baumarten versucht, zwei Aufnahmezeiträume innerhalb der letzten ca. 15 Jahre zu finden. Um die Ergebnisse vergleichbar zu machen, wurden mittlere jährliche Höhenzuwächse für die Aufnahmeperioden berechnet. Die jeweils mittleren Boxplots in Abbildung 55 bis 77 zeigen die simulierten Höhenzuwächse mit SILVA 2.2 nach der Höhenzuwachsschätzung auf Basis der Standort-Leistungsmodellierung nach Kahn (1994), wobei hell die Ergebnisse für die erste Simulationsperiode und dunkel die Ergebnisse für die jüngere Simulationsperiode dargestellt sind. Die rechten Boxplots stellen die simulierten jährlichen Höhenzuwächse auf Basis der in dieser Arbeit vorgestellten Methodik zur Einstellung des Standort-Leistungsbezuges in SILVA dar. Hellgrün gekennzeichnet ist wiederum die jeweils ältere, dunkelgrün die jüngere Simulationsperiode.

Um statistisch absicherbare Aussagen treffen zu können, wird im Anschluss an die grafische Gegenüberstellung in Abschnitt 4.4.3 der mittlere Höhenzuwachs aus Simulation und Realität mit Hilfe eines zweiseitigen Welch-Tests verglichen. Vor der Anwendung des Welch-Tests wurde mit Hilfe des Shapiro-Wilk-Tests in allen Fällen die Normalverteilungsannahme überprüft (Bortz, 2005).

Der Anwendung des Welch-Tests liegt folgende Annahme zugrunde:

$$X_i \sim N(\mu, \sigma^2), \quad Y_i \sim N(\nu, \tau^2)$$

wobei $\mu, \nu, \sigma^2, \tau^2$ unbekannt sind, $\sigma^2 \neq \tau^2$ aber zulässig ist. Geprüft wurde

$$H_0 : \mu = \nu \quad \text{gegen} \quad H_1 : \mu \neq \nu$$

Die Darstellung dieser Ergebnisse erfolgt tabellarisch in Abschnitt 4.4.3 (Tabelle 26 und 27).

4.4.2 Vergleichende Darstellung der mittleren Höhenzuwächse

a) Buche

Für die Baumart Buche wurden Daten der Wuchsreihe Rothenbuch/ Ebrach 640 zu Validierungszwecken herangezogen.

Zur Validierung der jungen Buchenbestände wurden die realen Höhenzuwächse der Versuchspartzellen 21 (Abb. 55) und 22 (Abb. 56) der Wuchsreihe Rothenbuch/ Ebrach 640 (Buchen-Durchforstungsversuch, Alter bei Versuchsanlage im Jahr 1986: 28 Jahre), herangezogen.

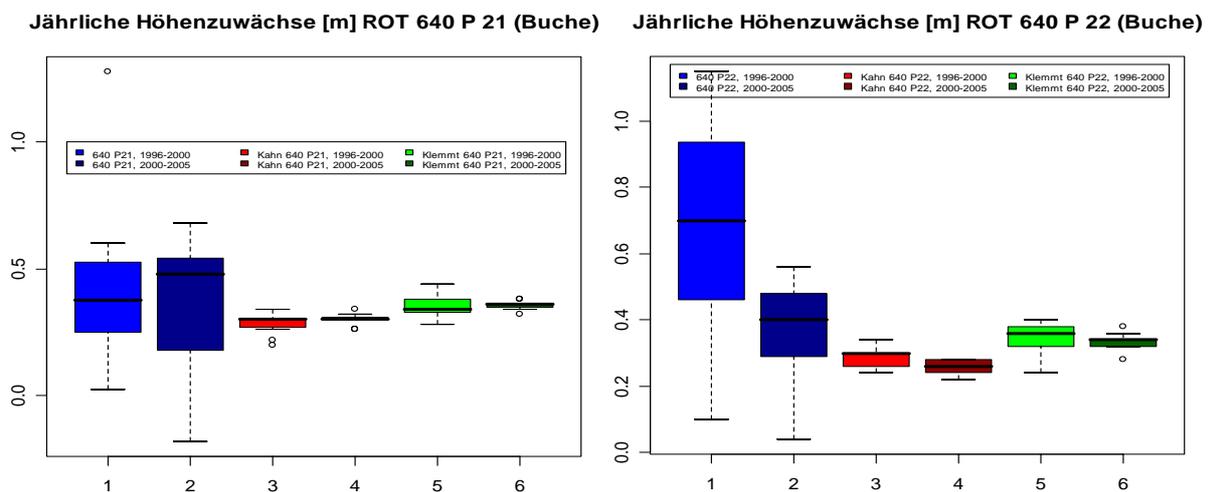


Abb. 55 (links) und 56 (rechts): Mittlere jährliche Höhenzuwächse junger Buchen auf den Parzellen 21 bzw. 22 der Versuchsfäche Rothenbuch/ Ebrach 640 im Vergleich zu den mit SILVA simulierten Höhenzuwächsen vor bzw. nach der Feinjustierung.

Aus den Abbildungen 55 und 56 wird ersichtlich, dass mit SILVA nach der modellhaften Einstellung des Standort-Leistungsbezuges nach Kahn (1994) die mittleren Höhenzuwächse unter den auf den Versuchsfächen beobachteten Höhenzuwächsen und unter den nach der in dieser Arbeit vorgestellten Methodik zur Einstellung des Standort-Leistungsbezuges simulierten Höhenzuwächsen liegen. Auffällig ist die große Spannweite der in der Realität beobachteten Höhenzuwächse im Vergleich zu den simulierten Werten.

Die Abbildungen 57 und 58 zeigen den Vergleich für mittelalte Buchen (Alter der Bäume bei Versuchsanlage im Jahr 1986: 50-65 Jahre).

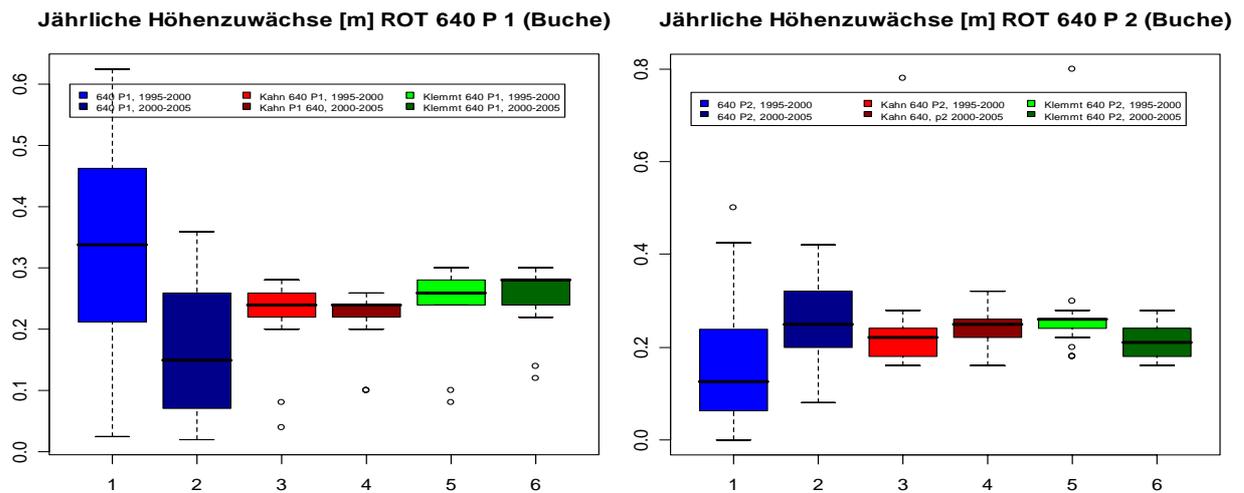


Abb. 57 (links) und 58 (rechts): Mittlere jährliche Höhenzuwächse mittelalter Buchen auf den Parzellen 1 bzw. 2 der Versuchsfläche Rothenbuch/ Ebrach 640 im Vergleich zu den mit SILVA simulierten Höhenzuwächsen vor bzw. nach der Feinjustierung.

Bei der Betrachtung der in den Abbildungen 57 und 58 dargestellten mittleren Höhenzuwächse für mittelalte Buchen fällt auf, dass beide Verfahren zur Einstellung des Standort-Leistungsbezuges dazu führen, dass die von SILVA prognostizierten Höhenzuwächse gut mit den real beobachteten Höhenzuwächsen der Versuchspartellen übereinstimmen.

In den Abbildung 59 und 60 sind die mittleren jährlichen Höhenzuwächse aus Realität und Simulation für alte Buchen (Alter der Bäume bei Versuchsanlage im Jahr 1986: 122 Jahre) dargestellt.

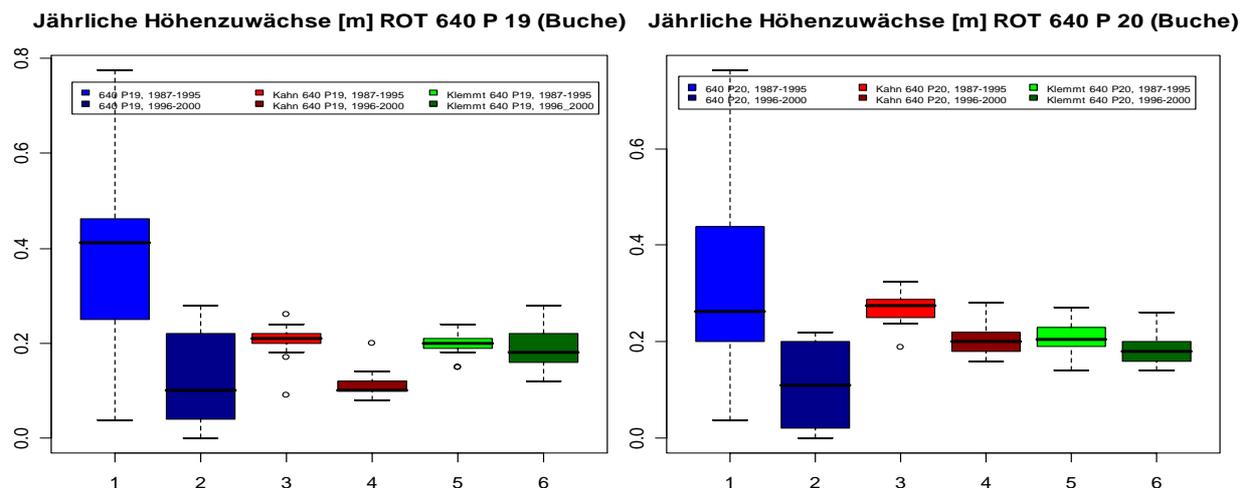


Abb. 59 (links) und 60 (rechts): Mittlere jährliche Höhenzuwächse alter Buchen auf den Parzellen 19 bzw. 20 der Versuchsfläche Rothenbuch/ Ebrach 640 im Vergleich zu den mit SILVA simulierten Höhenzuwächsen vor bzw. nach der Feinjustierung.

b) Eiche

Die Abbildungen 61 bis 66 zeigen die mittleren jährlichen Höhenzuwächse aus Realität und Simulation für mittelalte Eichen (Alter bei Versuchsanlage im Jahr 1981: 54 Jahre). Zu Vergleichszwecken wurde der Eichen-Durchforstungsversuch Rohrbrunn 620 herangezogen.

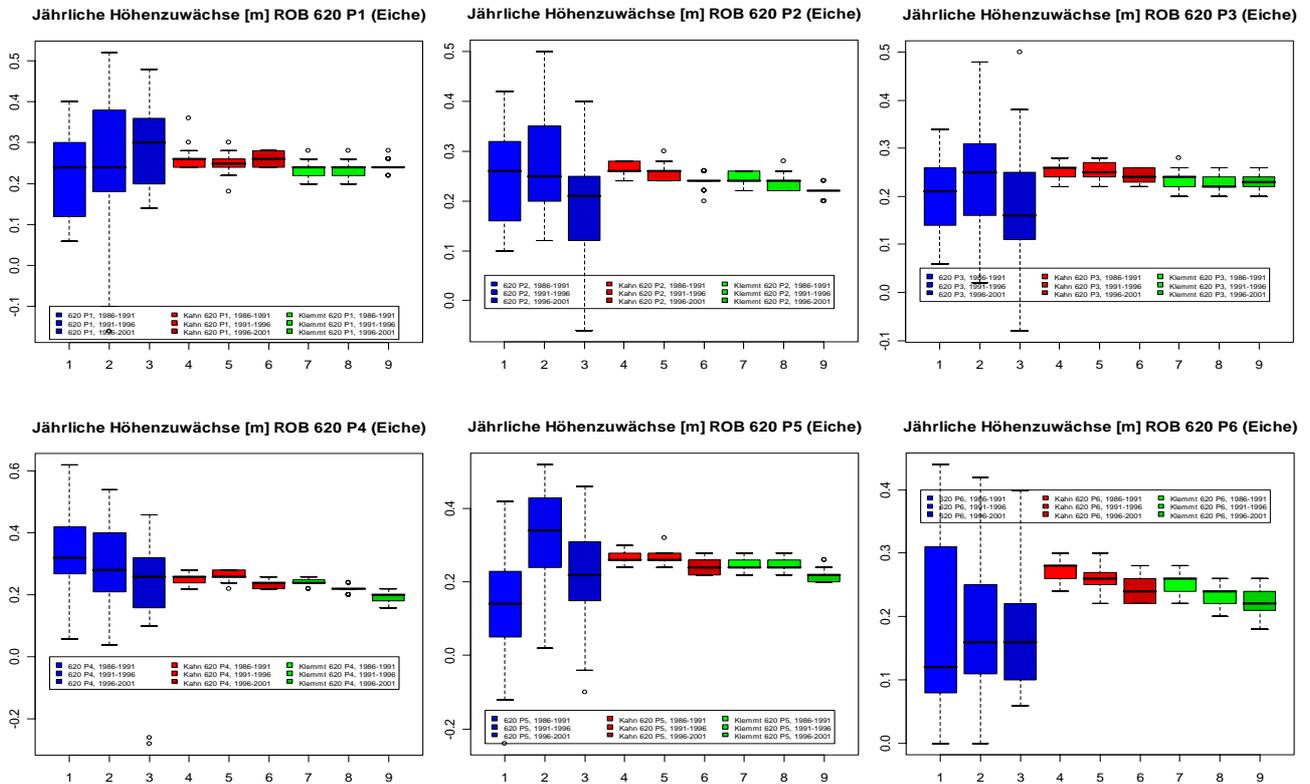


Abb. 61-66 (v. l. n. r. bzw. von oben nach unten): Mittlere jährliche Höhenzuwächse mittelalter Eichen auf den Parzellen 1 bis 6 der Versuchsfläche Rohrbrunn 620 im Vergleich zu den mit SILVA simulierten Höhenzuwächsen vor bzw. nach der Feinjustierung.

Die Abbildungen 61 bis 66 zeigen keine deutlich erkennbaren Unterschiede zwischen den von SILVA prognostizierten und den real beobachteten Höhenzuwächsen für die ausgewählten, mittelalten Eichen. Auffällig ist lediglich die vergleichsweise große Spannweite der in der Realität beobachteten Höhenzuwachsmessungen.

Alte Eichenreinbestände der Versuche Illertissen 39 (Parzelle 1) sowie Lohr 59 (Parzelle 1) und Lohr 60 (Parzelle 1) sind die Grundlage für den Vergleich von Realität und Simulationsergebnissen. Die Abbildungen 67 bis 69 zeigen die Ergebnisse der Simulationsläufe im Vergleich zu den Versuchsflächenergebnissen. Der Eichen-Durchforstungsversuch Illertissen wurde im Jahr 1885 angelegt, die aufstockenden Eichen waren damals bereits 43 Jahre alt. Der Eichen-Durchforstungsversuch Lohr 59 wurde 1901 angelegt. Das Alter der Eichen lag damals bei 67 Jahren. Die Versuchsfläche Lohr 60 wurde 1900 angelegt und ebenso

als Durchforstungsversuch konzipiert. Die Eichen auf dieser Fläche waren damals 42 Jahre alt.

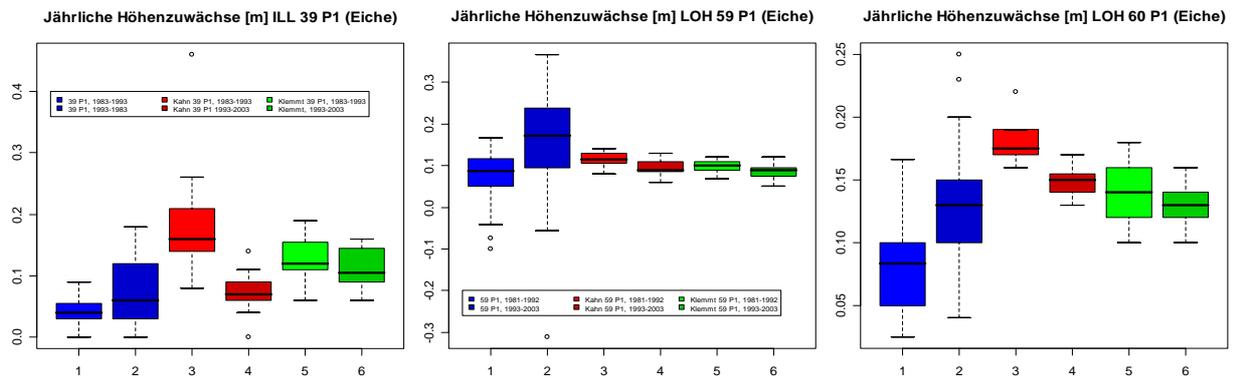


Abb. 67 – 69 (v.l.n.r.): Mittlere jährliche Höhenzuwächse alter Eichen auf der Versuchsfläche Illertissen 39, Parzelle 1 (Abb. 67), Lohr 59, Parzelle 1 (Abb. 68) bzw. Lohr 60, Parzelle 1 (Abb. 69) im Vergleich zu den mit SILVA simulierten Höhenzuwächsen vor bzw. nach der Feinjustierung.

Vergleicht man die von SILVA generierten Höhenzuwächse mit den realen Höhenzuwächsen auf den Versuchsflächen Illertissen 39 und Lohr 59 bzw. Lohr 60 so zeigt sich, dass zwischen der Höhenzuwachsschätzung auf Basis des Standort-Leistungsmodells nach Kahn und der Höhenzuwachsschätzung nach der Einsteuerung des Standort-Leistungsbezuges nach der in dieser Arbeit vorgestellten Methodik praktisch keine Unterschiede ergeben. Auch liegen die von SILVA nach beiden Methoden zur Einsteuerung des Standort-Leistungsbezuges generierten Höhenzuwächse in vergleichbaren Größenordnungen zu den real beobachteten Höhenzuwächsen.

c) Fichte

Für die Baumart Fichte wurden die Ergebnisse der Simulationsläufe mit den Höhenzuwächsen der Versuche Zusmarshausen 603, Weißenburg 613 sowie Zwiesel 111 (Parzelle 5) verglichen. Die Abbildungen 70 bis 72 zeigen die Ergebnisse exemplarisch. Der Fichten-Durchforstungsversuch Zusmarshausen 603 wurde im Jahr 1968 mit vierjährigen Fichten begründet (Abb. 70). 1975 wurde der Fichtendurchforstungsversuch Weißenburg 613 mit damals 53-jährigen Fichten angelegt (Abb. 71). Um für alte Fichten Vergleiche zu ermöglichen, wurde die Fichtenreinbestandsparzelle 5 des 1954 angelegten Bergmischwaldversuches Zwiesel 111 (Alter der Fichten bei Versuchsanlage: 70 Jahre) herangezogen (Abb. 72).

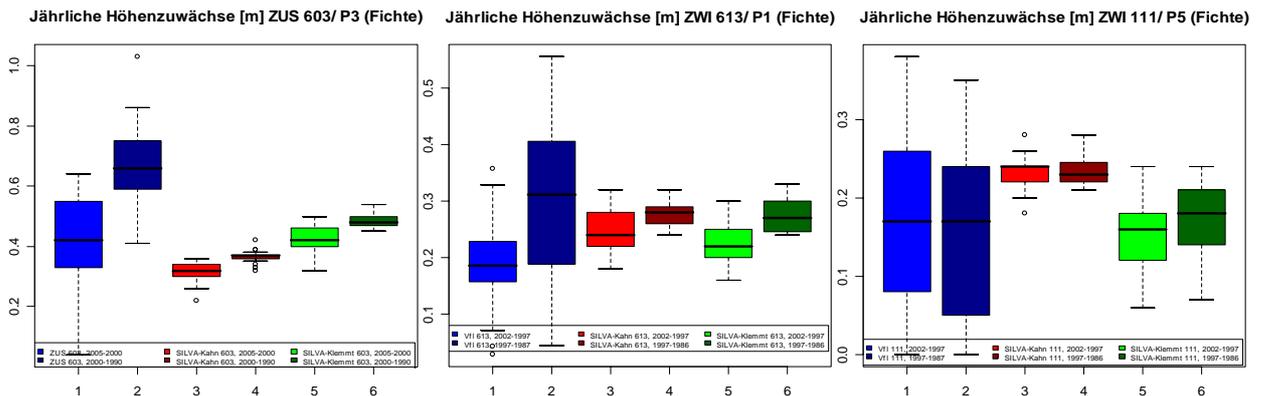


Abb. 70 – 72 (v.l.n.r.): Mittlere jährliche Höhenzuwächse junger, mittelalter und alter Fichten auf den Versuchsflächen Zusmarshausen 603 (links), Weißenburg 613 (mittig) sowie Ziesel 111 (rechts) im Vergleich zu den mit SILVA simulierten Höhenzuwächsen vor bzw. nach der regionalen standörtlichen Neujustierung.

Aus den Abbildungen 70 geht hervor, dass die von SILVA für junge Fichten nach der Einsteuerung des Standort-Leistungsbezuges gemäß Kahn (1994) generierten Höhenzuwächse im Vergleich zur Realität unterdurchschnittlich lagen. Dagegen zeigt die neue Einsteuerung des Standort-Leistungsbezuges eine bessere Annäherung an die Realität.

Mittelalte Fichten (Abb. 71) werden in Bezug auf ihren Höhenzuwachs sowohl nach der alten als auch nach der neuen Methode zur Einstellung des Standort-Leistungsbezuges in etwa gleich gut simuliert.

Für alte Fichten (Abb. 72) zeigt sich hingegen wiederum die Tendenz einer leichten Überschätzung des Höhenzuwachses nach der modellhaften Einstellung des Standort-Leistungsbezuges während die neue Einstellung zu tendenziell besseren Ergebnissen führt. In diesem Zusammenhang muss angemerkt werden, dass zu Validierungszwecken mangels weiterer Verfügbarkeit im Versuchsflächennetz des langfristigen ertragskundlichen Versuchswesens lediglich auf eine Versuchspartelle mit alten Fichten zurückgegriffen wurde.

d) Kiefer

Zur Validierung der Simulationsergebnisse wurden für die Baumart Kiefer die Versuchsflächen Bodenwöhr 610 (Parzelle 1 und 2) sowie Schnaittenbach 58 (Parzelle 1) herangezogen. Der Kiefern-Durchforstungsversuch Bodenwöhr 610 (Abb. 73 und 74) wurde 1974 mit damals 21-jährigen Kiefern begründet. Der Kiefern-Durchforstungsversuch Schnaittenbach 58 wurde im Jahr 1901 angelegt. Die aufstockenden Kiefern waren damals 35 Jahre alt.

Die Abbildungen 73 bis 75 zeigen die Höhenzuwächse auf den Versuchsflächen im Vergleich zu den Höhenzuwächsen der Simulationen.

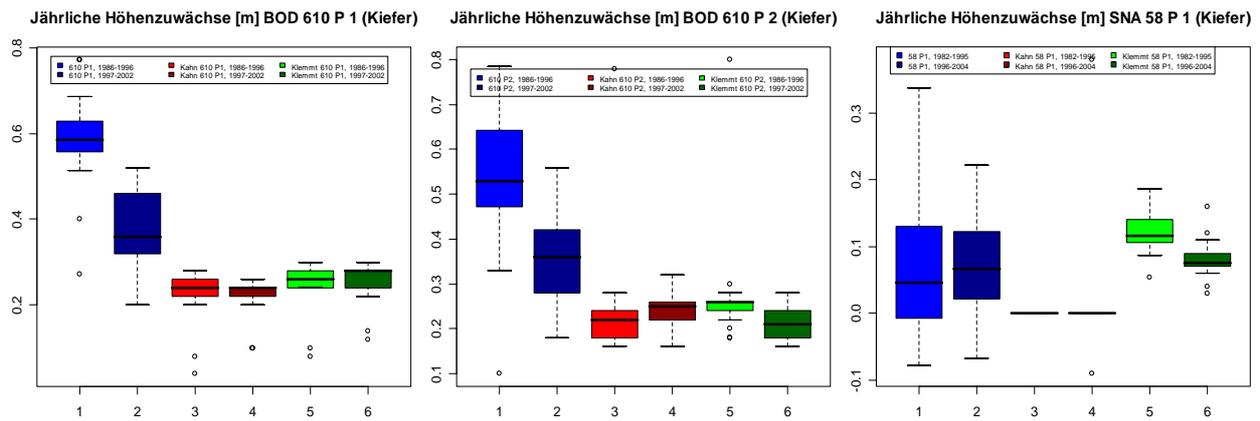


Abb. 73 - 75: Mittlere jährliche Höhenzuwächse junger und alter Kiefern auf den Versuchsfeldern Bodenwöhr 610 (links und mittig) sowie Schnaittenbach 58 (rechts) im Vergleich zu den mit SILVA simulierten Höhenzuwächsen vor bzw. nach der regionalen, standörtlichen Neujustierung.

Die Abbildungen 73 und 74 für junge Kiefern zeigen wiederum, dass die real beobachteten Höhenzuwächse deutlich über den simulierten Höhenzuwächsen lagen. Zwischen der alten und neuen Methode zur Einsteuerung des Standort-Leistungsbezuges bestehen praktisch keine Unterschiede bei den Simulationsergebnissen. Dagegen zeigt sich für alte Kiefernreinbestände eine deutliche Verbesserung der prognostizierten Höhenzuwächse beim Vergleich zwischen Realität und neuer Einstellungsmethodik (Abb. 75). Nach der Einsteuerung des Standort-Leistungsbezuges nach Kahn (1994) wurde in SILVA kein Höhenzuwachs für die Kiefern der Versuchsfeldfläche Schnaittenbach 58 (Parzelle 1) prognostiziert. Dagegen wurden von SILVA realitätsnahe Höhenzuwächse nach der in dieser Arbeit vorgestellten Methodik zur Einsteuerung des Standort-Leistungsbezuges vorhergesagt.

e) Lärche

Zur Validierung der Simulationsergebnisse wurden für die Baumart Lärche die Versuchsfeldflächen Mittelsinn 47 und 66 (jeweils Parzelle 1) herangezogen. Im Jahr 1895 wurde der Lärchen-Durchforstungsversuch Mittelsinn 47 mit damals 68-jährigen Lärchen begründet. Der Durchforstungsversuch Mittelsinn 66 wurde 1901 mit damals 85-jährigen Lärchen angelegt.

Die Abbildungen 76 und 77 zeigen die Höhenzuwächse auf den Versuchsfeldflächen im Vergleich zu den Höhenzuwächsen der Simulationen.

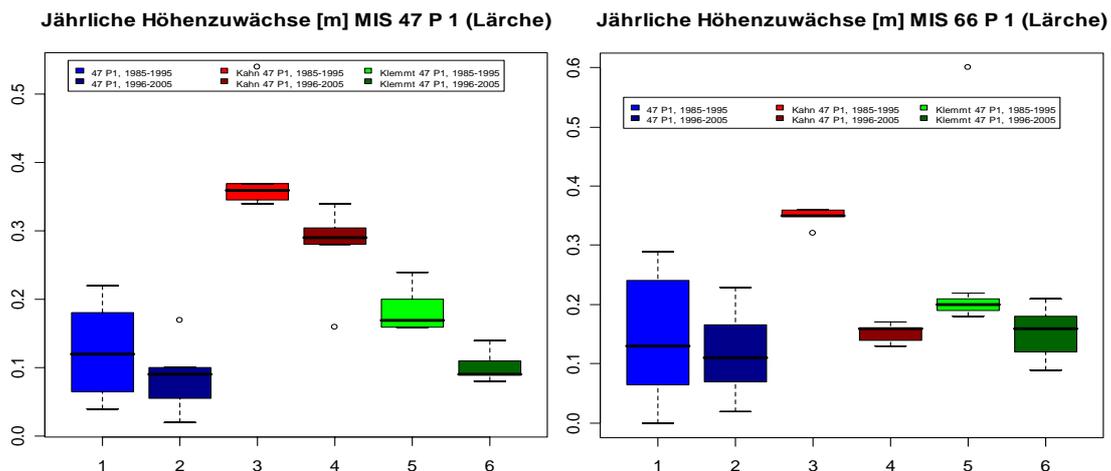


Abb. 76 - 77: Mittlere jährliche Höhenzuwächse alter Lärchen auf den Versuchsflächen Mittelsinn 47 (links) und 66 (rechts) im Vergleich zu den mit SILVA simulierten Höhenzuwächsen vor bzw. nach der regionalen standörtlichen Neujustierung.

Die Abbildungen 76 und 77 zeigen auch für die Baumart Lärche, dass SILVA nach der alten Methodik zur Einsteuerung des Standort-Leistungsbezuges die Höhenzuwächse für alte Bestände tendenziell leicht überschätzt, während die Anwendung der neuen Methodik dazu führt, dass realitätsnähere Höhenzuwächse durch SILVA geschätzt werden.

Zusammenfassend zeigen die Abbildungen 55 bis 77, dass über alle Baumarten hinweg die neue regionale, standortbezogene Feinjustierung des Standort-Leistungsbezuges in praktisch allen Fällen zu guten Übereinstimmungen zwischen tatsächlich beobachteten und simulierten Höhenzuwächsen geführt hat.

Dagegen zeigen die Simulationsläufe gemäß der Einsteuerung des Standort-Leistungsbezuges nach dem von Kahn (1994) vorgestellten Verfahren für junge Reinbestände eine leichte Unterschätzung der Höhenzuwächse. Für alte Bestände zeigt sich nach der modellhaften Einschätzung des Standort-Leistungsbezuges nach Kahn (1994) hingegen eine leichte Überschätzung der Höhenzuwächse im Vergleich zu real beobachteten Höhenzuwachsgängen. In diesem Zusammenhang muss darauf hingewiesen werden, dass die Einstellung der Standortparameter nach dem Modell Kahn für die Validierungsläufe mit SILVA gutachtlich erfolgte und z. T. mehrere Simulationsläufe erfordert hat. Bei der Anwendung der im Rahmen dieser Arbeit ermittelten AkP-Werte zur Einsteuerung des Standort-Leistungsbezuges in SILVA konnten dagegen direkt die Standortvariablen der forstlichen Praxis verwendet werden.

4.4.3 Statistische Überprüfung der mittleren Höhenzuwächse aus Realität und Simulation

Tabelle 26 enthält für die Laubbaumarten Buche und Eiche die in Abb. 55 bis 77 dargestellten Mittelwerte, die Standardabweichungen sowie die Ergebnisse der Anwendung des Welch-Tests. Für die Nadelbaumarten Fichte, Kiefer und Lärche sind die entsprechenden Werte in Tabelle 27 eingetragen.

Aus beiden Tabellen geht hervor, dass tendenziell die mittleren Höhenzuwächse, die nach der vorgestellten, neuen Methodik zur Einstellung des Standort-Leistungsbezuges eingesteuert wurden, näher an den real beobachteten mittleren Höhenzuwächsen liegen. Besonders deutlich wird diese Aussage für die Nadelbaumarten Fichte, Kiefer und Lärche.

Deutliche Unterschiede im Signifikanzniveau bei der Überprüfung zwischen Realität und Simulation konnten dabei allerdings lediglich für die Baumarten Kiefer und Lärche gefunden werden (Tab. 27).

Tab. 26: Mittlere jährliche Höhenzuwächse für die Baumarten Buche und Eiche in Realität und SILVA-Simulation sowie Ergebnisse der Anwendung des Welch-Tests.

Baumart	Altersgruppe	Name und Nummer der Versuchsfläche			Mittlerer beobachteter jährlicher Höhenzuwachs [m]	Mittlerer simulierter jährlicher Höhenzuwachs		Signifikanz-Niveau	Mittlerer simulierter jährlicher Höhenzuwachs		Signifikanz-Niveau	
		Parzelle	Aufnahmeperiode			nach Kahn (1994) [m]	p-Wert		nach Klemmt (2007) [m]	p-Wert		
Rotbuche (Fagus sylvatica L.)	jung	Rothenbuch 640	21	1996-2000	0,435	0,282	0,09659		0,351	0,37550		
			21	2001-2005	0,342	0,300	0,58340		0,356	0,84370		
		Rothenbuch 640	22	1996-2000	0,685	0,288	0,00200	**	0,349	0,00539	**	
			22	2001-2005	0,360	0,256	0,05024		0,333	0,58810		
	mittelalt	Rothenbuch 640	1	1996-2000	0,318	0,222	0,02144	*	0,254	0,08980		
			1	2001-2005	0,175	0,220	0,07796		0,250	0,03337	*	
		Rothenbuch 640	2	1996-2000	0,252	0,237	0,04554	*	0,269	0,05454		
			2	2001-2005	0,161	0,258	0,08264		0,211	0,13785		
	alt	Rothenbuch 640	19	1987-1995	0,300	0,269	0,00008	***	0,205	0,00006	***	
			19	1996-2005	0,166	0,207	0,47215		0,187	0,05043		
		Rothenbuch 640	20	1987-1995	0,378	0,204	0,37620		0,200	0,01158	*	
			20	1996-2005	0,125	0,110	0,00000	***	0,190	0,00006	***	
	Stieleiche (Quercus robur L.)	mittealt	Rohrbrunn 620	1	1991-1996	0,250	0,250	0,06409		0,234	0,55810	
				1	1997-2001	0,290	0,259	0,15020		0,243	0,03478	*
			Rohrbrunn 620	2	1991-1996	0,268	0,256	0,49790		0,238	0,62480	
				2	1997-2001	0,184	0,240	0,02572	*	0,218	0,16670	
Rohrbrunn 620			3	1991-1996	0,244	0,254	0,00244	**	0,230	0,03376	*	
			3	1997-2001	0,185	0,243	0,04085	*	0,227	0,10270		
Rohrbrunn 620			4	1991-1996	0,296	0,261	0,00080	***	0,218	0,00020	***	
			4	1997-2001	0,227	0,234	0,79510		0,197	0,25920		
Rohrbrunn 620			5	1991-1996	0,326	0,267	0,00003	***	0,246	0,00045	***	
			5	1997-2001	0,217	0,240	0,35400		0,217	0,99520		
Rohrbrunn 620			6	1991-1996	0,171	0,260	0,00153	*	0,233	0,00999	*	
			6	1997-2001	0,174	0,242	0,00215	*	0,221	0,02639	*	
alt		Illertissen 39	1	1983-1993	0,042	0,180	0,00000	***	0,129	0,00000	***	
			1	1994-2003	0,078	0,075	0,89060		0,114	0,00627	**	
		Lohr 59	1	1981-1993	0,068	0,113	0,01348	*	0,096	0,10500		
			1	1994-2002	0,148	0,095	0,12580		0,086	0,07704		
Lohr 60	1	1981-1993	0,087	0,180	0,00000	***	0,158	0,00005	***			
	1	1994-2002	0,130	0,148	0,09504		0,129	0,86100				

Tab. 27: Mittlere jährliche Höhenzuwächse für die Baumarten Fichte, Kiefer und Lärche in Realität und SILVA-Simulation sowie Ergebnisse der Anwendung des Welch-Tests.

Baumart	Altersgruppe	Name und Nummer der Versuchsfläche		Aufnahmeperiode	Mittlerer beobachteter jährlicher Höhenzuwachs [m]	Mittlerer simulierter jährlicher Höhenzuwachs nach Kahn (1994)			Mittlerer simulierter jährlicher Höhenzuwachs nach Klemmt (2007) [m]		
		Parzelle				p-Wert	Signifikanz-Niveau	p-Wert	Signifikanz-Niveau		
Fichte (Picea abies L.)	jung	Weißenburg 613	1	1986-1996	0,298	0,254	0,30910		0,256	0,26290	
			1	1997-2002	0,196	0,261	0,00128	*	0,213	0,22072	
			2	1986-1996	0,286	0,243	0,13892		0,246	0,15467	
			2	1997-2002	0,210	0,249	0,02140	*	0,229	0,34560	
			3	1986-1996	0,304	0,277	0,22270		0,275	0,20018	
	mittelalt	Zusmarshausen 603	3	1990-1999	0,676	0,367	0,00078	***	0,486	0,00399	**
			3	2000-2005	0,419	0,311	0,08264		0,421	0,87620	
	alt	Zwiesel 111	5	1986-1997	0,159	0,234	0,02443	*	0,174	0,44244	
5			1998-2005	0,178	0,233	0,00688	**	0,153	0,08234		
Kiefer (Pinus sylvestris L.)	jung	Bodenwöhr 610	1	1986-1996	0,581	0,031	0,00000	***	0,291	0,00000	***
			1	1997-2002	0,372	0,029	0,00000	***	0,284	0,00858	**
			2	1986-1996	0,536	0,084	0,00000	***	0,327	0,00000	***
			2	1997-2002	0,363	0,063	0,00000	***	0,289	0,00005	***
	alt	Schnaittenbach 58	1	1982-1994	0,064	0,016	0,00746	**	0,122	0,01583	*
			1	1995-2002	0,071	0,000	0,00050	**	0,081	0,58460	
Europäische Lärche (Larix decidua L..)	alt	Mittelsinn 47	1	1985-1994	0,124	0,381	0,00001	***	0,184	0,07316	
			1	1995-2005	0,084	0,280	0,00001	***	0,101	0,41430	
		Mittelsinn 66	1	1985-1994	0,145	0,351	0,00005	***	0,234	0,07956	
			1	1995-2005	0,121	0,152	0,15150		0,149	0,23530	

4.5 Auswertungsergebnisse der Online-Befragung zur vergleichenden Darstellung der Haltung von Forstinventurdaten in Deutschland

An der in Abschnitt 3.2.4 beschriebenen Online-Befragung haben in der Zeit zwischen dem 01.08.2005 und 26.09.2005 25 Personen bzw. Organisationen teilgenommen. Ein Datensatz war dabei fehlerhaft und unvollständig und wurde aufgrund des Wunsches des Befragungsteilnehmers, aus der Datenbank entfernt. 24 Online-Fragebögen waren vollständig und fehlerfrei ausgefüllt. Dies entspricht einer Rücklaufquote von 40 % und liegt damit im Vergleich zu anderen Online-Befragungen (vgl. z. B. Clasen und Stricker, 2004 oder Richter, 1998) im mittleren Bereich. Die 24 korrekt ausgefüllten Fragebögen bilden die Grundlage für die weiteren Auswertungen.

Aus Abbildung 78 (a) geht hervor, dass aus nahezu allen Bundesländern Personen bzw. Organisationen an der Online-Befragung teilgenommen haben. Lediglich aus Bremen, Nordrhein-Westfalen und dem Saarland hat niemand teilgenommen. Die häufigsten Teilnahmen erfolgten aus Bayern.

Abbildung 78 (b) zeigt, dass 7 Vertreter staatlicher Forstbetriebe, 11 Vertreter kommunaler Forstbetriebe und 5 Vertreter privater Forstbetriebe an der Online-Befragung teilgenommen haben.

Auf die Frage nach der Größe der von den Forstorganisationen bewirtschafteten Waldflächen variierten die Flächenangaben zwischen 825 ha und 722.000 ha. Der Mittelwert der Flächenangaben lag bei 64.651 ha. Die kleinste Flächenangabe zeigt, dass durch die Betriebsauswahl auf Basis der Internet-Recherche nicht in allen Fällen sichergestellt werden konnte, dass es sich bei den zertifizierten Flächen der Forstorganisationen um Waldflächen handelte, sondern dass durch die eingeführte Mindestflächengröße von 1.250 ha (vgl. Abschnitt 3.2.4) lediglich sichergestellt werden konnte, dass die Forstorganisation eigenes Personal, das sich vornehmlich mit der Bewirtschaftung der dem Forstbetrieb zugehörigen Flächen beschäftigt, besitzt. In der Summe ergeben die von den Befragungsteilnehmern angegebenen Waldflächen eine Fläche von 1.549.474 ha. Nach der Zweiten Bundeswaldinventur von 2001 besitzt die Bundesrepublik Deutschland eine Waldfläche von 11.075.799 ha (Bundeswaldinventur, 2005). Demnach haben an der Online-Befragung Forstorganisationen teilgenommen, die nach eigenen Angaben 13,99 % der Waldfläche in Deutschland betreuen.

Frage 7 (vgl. Tabelle 5) wollte in Form einer offenen Frage in Erfahrung bringen, seit wann in den Wäldern der Forstorganisationen Inventurdaten erhoben werden. Die Antworten auf diese Frage variierten sehr stark zwischen 1774 und 1994. Die Antworten hatten dabei zum Teil alphanumerischen Charakter. Auf weitergehende Auswertungen zu dieser Frage wurde daher verzichtet.

Abbildung 78 (c) zeigt, dass die weit überwiegende Zahl der Befragungsteilnehmer im zehnjährigen Turnus Forstinventuren durchführen.

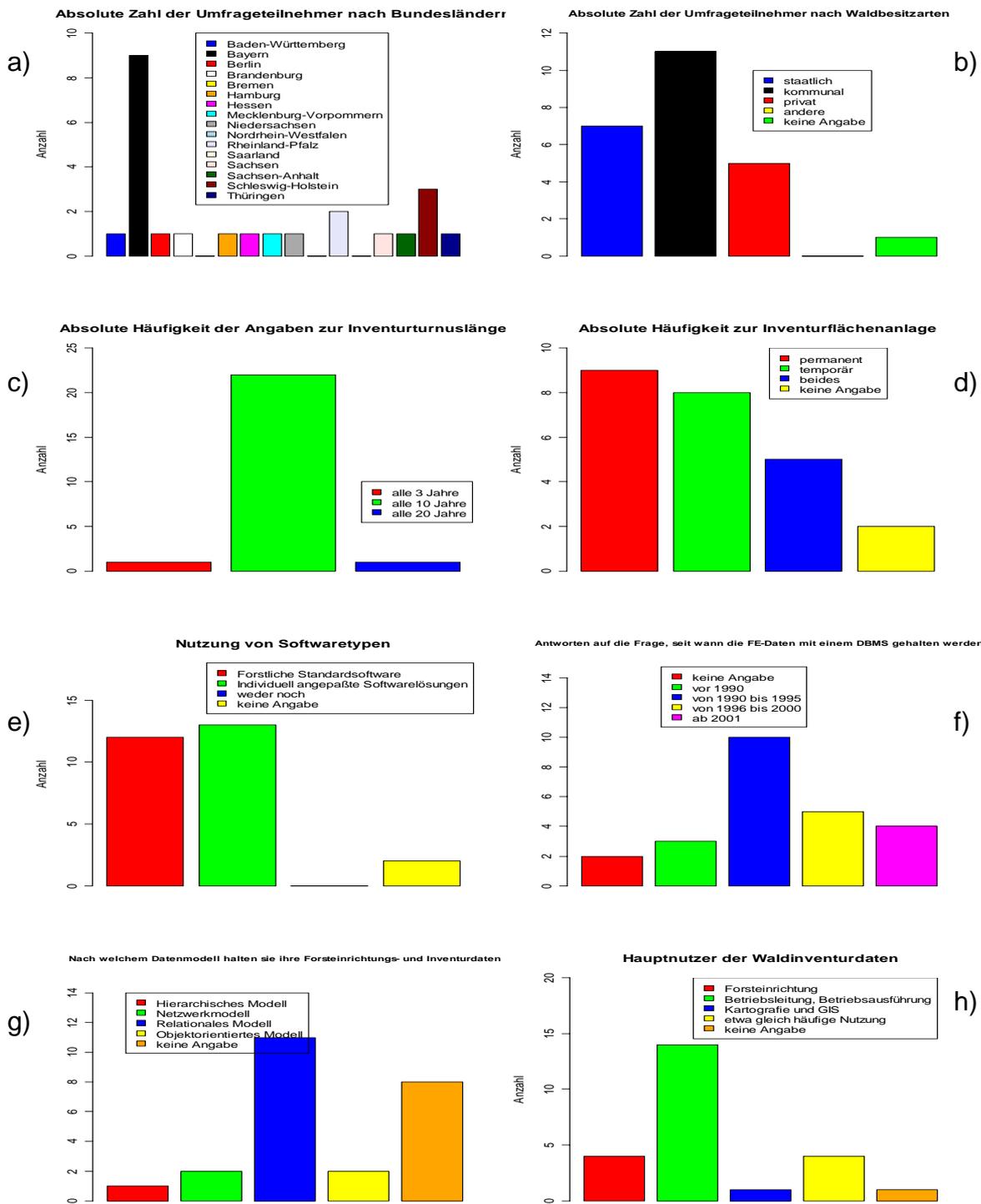


Abb. 78 (a-h): Absolute Anzahl der Antworten auf die Fragen 3 (a), 4 (b), 8 (c), 11(d), 12(e), 14(f), 17(g) und 24 (h) der durchgeführten Online-Befragung zum Vergleich der Datenhaltung von Forsteinrichtungs- und Inventurdaten in Deutschland. Die Fragen sind in Abschnitt 3.2.4 (Tabelle 5, Seite 82) aufgeführt.

Von den 24 Befragungsteilnehmern haben drei Teilnehmer auf die Frage „Welche Hilfsmittel verwenden Sie derzeit im Gelände zur Erfassung der Waldinventurdaten“ die Angabe gemacht, dass sie keine speziellen Hilfsmittel nutzen. Für sieben Teilnehmer kann sicher die Aussage getroffen werden, dass sie mobile Datenerfassungsgeräte zur Datenaufnahme im Gelände benutzen. Vier Befragungsteilnehmer gaben an, die Waldinventurdaten im Gelände auf Papier zu erfassen und später in einen Rechner zu übertragen. Zwei Befragungsteilnehmer gaben an, dass in ihren Organisationen die Inventurdaten durch externe Firmen aufgenommen werden.

Aus den Antworten auf die Frage „Welches Inventurverfahren nutzen sie?“ wurde ersichtlich, dass praktisch alle Forstorganisationen Stichprobenverfahren durchführen, die eine betriebs- bzw. landesspezifische Ausgestaltung erfahren haben. Aufgrund der freien Gestaltungsmöglichkeit ergaben sich zu dieser Frage zahlreiche Antwortkombinationen, die darauf schließen lassen, dass zahlreiche Befragungsteilnehmer das gleiche Verfahren nutzen, es allerdings unter anderem Namen in der Befragung benannt haben.

Abbildung 78 (d) zeigt die Antworten zur Erhebung der Inventurdaten auf permanenten oder temporär angelegten Probeflächen. Ersichtlich wird, dass praktisch die gleiche Anzahl an Befragungsteilnehmern entweder auf permanenten oder temporär angelegten Flächen Inventurdaten erheben. Etwa 20 % der teilnehmenden Forstorganisationen erheben sowohl auf temporären als auch auf permanenten Flächen Inventurdaten. Zwei Befragungsteilnehmer machten keine Angaben zu dieser Frage.

Frage 12 wollte von den Befragungsteilnehmern in Erfahrung bringen, welche Art von Software für die Verarbeitung der Forstinventur- und Forsteinrichtungsdaten verwendet wird. Die in Abbildung 78 (e) dargestellten Antwortmöglichkeiten waren vorgegeben, wobei Mehrfachantworten möglich waren. Abbildung 78 (e) zeigt, dass nahezu gleichhäufig „Forstliche Standardsoftware“ als auch „individuell angepasste Softwarelösungen“ bei den Befragungsteilnehmern zum Einsatz kommen. Nur ein Befragungsteilnehmer hat die Möglichkeit wahrgenommen, zwei Antwortoptionen zu wählen.

In direktem Zusammenhang mit Frage 12 stand die anschließende Frage nach dem Namen der verwendeten Software. Ersichtlich wurde, dass die Befragungsteilnehmer eine Vielzahl an unterschiedlichen Softwareprodukten verwenden. Die häufigsten Mehrfachnennungen fanden sich für die verschiedenen Produkte der Firma Forstware. Das Produkt Waldkat der Firma Greenlab wurde zweimal benannt. Acht Befragungsteilnehmer haben keine Angaben für diese Frage gemacht.

Gruppiert man die gemachten Angaben auf die Frage „Seit wann werden in ihrer Organisation mit einem Datenbankmanagementsystem Inventurdaten verwaltet?“

nach vier Zeitphasen, so ergibt sich das in Abbildung 78 (f) dargestellte Bild. Am häufigsten wurde bei den teilnehmenden Forstorganisationen zwischen 1990 und 1995 begonnen, die Inventurdaten mit einem Datenbankmanagementsystem zu verwalten. Bringt man die Antworten auf diese Frage in Verbindung mit der Waldbesitzart, die wiederum sehr eng mit der verwalteten Flächengröße zusammenhängt, so sieht man, dass alle staatlichen Forstorganisationen in diesem Zeitbereich Angaben gemacht haben.

Die Frage nach der Anzahl der aktuell gehaltenen Baumdatensätze war wiederum als offene Frage formuliert. Dies führte zu einem sehr heterogenen Antwortverhalten. Die Antworten lagen z. T. als verbale Beschreibungen, z. T. aber auch als präzise Zahlenangaben vor. Die größte angegebene Zahl lag bei 8.13 Millionen Baumdatensätzen. Auf weitergehende Auswertungen wurde für diese Frage verzichtet.

Auf die Frage „Wie erfolgt die Datenhaltung“ haben 84 % bzw. 20 Befragungsteilnehmer mit der Antwort „zentral geantwortet. Jeweils zwei Befragungsteilnehmer haben die Antwortmöglichkeit „dezentral“ bzw. „keine Angabe“ gewählt.

Auf Frage 16 „Bitte benennen sie die Architektur des Datenbankmanagementsystems“ haben drei Befragungsteilnehmer mit „Fileserver-Lösung“ und 12 mit „Client-Server-Lösung“ geantwortet. Neun Befragungsteilnehmer machten von der Möglichkeit Gebrauch, keine Angaben zu dieser Frage zu machen.

Abbildung 78 (g) zeigt die Antworten zu Frage 17 nach dem verwendeten Datenmodell. Die Mehrzahl der Befragungsteilnehmer hält die Forstinventurdaten nach dem relationalen Modell. Auffällig ist die hohe Zahl der Antworten bei der Antwortmöglichkeit „keine Angabe“.

Auf die Frage nach dem Hersteller des verwendeten Datenbankmanagementsystems antworteten vier Teilnehmer mit „Oracle“, sieben mit „Microsoft“, zwei mit „MySQL“ sowie fünf mit „andere“. Bei der Unterfrage zur Antwortoption „andere“ entfielen zwei Antworten auf „PostgreSQL“.

Frage 19 wollte in Erfahrung bringen, wie viele Benutzer insgesamt Zugriff auf die Forsteinrichtungs- und Inventurdaten haben. Die Antworten variierten zwischen 1 und 800, bei der Hälfte der Befragungsteilnehmer haben weniger als elf Personen Zugriff auf diese Daten, bei drei Forstorganisationen hingegen können mehr als 100 Personen auf die Forstinventurdaten zugreifen.

In Ergänzung zur vorhergehenden Frage ist Frage 20 („Wie viele Benutzer haben im Durchschnitt gleichzeitig Zugriff auf die Forsteinrichtungsdaten?“) zu sehen. 16 Teilnehmer machten Angaben mit weniger als drei Benutzern, sechs Teilnehmer machten keine Angabe auf diese Frage.

Auf die Frage „Wie viele Personen sind mit der Administration des Datenbankmanagementsystems beschäftigt (Vollzeit)?“ variierten die erhaltenen

Antworten zwischen null und vier Personen. Fünf Befragungsteilnehmer gaben an, dass eine Person in Vollzeit mit der Administration des Datenbankmanagementsystems beschäftigt ist. Aus 8 Antworten geht hervor, dass es sich um eine Teilzeitbeschäftigung handelt.

Auf Frage 22 „Welche Techniken nutzen sie zur Auswertung der Daten?“ wählten 17 Teilnehmer die Antwortoption „Klassische Abfragesprachen“, sieben Teilnehmer machten keine Angaben.

Die anschließende Frage nach den verwendeten Datenbanksprachen führte in 15 Fällen zur Antwort „SQL“, neun Teilnehmer gaben keine Antwort auf diese Frage.

Abb. 78 (h) zeigt die Antworten auf die Frage „Wer nutzt in ihrer Organisation die Waldinventurdaten am häufigsten“. 14 Nennungen entfielen bei der Beantwortung dieser Frage auf „Betriebsleitung und –ausführung“, vier Nennungen entfielen jeweils auf „Forsteinrichtung“, eine auf „Kartografie und GIS“. Vier Nennungen entfielen auf „etwa gleichhäufige Nutzung“, ein Teilnehmer machte keine Angaben.

Auf Frage 25 „Werden die erhobenen Daten standardmäßig mit einem GIS-System weiterverarbeitet haben elf Teilnehmer mit „ja“ und neun Teilnehmer mit „nein“ geantwortet.

An alle Teilnehmer, die auf Frage 25 mit „ja“ geantwortet haben, wurden drei weitere Fragen gestellt.

In der ersten Zusatzfrage wurde nach den verwendeten GIS-Produkten gefragt. Acht Antworten entfielen hierbei auf Produkte der Firma ESRI, zwei Antworten entfielen auf PIA-GIS sowie eine Antwort auf FOGIS. Die zweite Zusatzfrage wollte in Erfahrung bringen, ob in dem gleichen Datenbanksystem die Verwaltung sowohl der Inventur- als auch der Geodaten möglich wäre. Sechs Befragungsteilnehmer antworteten mit „ja“, drei mit „nein“, zwei machten keine Angaben. Die sechs Befragungsteilnehmer, die mit „ja“ geantwortet haben, sind identisch mit denen, die auf Frage 18 mit „Oracle“ oder „PostgreSQL“ geantwortet haben. Die letzte Zusatzfrage war wie folgt formuliert „Stellen sie ihren Benutzern derzeit die Waldinventurdaten in Form einer web-basierten GIS-Lösung (Web-GIS) zur Verfügung?“. Zwei Antworten auf diese Frage entfielen dabei auf „ja“, sieben auf „nein“, zwei Teilnehmer machten keine Angaben.

Nachfolgend wird versucht, die Ergebnisse der Online-Befragung zu interpretieren. Zu Beginn wird allerdings noch einmal darauf hingewiesen, dass aufgrund der Auswahl der potenziellen Befragungsteilnehmer in Verbindung mit der freiwilligen Teilnahmemöglichkeit die Repräsentativität der Aussagen nur sehr bedingt gegeben ist und eine Übertragbarkeit auf Einzelbetriebe oder größere Flächeneinheiten nur eingeschränkt möglich ist.

Die Online-Befragung hatte das Ziel, einen Überblick über die Datenhaltung der Forstinventurdaten in Deutschland als Grundlage für eine eventuelle Übertragbarkeit einer Softwarelösung zur regionalen, standortbezogenen Feinjustierung des Wachstumsmodells SILVA auf Basis von Forstinventurdaten, zu gewinnen. Dieser Zusammenhang wurde den potenziellen Befragungsteilnehmern im Vorfeld der Durchführung der Online-Befragung mitgeteilt. Die Teilnahme von 40 % der angeschriebenen Forstbetriebe wird daher als Anzeichen dafür gedeutet, dass bei den Forstorganisationen ein großes Interesse an dieser Thematik besteht. Die Teilnahme aus nahezu allen Bundesländern zeigt, dass diese Thematik eine weit über Bayern hinausgehende Bedeutung hat. Eventuell ist die Teilnahme aus ganz Deutschland auf die Bekanntheit des Wachstumsmodells SILVA aufgrund zahlreicher Publikationen und Vortragsveranstaltungen zurückzuführen. Die Teilnahme von Forstorganisationen aller Waldbesitzarten zeigt, dass dieses Interesse besitzartübergreifend vorhanden ist. Betrachtet man die Zahlen zur Waldflächengröße in Verbindung mit amtlichen Vergleichszahlen z. B. der Bundeswaldinventur (Bayerische Landesanstalt für Wald und Forstwirtschaft, 2005) oder des Statistischen Bundesamtes (Statistisches Bundesamt, 2006) so stellt man fest, dass tendenziell größere Forstorganisationen an der Befragung teilgenommen haben. In Verbindung mit den Ergebnissen zu den verwendeten Hilfsmitteln bei der Aufnahme der Waldinventurdaten im Gelände wird dies als Folge der effizienteren Organisation und der längeren Tradition der EDV-basierten Datenhaltung in größeren Forstbetrieben gesehen. Die kleineren Forstbetriebe unter den Befragungsteilnehmern scheinen erst seit kürzerem die elektronische Datenerfassung sowie die effiziente Datenhaltung mit Hilfe elektronischer Datenverarbeitungstechniken eingeführt zu haben. Die Betriebe mit größeren Betriebsflächen scheinen dagegen schon eine längere Erfahrung im Zusammenhang mit datenverarbeitungsgestützter Aufnahme und Haltung von Waldinventurdaten zu besitzen.

Im Hinblick auf die Inventurverfahren scheinen bei den einzelnen Forstorganisationen, die an der Umfrage teilgenommen haben, zahlreiche Gemeinsamkeiten zu bestehen, die lediglich aufgrund regional oder lokal anderer ökologischer, ökonomischer oder sozialer Rahmenbedingungen eine Abwandlung erfahren. Alle Befragungsteilnehmer gaben an, Stichprobenverfahren auf Bestandes- oder Probepunktebene zu nutzen. Dies könnte eine Grundlage für z. B. die Entwicklung von Empfehlungen für den einheitlichen Aufbau und das Design von Forsteinrichtungsdatenquellen in Deutschland darstellen.

Die Antworten auf die Frage nach der bisher verwendeten Software zeigen, dass für den Bereich der Forstinventur entweder „Forstliche Standardsoftware“ oder „Individuell angepasste Softwarelösungen“ verwendet wurden. Bei den

Befragungsteilnehmern hat sich diesbezüglich kein klarer Trend „herauskristallisiert“. Bezüglich der Datenhaltung von Forstinventurdaten sowie der Datenverarbeitungs-umgebung fiel bei den Antworten der Befragungsteilnehmer auf, dass die Anzahl der Antwortmöglichkeit „keine Angabe“ deutlich zunahm bzw. die Qualität der Antworten insbesondere bei offenen Fragen stark voneinander abwich. Dies steht in engem Zusammenhang mit den Antworten zur Zahl der mit der Haltung der Forstinventurdaten betrauten Personen. Insbesondere bei Befragungsteilnehmern mit wenigen Beschäftigten zur Betreuung der Datenbankumgebungen erschienen die Antworten teilweise unpräziser bzw. häufte sich der Anteil von „keine Angabe“-Antworten. Dies könnte eine Folge davon sein, dass in den teilnehmenden Forstorganisationen mit wenigen EDV-Spezialisten für diesen Aufgabenbereich, die Antworten durch Personen gegeben wurden, denen spezielle datenverarbeitungstechnische Hintergründe nur bedingt vorliegen. Z. T. wiesen diese Organisationen aber auch darauf hin, dass bei ihnen die Datenerfassung, Datenhaltung und z. T. die (Vor-) Verarbeitung der Forstinventurdaten in externe Unternehmen ausgelagert wurde.

Bei den Antworten zur Datennutzung fiel auf, dass in den meisten teilnehmenden Organisationen Betriebsleitung und Betriebsausführung die Waldinventurdaten nutzen während nur ein deutlich geringerer Teil der Forstorganisationen angab, dass die Waldinventurdaten vornehmlich von den vorverarbeitenden Stabsstellen der Forsteinrichtung bzw. Kartographie genutzt werden. Bei letzterer Aussage gilt es allerdings im Einzelfall stets den organisatorischen Aufbau der teilnehmenden Forstorganisation zu betrachten.

Zu den Fragen zur standardmäßigen Weiterverarbeitung und Vorverdichtung der Waldinventurdaten mit Hilfe geographischer Informationssysteme fiel auf, dass dieser Arbeitsschritt noch keinen Standard bei den teilnehmenden Forstbetrieben darstellt. Für die größeren Forstorganisationen auffällig ist, dass sie diesen Schritt durch die Schaffung einer geodatenfähigen Infrastruktur derzeit aufbauen. Die netzbasierte Präsentation der Ergebnisse von Waldinventuren und Forsteinrichtungen sowie der Benutzerdialog über Netzwerkumgebungen stellt hingegen derzeit bei den Befragungsteilnehmern noch den Ausnahmefall dar.

5. Erstellung einer Softwarelösung zur Feinjustierung von SILVA

5.1 Grundlegende Konzeption

Auf der Grundlage der in den Abschnitt 3 vorgestellten Methodik und der in Abschnitt 4 vorgestellten Ergebnisse wird im Rahmen dieser Arbeit eine Softwarelösung erstellt, die die für Simulationen und Prognosen mit SILVA benötigten Parameter einer standortbezogenen Wuchsleistungseinschätzung der Baumarten semiautomatisch aus Forstinventurdaten ableitet. Der semiautomatische Aspekt bezieht sich dabei auf die in gewonnene Erkenntnis, dass die Ergebnisse der Berechnungsläufe - insbesondere der Generierung der Entscheidungsbäume - der Einschätzung und Evaluierung durch Experten bedürfen. Die Ergebnisse der erstellten Softwarelösung sollen dem Benutzer vor Übernahme in SILVA zur Überprüfung vorgestellt werden. Der Benutzer kann diese akzeptieren, in Werten und Kategorien ändern oder verwerfen.

Bei der Ergebnispräsentation soll dabei darauf geachtet werden, dass diese dem Benutzer in möglichst einfacher Art und Weise zur Verfügung gestellt werden. Die Ergebnisse sollen dabei nicht nur in reiner Zahlenform dargestellt werden sondern bereits grafisch aufbereitet zur Verfügung stehen. Um dem Benutzer erste Auswirkungen der Berechnungen aufzeigen zu können, wird neben der zahlenmäßigen und grafischen Ergebnisdarstellung noch eine einfache GIS-Komponente in das Programmsystem integriert. Diese soll insbesondere den praxisorientierten Anwender in die Lage versetzen, die Auswirkungen der Klassifikation kartenmäßig in bereits bekannter Form zu betrachten. Hierdurch erhält die Anwendung neben der Feinjustierungsfunktion einen eigenen Gebrauchswert.

Als Programmiersprache für die Entwicklung der Anwendung kommt JAVA der Firma Sun Microsystems zum Einsatz. Maßgebend für die Wahl von JAVA sind folgende Eigenschaften der Programmiersprache (nach Jobst (2002), Flanagan (2004) sowie nach Balzert (2005)):

- Der Java-Compiler generiert sogenannten JAVA-Bytecode. Dieser Bytecode stellt eine symbolische Sprache dar, die unabhängig von der verwendeten Hardware ist. Zur Ausführung dieses Codes wird eine spezielle Umgebung herangezogen, die als JAVA-Laufzeitumgebung (Runtime-Environment) oder JAVA-Plattform bezeichnet wird. Hierdurch wird sichergestellt, dass die zu erstellende Anwendung auf allen gängigen Rechner- und Betriebssystemen lauffähig wird.

- JAVA gehört zu den objektorientierten Sprachen. Die Grundidee der objektorientierten Programmierung ist die softwaretechnische Abbildung in einer Art und Weise, wie der Mensch auch Dinge der realen Welt erfährt. Dies erfolgt mit dem Ziel, die Softwarequalität zu erhöhen und Fehler zu minimieren. JAVA ist allerdings im Gegensatz zum Beispiel zu Smalltalk oder Eiffel nicht vollständig objektorientiert. Die Grunddatentypen beispielsweise sind keine Objekte. Weiterhin unterstützt JAVA keine direkte Mehrfachvererbung. Letztere Einschränkungen sind allerdings für die Erstellung der Anwendung dieser Arbeit irrelevant.
- JAVA bietet darüber hinaus die Möglichkeit Klassen zu erstellen, die in unterschiedlichen Ausführungsumgebungen ablaufen. So lassen sich beispielsweise Applets in Webbrowsern, die JAVA unterstützen, ausführen. Das Sicherheitskonzept von JAVA (Code-Verifier, Class-Loader sowie diverse Security-Manager) wird dabei genutzt, um unbekannte Klassen keinen Schaden anrichten zu lassen, was insbesondere bei Applets wichtig ist.
- JAVA-Programme lassen sich modular aufbauen. Der Quellcode kann dabei auf verschiedene Dateien aufgeteilt werden, wobei JAVA keine sog. Header-Files kennt. Die Konsistenz der Module wird in JAVA direkt gegen den tatsächlichen Quellcode geprüft.
- JAVA besitzt eine hochgradige Netzwerkfähigkeit. Mit JAVA ist beispielsweise der Zugriff auf Funktionen der Browser sowie die Programmierung von Sockets möglich. Hierdurch ermöglicht JAVA den Aufbau von verteilten Anwendungen, was insbesondere für den Aufbau von Unternehmensanwendungen an Bedeutung gewinnt (Haas und Schreiner, 2002).

In die Softwareumgebung von SILVA - insbesondere in der Version 3 (vgl. Seifert et. al., 2005) - passt sich die JAVA-Applikation dieser Arbeit deswegen gut ein, da bereits im aktuellen SILVA 3-Entwicklungsprojekt die Programmiersprache JAVA für die Erstellung des Software-Clients zum Einsatz kommt.

Mit Abschluss der Entwicklungsarbeiten vorliegender Arbeit liegt eine selbständig lauffähige Softwarelösung vor, die weitgehend unabhängig von der verwendeten Hard- und Softwareumgebung des potenziellen Nutzers ist. Dies soll eine einfache Anwendbarkeit bei einer großen Zahl potenzieller Nutzer ermöglichen.

Mit der erstellten Softwarelösung soll es z. B. möglich sein, für SILVA in der Version 2.2 die standortbezogenen Leistungsparameter für die einzelnen Baumarten zu ermitteln. Gleichzeitig erhält der Benutzer in diesem Fall eine Softwarelösung, die das Wissen um die standörtlichen Potenziale der Baumarten im eigenen Forstbetrieb erweitert.

Für die weitere Verwendung in SILVA 3 liegen lauffähige Klassen vor, die sich ohne großen Aufwand in die Client-Umgebung von SILVA 3 integrieren lassen.

Die erstellte Softwarelösung wird der Lesser General Public License (LGPL) unterstellt. Die LGPL setzt ähnlich wie die GPL die Idee des Copyleft ein, um die Freiheit der lizenzierten Software zu schützen. Dabei sind die Freiheiten, die Software für einen beliebigen Zweck zu nutzen, die Software zu vervielfältigen und weitergeben zu dürfen, die Software nach eigenen Bedürfnissen zu ändern und die geänderten Versionen weiterzugeben, gemeint. Veränderte Versionen müssen allerdings ebenfalls unter der LGPL lizenziert werden. Grundsätzlich darf eine unter LGPL lizenzierte Software nur mit Quelltext vertrieben werden oder mit der Zusage, den Quelltext auf Anfrage nachzureichen. Im Gegensatz zur GPL dürfen alle Programme, die LGPL-lizenzierte Software nur extern benutzen ihre eigene Lizenz behalten. Damit eignet sich die LGPL besonders als Lizenz für Bibliotheken, deren Benutzung man auch den Programmierern proprietärer Programme erlauben will. Durch die Nutzung dieses Lizenzierungsmodells wird es für potenzielle Nutzer möglich, die in den Klassen des Programms implementierte Funktionalität innerhalb eigener Anwendungen zu nutzen. Die Klassen des Programms können demnach z. B. in den SILVA 3-Client integriert werden bzw. an eigene Oberflächen angepasst werden, ohne bei einer Weitergabe den vollständigen SILVA 3-Quellcode weitergeben zu müssen.

Aufgrund des gewählten Lizenzierungskonzeptes wird bei der Entwicklung der Softwarelösung konsequent darauf geachtet, dass die verwendeten Klassen bzw. Bibliotheken den Lizenzierungsrichtlinien der GPL bzw. LGPL entsprechen und kostenlos zur Verfügung stehen.

Im Rahmen dieser Arbeit wird ein Prototyp entwickelt. In der Softwareentwicklung ist der Begriff Prototyp nicht eindeutig definiert. Gemäß einer Definition aus der Technik ist ein Prototyp eine ausführbare und evaluierbare Version eines Systems oder Systemteils oder eines Werkzeuges, um ausgewählte Merkmale des späteren, operationellen Systems zu evaluieren. Die nachfolgenden Beschreibungen des Prototyps basieren dabei auf dem Entwicklungsstand zum 01.12.2006 und beschreiben die grundlegende Funktionalität der dieser Arbeit beigelegten Softwarelösung. Die aktuelle Version der Software kann am Lehrstuhl für Waldwachstumskunde oder beim Autor dieser Arbeit nachgefragt werden.

Die im Rahmen dieser Arbeit erstellte Softwarelösung zur regionalen standortbezogenen Feinjustierung des Wachstumsmodells SILVA mit Hilfe von Forstinventurdaten erhält die Bezeichnung „Silva-PEM“ bezeichnet. PEM steht dabei als Abkürzung für „Parameter Estimation Module“.

5.2 Technische Konstruktion

Wie in Abschnitt 5.1 bereits ausführlich erläutert, wird Silva-PEM mit Hilfe der Programmiersprache JAVA entwickelt. Als Entwicklungsumgebung wird Eclipse (<http://www.eclipse.org>) in der Version 3.0 verwendet.

Jede JAVA-Applikation enthält einmal die Methode `main()` als zentralen Einsprungspunkt in die Anwendung (Balzert, 2005). Im Modul `SilvaPEM.java` wird diese Methode in der Klasse `SilvaPEM` implementiert. `SilvaPEM` stellt demnach den zentralen Einsprungspunkt der Anwendung dar, der für den Programmstart (Methode `main(java.lang.String[] args)`), die grundlegende Bildschirmdarstellung (Konstruktor `SilvaPEM` bzw. Methode `addChild(javax.swing.JInternalFrame child, int x, int y)`) sowie für die Benutzerinteraktion (Methode `actionPerformed(java.awt.event.ActionEvent event)`) verantwortlich ist. Innerhalb des Konstruktors wird das grundlegende Erscheinungsbild des Fensters sowie das Aussehen der Menüleiste festgelegt. Die Methode `actionPerformed` entspricht der Implementierung eines Event-Listeners und reagiert auf Tastatur- oder Mauseingaben innerhalb des Hauptfensters der Anwendung.

Im Menü Datei können grundlegende Datenoperationen wie das Einlesen von Daten oder das Speichern oder Drucken von Ergebnissen durchgeführt werden. Daten können aus zwei Datenquellen eingelesen werden. Zum einen ist ein direkter Datenbankzugriff implementiert, der über eine JDBC-ODBC-Bridge das Einlesen der Daten bewerkstelligt und die entsprechenden Datenbankinformationen in programminterne Objekte überführt. Die entsprechende Funktionalität ist im Modul `JDBC_ODBCDataReader.java` der Anwendung implementiert. Zum anderen ist die Anwendung Silva-PEM in der Lage, XML-Dateien über den Aufruf eines entsprechenden XML-Parsers abzuarbeiten. Aufgrund der Empfehlung von Hein und Zeller (2003) wird in diesem Fall die SAX-API (Modul `CallSAX.java`) zur Abarbeitung der linearisierten Daten verwendet. Innerhalb des letztgenannten Moduls wird über den Aufruf eines Factory-Objektes ein `SAXParser`- bzw. ein `XML-Reader`-Objekt erzeugt. Abschließend werden die Daten über entsprechende `ContentHandler` eingelesen.

Für jeden Baumdatensatz wird dabei ein eigenes Objekt generiert, welches im Modul `TreeDataSet.java` definiert ist. Die Verwaltung aller zusammengehörigen Baumdatensätze erfolgt mit Hilfe einer Umsetzung des sog. Iterator-Patterns (Jobst, 2002), welches in JAVA in der Aufzählungsschnittstelle „Enumeration“ implementiert ist. Die programminterne Verwaltung der Bäume erfolgt über diese Schnittstelle innerhalb eines Objektes der Klasse `TreeList`, die wiederum innerhalb des Moduls `ObjectManagement.java` implementiert ist.

Zur Neuberechnung der Werte der Parameter A, k und P (vgl. Abschnitt 3.2.3) werden die Module `NonlinearRegression.java`, `AKPValues.java` und `MeanAKPValuesVisualiser.java` benutzt. Das Modul `NonlinearRegression.java` nimmt die Liste der Baumobjekte auf, wandelt sie entsprechend um und übergibt sie der freien Statistiksoftware R (R Development Core Team, 2006) zusammen mit den notwendigen R-Befehlen. Der Datenaustausch erfolgt dabei mit Hilfe der JRI-Schnittstelle, die standardmäßiger Bestandteil des R-packages „rJAVA“ ist. Exemplarisch ist der Quellcode des Moduls „`NonlinearRegression.java`“ in Anhang 9.5 abgedruckt. Das Berechnungsergebnis von R wird von Silva-PEM wieder übernommen und in einem Objekt der Klasse `AKPValues` abgelegt. Das Modul `MeanAKPValuesVisualiser.java` übernimmt dieses Ergebnisobjekt und steuert die Anzeige sowie evtl. Änderungsmöglichkeiten für den Benutzer.

Bei Akzeptanz durch den Benutzer stellen die errechneten Koeffizienten der Parameter A, k und P die Grundlage für die Klassifikation der Daten dar (vgl. hierzu Abb. 17 und 18, Seite 75). Hierzu wird den Instanzen der Klasse `TreeDataSet` die Klassenzugehörigkeit in der Elementvariable Klasse zugewiesen. Dies erfolgt über ein Objekt der Klasse `Classification`. Diese Klasse ist im Modul `Classification.java` implementiert. Im vorliegenden Prototyp werden die Eingangsdaten umgewandelt in ein ARFF-File, welches entweder im Verzeichnis der Ausgangsdaten (Datengrundlage: XML-Datei) oder im Installationsverzeichnis von Silva-PEM (Datengrundlage: Datenbank) abgelegt wird. Dieser Zwischenschritt ist notwendig, da dem Klassifikationsalgorithmus in WEKA das Skalenniveau der einzelnen Variablen explizit übergeben werden muss.

Der angewendete C4.5-Algorithmus zur Erstellung von Entscheidungsbäumen entstammt dem WEKA-Framework (<http://www.cs.waikato.ac.nz/ml/weka/>). Bei diesem handelt es sich um ein Data-Mining-Framework, das unter der Leitung von Witten und Frank an der University of Waikato (Neuseeland) entwickelt wurde. Es enthält zahlreiche Data-Mining-Algorithmen bzw. Algorithmen aus dem Bereich des maschinellen Lernens. Dieses Framework liegt mit seinen Algorithmen im JAVA-Quellcode vor und darf lizenzgemäß weiterverwendet werden. Bei Witten und Frank (2005) sind die einzelnen Algorithmen und Techniken erläutert.

Durch den Aufruf der Klasse `J48`, die sich im Package `weka.classifier.trees` der Anwendung Silva-PEM befindet, wird die Ausführung des C4.5-Algorithmus angestoßen. Die weiteren Startwerte werden aus dem Modul `Preferences.java` (Mindestanzahl der Fälle pro Blatt etc.) entnommen und dem Objekt der Klasse `J48` übergeben. Nach erfolgreicher Ausführung des Algorithmus liegt ein Entscheidungsbaum in der Notation wie in den Abschnitten 4.2.1 bis 4.2.7 für die Ausführung des Algorithmus C4.5 dargestellt, vor. Mit Hilfe der Klasse

ParseDecisionTree im gleichnamigen Modul wird dieser Baum in ein XML-File überführt und steht zur Weiterbearbeitung mit anderen Klassen bereit. Dieser Schritt stellt im Programmfluss auch die Verbindung zur Anzeige von vorberechneten Berechnungsergebnissen dar.

Über Instanzen der Klasse DecisionTreeReader werden vorberechnete Entscheidungsbäume in das Programmsystem eingelesen. Für die Verteilung bzw. ergänzende Berechnung von darzustellenden Größen sind die Objekte der Klasse DecisionTreeAnalyzer verantwortlich. Die Darstellung der Entscheidungsbäume erfolgt mit Hilfe eines Objektes der Klasse DecisionTreeVisualizer. In diese Klasse implementiert ist eine Methode mit der Bezeichnung MouseEvent, welche auf Mausereignisse reagiert. Beim „Klicken“ auf ein Blatt- oder Knotenelement des dargestellten Entscheidungsbaumes wird eine Instanz der Klasse AKPTable instanziiert, die in einem nichtmodalen Fenster die Ergebnisse einer Neu- oder Vorberechnung anzeigt (vgl. hierzu Abb. 84). Nach Ansicht aller Blätter bzw. Knotenelemente wird, je nach gewählter Voreinstellung im Menüpunkt Ergebnisse eine Instanz der Klasse WriteResultsToDatabase bzw. WriteResultsToFile instanziiert, die für die Ausgabe der AkP-Werte und der generierten Regeln in den entsprechenden Ausgabezweig verantwortlich ist.

Durch die Wahl des Menüpunktes „GIS“ im Menü „Ansicht“ wird ein Objekt der Klasse ShowMapResult generiert, welches die entsprechenden Klassen der GeoTools-Library (<http://geotools.codehaus.org/>) nutzt, um die Ergebnisse des Berechnungslaufes kartografisch einfach darzustellen. Im vorliegenden Prototyp der Anwendung ist allerdings lediglich die Anzeige von im Vorfeld der Berechnung erstellten Karten (Menüzweig „Öffnen“) möglich. Dadurch, dass die GeoTools-Bibliothek die Bibliothek JTS (Java Topology Suite) nutzt, sollte eine Erweiterung zur dynamischen Erstellung von thematischen Karten auf Basis neuer Berechnungsergebnisse möglich sein.

5.3 Programmtechnische Dokumentation

Sämtliche Klassen der Anwendung Silva-PEM sind mit Hilfe von JavaDOC dokumentiert. JavaDOC ist ein von Sun Microsystems entwickeltes, kostenfrei zugängliches Software-Dokumentationswerkzeug, das automatisch HTML-basierte Dokumentationsdateien auf der Grundlage spezieller Quelltextkommentare erstellt. Die vollständige HTML-Dokumentation befindet sich für alle Klassen der Anwendung Silva-PEM auf beiliegendem Datenträger. Aufgrund des z. T. erheblichen Umfangs

der Dokumentation wird darauf verzichtet, diese schriftlich der Anlage zu dieser Arbeit beizufügen.

5.4 Benutzerdokumentation

Der schriftlichen Version dieser Arbeit liegt eine Version der Softwarelösung Silva-PEM zur Ansicht weiterer Ergebnisse bei. Diese Version entspricht dem Entwicklungsstand zum 01.12.2006 und ermöglicht die Anzeige von vorberechneten Fällen. Nachfolgend wird anhand von Bildschirmabzügen ein Programmlauf für den Benutzer erläutert. Als Beispiel wurde wiederum die Berechnung für die Baumart Fichte im Bayerischen Wuchsbezirk 4.2 („Südliche Fränkische Platte“) herangezogen (vgl. Abschnitt 3.2.3). Der dazugehörige Testdatensatz befindet sich auf beiliegendem Datenträger im Verzeichnis „/Testdatensatz/“.

Abbildung 79 zeigt das Startfenster der Anwendung Silva-PEM. Das interne Fenster wird standardmäßig 15 Sekunden angezeigt.



Abb. 79:
Startbildschirm der
Anwendung Silva-
PEM.

Nachdem der Startbildschirm nicht mehr im Hauptfenster der Anwendung angezeigt wird, kann der Benutzer aus den Menüpunkten „Datei“, „Ansicht“, „Ergebnisse“, „Voreinstellungen“ sowie „Hilfe“ auswählen.

Der Menüpunkt „Datei“ dient dazu, grundlegende Datenoperationen wie das Laden der Daten oder das Speichern oder Drucken von Ergebnissen durchzuführen.

Im Menüpunkt „Ansicht“ kann aus den Ansichtsoptionen „DataExplorer“ (Standard), „Entscheidungsbaum“ oder „GIS“ ausgewählt werden.

Der Menüpunkt „Ergebnisse“ dient der Steuerungsmöglichkeit der Ausgaben. Die Ergebnisse können entweder für SILVA 2.2 in Dateiform oder für SILVA 3 in Form von Datenbanktabellen ausgegeben werden. Mit Hilfe des Menüpunktes „Voreinstellungen“ können grundlegende Einstellungen zum Erscheinungsbild des Programms (Anmutung) sowie zu den voreingestellten Start- und Steuergrößen (im vorliegenden Prototyp noch nicht vollständig implementiert) vorgenommen werden. Der Menüpunkt Hilfe (im vorliegenden Prototyp noch nicht vollständig implementiert) vermittelt Informationen zum Programm auf unterschiedlichen Informationsebenen.

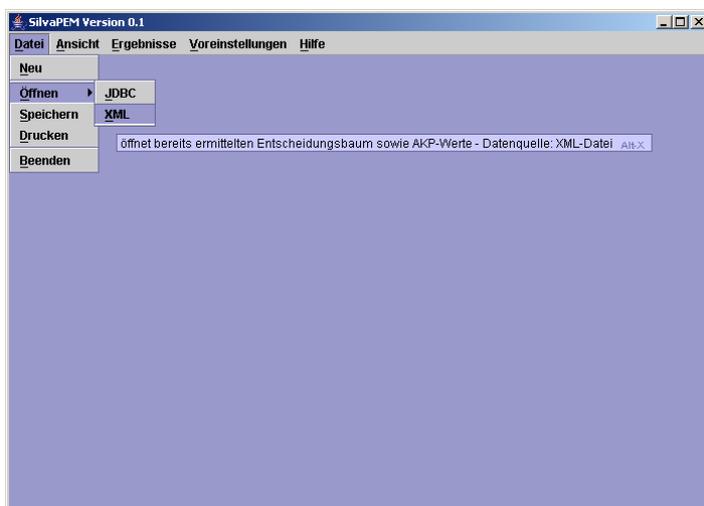


Abb. 80: Hauptfenster der Anwendung Silva-PEM bei geöffnetem Menü Datei-Öffnen-XML.

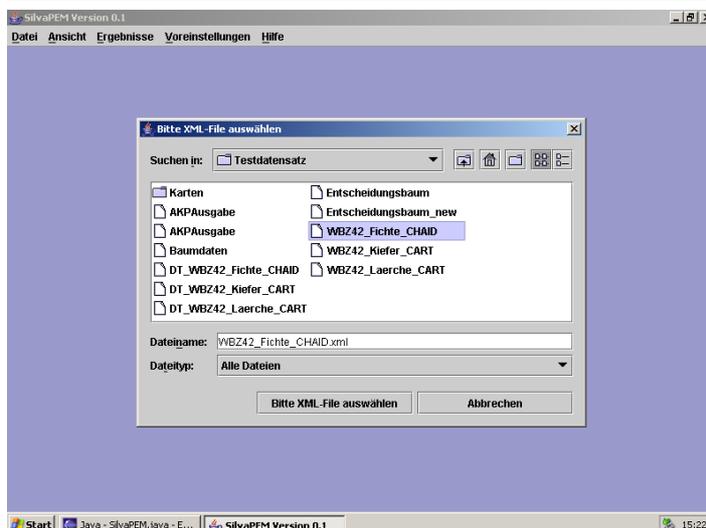


Abb. 81: Dateiauswahldialog im Hauptfenster der Anwendung Silva-PEM.

Abbildung 80 zeigt das Hauptfenster bei geöffnetem Menüpunkt Datei. Zu erkennen ist die Anzeige des sog. ToolTips für den Menüpunkt „Datei-Öffnen-XML“. Bei ToolTips handelt es sich um nichtmodale Anzeigen, die dem (ungeübten) Benutzer

die Reaktionen des Programms bei der Auswahl z. B. des Menüpunktes anzeigen. Hiermit soll für den (ungeübten) Benutzer der Umgang mit der Anwendung erleichtert werden. ToolTips wurden daher für alle Menüpunkte des Programms Silva-PEM implementiert.

Abbildung 81 zeigt den Dateiauswahldialog, der nach Anwahl des Menüpunktes „Datei-Öffnen-XML“ erscheint. Der Benutzer wird aufgefordert, ein XML-File auszuwählen. Silva-PEM akzeptiert dabei bei der Auswahl sowohl die Wahl einer Entscheidungsbaumdatei (DT_xxx.xml) als auch einer Datendatei (WBZxx_xxx.xml).

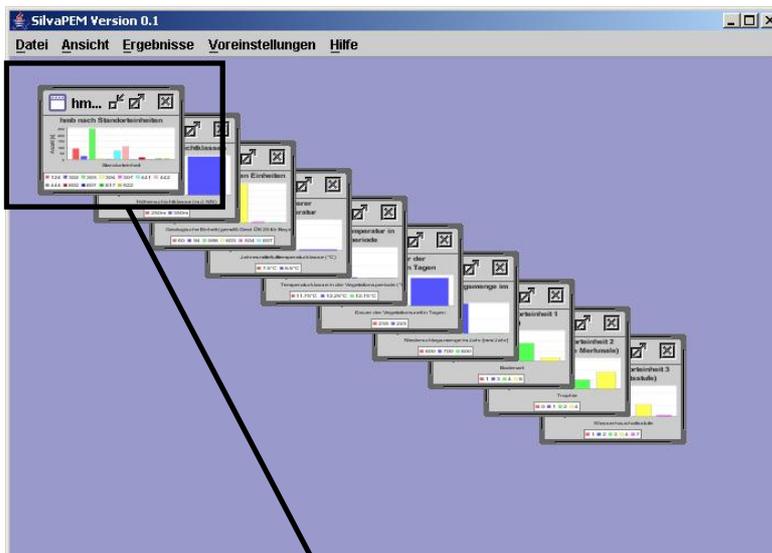


Abb. 82:
Bildschirmabzug von
Silva-PEM in der
sogenannten
„DataExplorer“-
Ansicht.

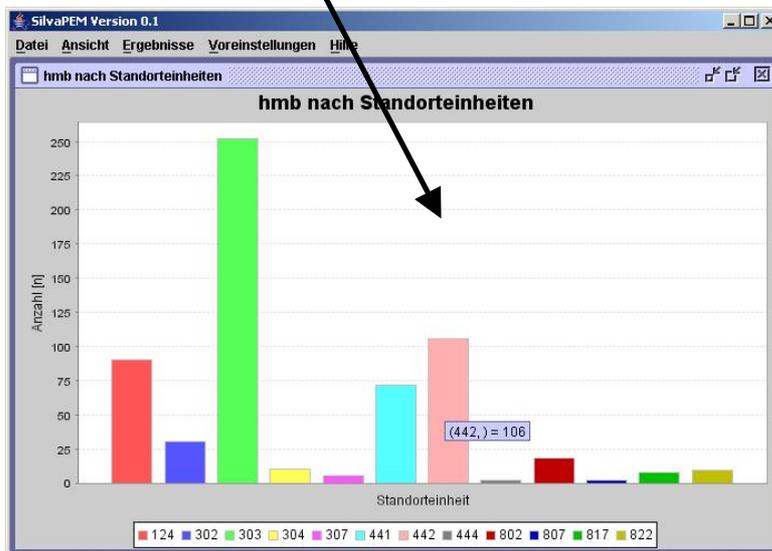


Abb. 83:
Bildschirmabzug von
Silva-PEM in der
sogenannten
„DataExplorer“-
Ansicht. Das
nichtmodale
Kindfenster zur
Anzeige der absoluten
Zahl der Höhenmess-
bäume nach Standort-
einheiten wurde auf
die Größe des Arbeits-
bereiches vergrößert.

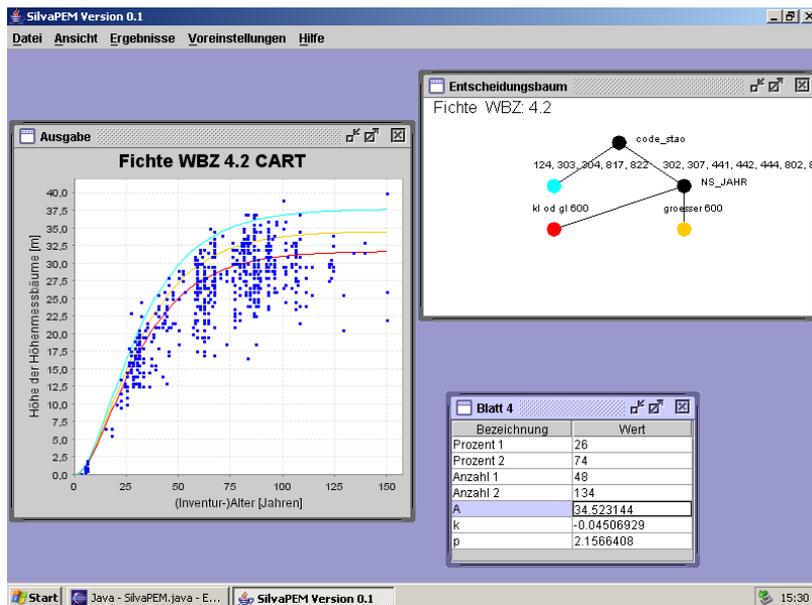


Abb. 84: Bildschirmabzug von Silva-PEM bei gewählter Ansichtoption „Entscheidungsbaum“.

Aufgrund der standardmäßigen Vorsteinstellung der Option „DataExplorer“ im Menü „Ansicht“ wird dem Benutzer nach erfolgter Dateiauswahl ein Bildschirm gemäß Abbildung 82 bzw. vergrößert gemäß Abbildung 83 gezeigt. Dieses Vorgehen zwingt den Benutzer, sich mit dem Datenmaterial für die Berechnungen zu beschäftigen. Er muss jedes Fenster aktiv zumindest schließen, um mit Silva-PEM weiterarbeiten zu können.

Nachdem der Benutzer die Möglichkeit erhalten hat, das Datenmaterial anzusehen, erfolgt die programminterne Bearbeitung der Daten. Das Ergebnis (Neuberechnung oder Darstellung einer bereits im Vorfeld des Programmlaufes durchgeführten Berechnung) wird mit Hilfe eines eigenen nichtmodalen Fensters mit dem Entscheidungsbaum bzw. einem Fenster für die einzelnen Datenpunkte dargestellt. Abbildung 84 zeigt dieses Ergebnis für den Berechnungslauf gemäß den Abbildungen 17 und 18 (Seite 75) für die Baumart Fichte im Wuchsbezirk 4.2.

Im Kindfenster „Entscheidungsbaum“ (Abb. 84, oben rechts) wird der Entscheidungsbaum dargestellt. Die Blätter bzw. die dazugehörigen Datensätze erhalten hier eine eigene Farbe zugewiesen, die wiederum im Ausgabefenster (Abb. 84, linkes Kindfenster) oder in der GIS-Darstellung (Abb. 85) verwendet wird. Die Knotenelemente des Entscheidungsbaumes bleiben schwarz gefärbt.

Bei Auswahl eines Blattes bzw. eines Knotens mit Hilfe eines Mausklicks öffnet sich ein weiteres Kindfenster, welches die dazugehörigen Werte für den generierten Entscheidungsbaum anzeigt (vgl. Abb. 84, Kindfenster unten rechts). In dieses Fenster werden die relativen und absoluten Häufigkeiten der Zuordnung zu Klasse 1

oder 2 sowie die Werte der Parameter A, k und P der Wachstumsfunktion nach Chapman-Richards eingetragen. Wie für dieses Kindfenster zu sehen ist, können diese Werte hier vom Benutzer aufgrund von lokalem Expertenwissen editiert werden. Die in diesen Fenstern eingetragenen Werte können direkt für Prognoseläufe im Wuchsmodell SILVA verwendet werden.

Um dem Benutzer die Auswirkungen der Klassifikation praxisnah für betriebliche Entscheidungsprozesse zu zeigen, wurde zudem eine GIS-Ansicht in Silva-PEM implementiert (Abb. 85). In den Farben der Entscheidungsbaumklassifikation werden hier mit Hilfe eines nichtmodalen Kindfensters die Klassifikationsergebnisse kartenmäßig dargestellt.

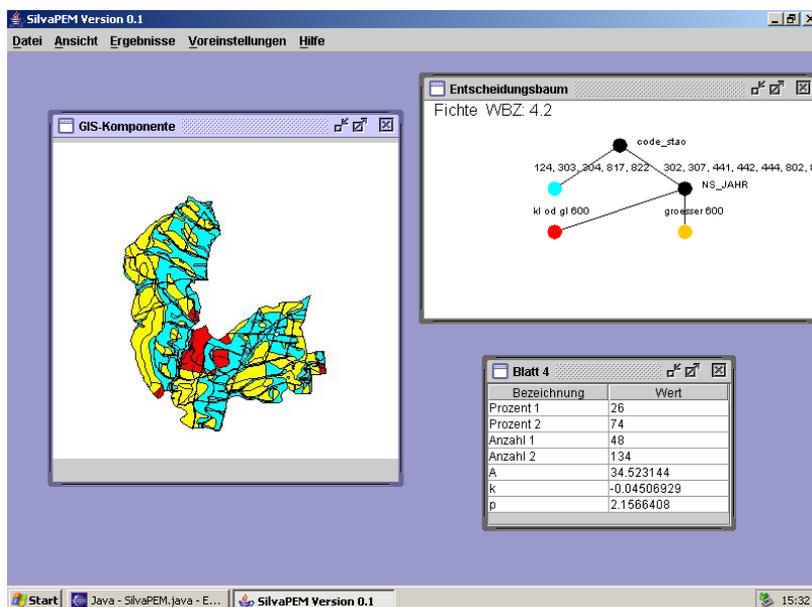


Abb. 85: Bildschirmabzug von Silva-PEM in der GIS-Ansicht.

In Abbildung 85 zu sehen ist die Verschneidung eines Ausschnittes einer Standortkarte eines im Wuchsbezirk 4.2 liegenden Staatswaldkomplexes mit der entsprechenden Niederschlagskarte des Klimaatlas für Bayern (BayFORKLIM, 1996). Wie bereits im vorhergehenden Abschnitt angeführt wurde, ist im vorliegenden Prototyp lediglich die Darstellung einer vorbereiteten Karte mit Hilfe der freien Bibliothek GeoTools möglich. Da diese JAVA-Bibliothek die JavaTopologySuite (JTS) (<http://www.vividsolutions.com/jts/jtshome.htm>) nutzt, wäre eine Erweiterung zur automatischen Verschneidung und Darstellung derartiger Karten oder anderer thematischer Karten innerhalb von Silva-PEM technisch realisierbar.

Aufgrund rechtlicher Bedenken wurde die dieser Arbeit beiliegende Programmversion dahingehend abgeändert, dass keine Geodaten in Form von Shapefiles angezeigt werden können sondern lediglich die im Vorfeld erstellten thematischen Karten auszugsweise in Form von JPEG-Files angezeigt werden

können. Dadurch wird es dem Benutzer exemplarisch ermöglicht, sich die Ergebnisse in Form thematischer Karten anzeigen zu lassen, ohne dass die für die Erstellung der Arbeit zur Verfügung gestellten Geodaten weitergegeben werden müssen.

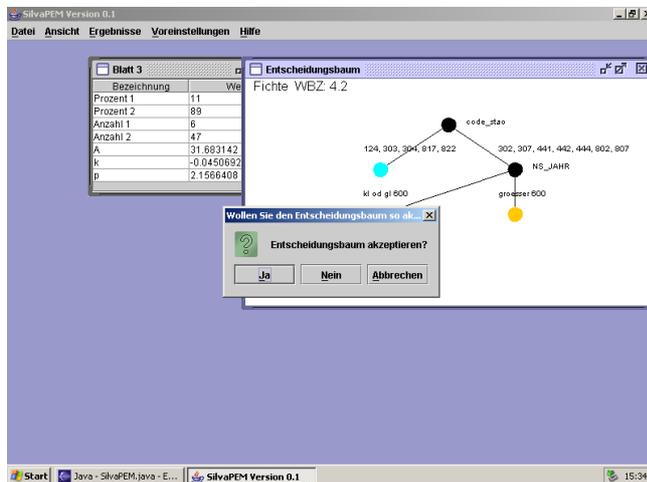


Abb. 86: Bildschirmabzug von SilvaPEM mit Anzeige des nicht modalen Dialogfensters zur Abfrage, ob der Entscheidungsbaum akzeptiert werden soll oder nicht.

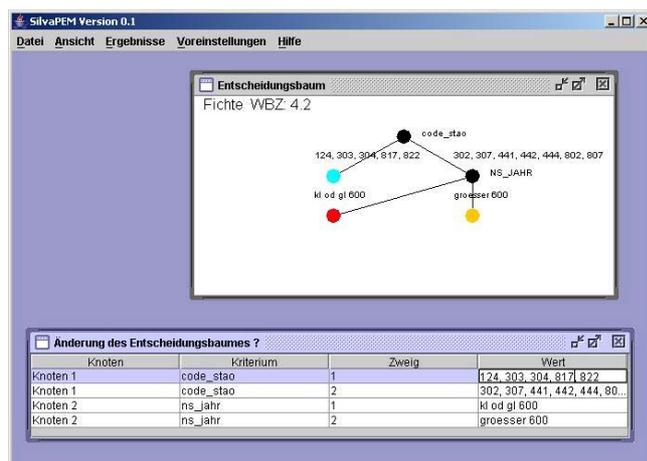


Abb. 87: Bildschirmabzug von SilvaPEM mit Anzeige der Editiermöglichkeit, wenn der Entscheidungsbaum geändert werden soll (Auswahl „Nein“ nach Abb. 86).

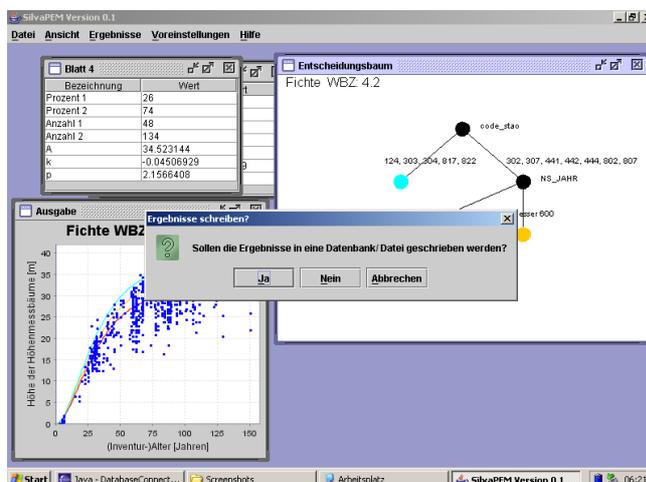
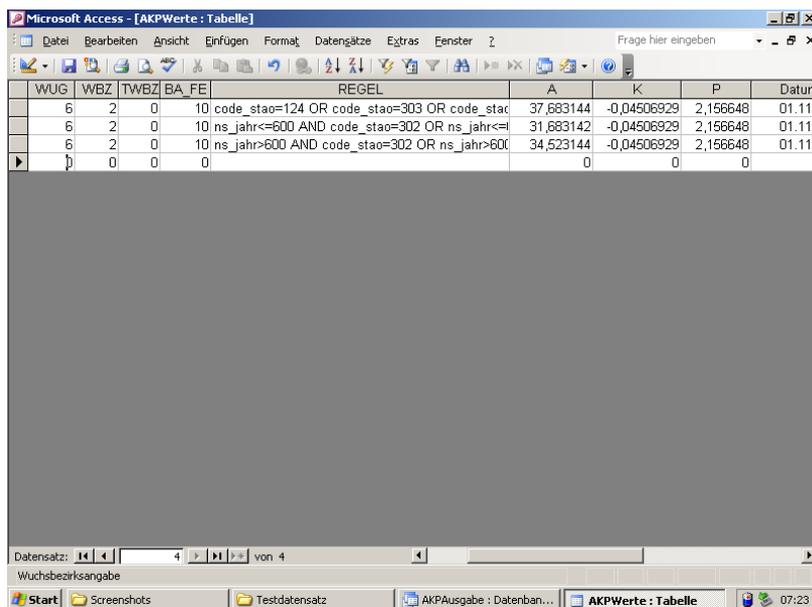


Abb. 88: Bildschirmabzug von SilvaPEM mit Anzeige des modalen Dialogfensters zur Abfrage, ob die angezeigten Ergebnisse in eine Datenbank oder Datei geschrieben werden sollen.

Wurden alle Knoten und Blätter des Entscheidungsbaumes durch den Benutzer angesehen, so erscheint ein Dialogfenster gemäß Abbildung 86, mit welchem er entscheiden kann, ob er den Entscheidungsbaum wie angezeigt akzeptieren möchte oder nicht. Im konkreten Fall nach Abb. 84 würde der Benutzer vermutlich die Standorteinheiten 444 und 817 innerhalb des Hauptknotenelementes aufgrund seines Erfahrungswissens austauschen. Bei der Auswahl der Option „Nein“ im entsprechenden Dialogfeld nach Abbildung 86 erscheint ein zusätzliches modales, editierbares Fenster. Dies ist in Abbildung 87 dargestellt. In diesem hat der Benutzer die Möglichkeit, die Werte innerhalb der einzelnen Knotenelemente auszutauschen. Den Vorgang der Änderung von Werten innerhalb eines Entscheidungsbaumes schließt der Benutzer durch Schließen des Fensters und erneute Auswahl aller Blatt- und Knotenelemente im Entscheidungsbaumfenster ab.

Bei Akzeptanz des Entscheidungsbaumes bzw. nach abgeschlossener Änderung des Entscheidungsbaumes gelangt der Benutzer zu einem Fensterbild gemäß Abbildung 87. An dieser Stelle wird gefragt, ob die Daten in eine Datenbank zur Weiterverwendung mit SILVA 3 oder in eine Datei zur Weiterverwendung mit SILVA 2.2 geschrieben werden sollen.

Bei Auswahl der Optionen „Ja“ im Dialogfenster gemäß Abb. 88 werden bei Wahl der Ergebnisoption „Ergebnisse für SILVA 3 schreiben“ die Ergebnisse über die bereits erläuterte JDBC-ODBC-Bridge in eine relationale Datenbank geschrieben. Abb. 89 zeigt exemplarisch die Ergebnisausgabe eingetragen in eine Microsoft-Access-Datenbank.



WUG	WBZ	TWBZ	BA	FE	REGEL	A	K	P	Datum
6	2	0	10	code_stao=124 OR code_stao=303 OR code_stao=302 OR ns_jahr<=600 AND code_stao=302 OR ns_jahr<=600	37,683144	-0,04506929	2,156648	01.11.11	
6	2	0	10	ns_jahr>600 AND code_stao=302 OR ns_jahr>600	31,683142	-0,04506929	2,156648	01.11.11	
6	2	0	10	ns_jahr>600 AND code_stao=302 OR ns_jahr>600	34,523144	-0,04506929	2,156648	01.11.11	

Abb. 89: Bildschirmabzug der Ergebnisausgabe in eine MS-Access-Datenbank.

In Abbildung 89 zu erkennen ist die Tabelle AKP-Werte, in die für die gewählte Baumart (BA_FE) der Wuchsregion (WUG, WBZ, TWBZ) die maschinell generierten Regeln sowie die ermittelten Werte der Parameter A , k und P der Wachstumsfunktion nach Chapman-Richards eingetragen sind. Des Weiteren ist am rechten Bildschirmrand ein Datumseintrag zu erkennen, der von Silva-PEM zum Zeitpunkt des Eintrags in die Datenbank generiert wird.

6. Diskussion

6.1 Zugrundeliegende Daten

Vorliegende Arbeit nutzt Forstinventurdaten, um auf Basis der Alters-Höhenentwicklung der tatsächlich höhengemessenen Bäume einen regionalen, standortbezogenen Leistungsbezug als Grundlage für eine Feinjustierung des Wachstumsmodells SILVA abzuleiten.

Wie in Tabelle 1 gegenübergestellt, unterscheiden sich Forstinventurdaten sowohl in qualitativer als auch in quantitativer Hinsicht stark von Daten, die auf langfristigen, ertragskundlichen Versuchsflächen aufgenommen werden. Ein wesentlicher Unterschied besteht in der Zeitreiheneigenschaft, wie nachfolgend kurz dargestellt wird.

Beim Inventurverfahren im Staatswald Bayerns handelt es sich im Regelfall um die Aufnahme konzentrischer Probekreise mit gestaffelten Kluppschwellen. Die Anlage der Probekreise erfolgt temporär oder permanent nach den Vorgaben der Inventurverantwortlichen (vgl. z. B. Forstdirektion Oberfranken-Mittelfranken, 2001). Dies ist maßgebend für einen Teil der Eigenschaften der Inventurdaten. Werden Stichprobenpunkte temporär angelegt, so wird bei einer Wiederholungsinventur so gut wie nie der gleiche Baum in seinen Dimensionsveränderungen erfasst. Bei permanent angelegten Inventuren hingegen ist das Auffinden des gleichen Baumes bei der Wiederholungsaufnahme möglich. Da die permanente Markierung des Stichprobenmittelpunktes mit einem höheren finanziellen Aufwand verbunden ist (Zöhrer, 1980), wurden in der Vergangenheit im Staatswald Bayerns beide Inventurverfahren angewendet. Aufgrund des Aufnahmeverfahrens sowie der Datenhaltung in der Zentralen Forsteinrichtungsdatenbank liegen mit den Daten der Forstbetriebsinventur für den Staatswald Bayerns keine echten Zeitreihendaten vor, wie sie beispielsweise im langfristigen ertragskundlichen Versuchswesen durch die dauerhafte Beobachtung ein und desselben Baumes über der Zeit entstehen.

Daten, die durch kontinuierliche Beobachtung eines Objektes entlang der Zeitachse entstehen, werden als Längsschnittdaten (longitudinal data) bezeichnet. Bei den Daten der bayerischen Stichprobeninventur handelt es sich allerdings lediglich bei Vorliegen permanent angelegter Inventuren um Längsschnittdaten. Bei temporär angelegten Stichproben handelt es sich um Querschnittdaten (cross-sectional-data). Derartige Daten entstehen im langfristigen ertragskundlichen Versuchswesen z. B. bei der Beobachtung von Wuchsreihen (Pretzsch, 2001). Datenquellen, die sich aus unechten und echten Zeitreihendaten zusammensetzen, werden als Panel-Daten

oder als time-series cross-sectional data bezeichnet (Schlittgen und Streitberg, 2001).

Von Gadow (2003) bezeichnet mit dem Begriff Intervalldaten, Daten die letzterer Kategorie angehören. Er sieht in der temporären Anlage und mindestens einmaligen Wiederholungsaufnahme von kleinen Versuchsflächen ein großes Potenzial für die Zukunft der Waldwachstumsforschung. Er weist allerdings auch darauf hin, dass es ganz wichtig ist, waldbauliche Maßnahmen möglichst gleichzeitig mit dem Aufnahmezeitpunkt durchzuführen, um waldwachstumskundlich sinnvolle Aussagen treffen zu können.

Im Rahmen eines betrieblichen Inventurverfahrens im bayerischen Staatswald wird gitternetzartig ein meist quadratisches Stichprobennetz über die Flächen eines Forstbetriebes gelegt, welches mit dem dort gewonnenen Datenmaterial Aussagen für verschiedene Stratenebenen zulässt. Die Gitternetzweite orientiert sich dabei an der standörtlichen Ausgangssituation, an der Struktur der Bestockung sowie am Informationsbedürfnis der mit der Waldbewirtschaftung betrauten Personen (Kurth, 1994). Bei der Wahl geeigneter Gitternetzabstände ermöglicht die Verteilung der Stichprobenpunkte die repräsentative Erfassung von Baumdaten auf nahezu allen Standorteinheiten über dem gesamten vorkommenden Altersspektrum. Zum Stichtag einer Inventur liegen somit Daten vor, in denen räumlich nebeneinander liegende Bestände das zeitliche Nacheinander der Bestände repräsentieren. Dies entspricht der Idee von Wuchsreihen im ertragskundlichen Versuchswesen. Inventurdaten unterscheiden sich von Versuchsflächendaten, die auf Wuchsreihen aufgenommen werden, durch einen geringeren Detailgrad. Beispielsweise werden an Inventurpunkten keine Größen wie Kronenansatzhöhen oder Kronendimensionen erfasst. Ein weiterer großer Unterschied besteht in der Größe der aufgenommenen Flächen. Sowohl wegen des geringeren Detailgrads als auch wegen der geringeren Flächengröße ermöglichen Inventurdaten praktisch keine Aussage zur räumlichen Wuchskonstellation der Einzelbäume im Bestand. Der Vorteil der Auswertung von Inventurdaten besteht hingegen darin, dass einerseits am Inventurpunkt selbst die standörtlichen Verhältnisse vergleichsweise homogen sind und andererseits die große Anzahl an aufgenommenen Punkten standörtlich und regional repräsentative Aussagen ermöglicht.

Nach Pretzsch (2001) sind echte Zeitreiheninformationen aufgrund der Langlebigkeit der untersuchten Baumobjekte mit dem Nachteil behaftet, dass sie unter Umständen historische Wuchsbedingungen und Behandlungsprogramme abbilden. Sie besitzen allerdings den Vorteil, dass sie die gleiche, bekannte Vorgeschichte besitzen. Unechte Zeitreihen hingegen reflektieren aktuelle Wuchsbedingungen und moderne Behandlungsprogramme, die Vorgeschichte kann aber unter Umständen sehr unterschiedlich sein.

Die Verwendung von Daten räumlich nebeneinander liegender Flächen als Ersatz für ein zeitliches Nacheinander muss aufgrund der Tatsache, dass eine Konstanz der Standortbedingungen nicht mehr gegeben ist (Pretzsch, 1985; Prietzel et al., 1997), kritisch hinterfragt werden. Dies gilt insbesondere für die Verwendung von Forstinventurdaten als Grundlage für die regionale, standortbezogene Feinjustierung von einzelbaumbasierten Wachstumsmodellen. Biber (1996) hat sich u. a. mit der Verwendung von unechten Zeitreihendaten zur Justierung bzw. Parametrisierung eines derartigen Modells beschäftigt. Er sieht für ein Wachstumsmodell, das mit Daten einer Mischbestands-Wuchssreihe parametrisiert wurde, einen positiven Effekt, da es die aktuellen Entwicklungstendenzen von Bäumen definierter Gestalt unter bestimmter Konkurrenz gut widerspiegeln kann, da die Daten nicht wie bei Versuchsflächen unterschiedlich aktuell sind. Die Wuchssreihendaten aller beobachteten Altersstufen beinhalten hingegen eine Wuchsdynamik, die durch die Umweltbedingungen geprägt wird. Für diese Umweltbedingungen kann zumindest für wenige Jahre Geltung bzw. Konstanz angenommen werden. Er folgert, dass daher ein derartig parametrisiertes Modell für kurzfristige Prognosen ein schlagkräftiges Werkzeug sein kann. Für langfristige Prognosen sind derartig eingestellte Modelle allerdings nicht geeignet. „Je weniger die Annahme der Wuchssreihe als Zeitreihe zutrifft, in desto größerem Ausmaß führen längere Prognosen selbst bei weiterhin konstant bleibenden Standortbedingungen zu systematischen Fehleinschätzungen“ (Biber, 1996).

Wie in den Abschnitten 2.1.1 und 2.1.2 dieser Arbeit ansatzweise aufgeführt, haben zahlreiche Autoren zum Teil erhebliche Mehrzuwächse gegenüber gebräuchlichen Ertragstafeln aufgrund verbesserter standörtlicher Wachstumsbedingungen gefunden. Die zunehmende, z. T. verbessernd wirkende Veränderung der Standorte der letzten Jahrzehnte hat für das Wachstum von Einzelbäumen die Konsequenz, dass Bäume mit gleichem physischem Alter höhere Höhen und Durchmesser aufweisen als vergleichbar alte Bäume vor mehreren Jahrzehnten. Diese Tatsache hat erhebliche Auswirkungen auf die Entwicklung der Alters-Höhenbeziehung.

Trägt man von Längs- und Querschnittsdaten die Höhen über den entsprechenden Altersangaben auf und gleicht die beiden Punktwolken mit der Wachstumsfunktion nach Chapman-Richards aus, so sieht man einen charakteristisch unterschiedlichen Verlauf beider Kurven (vgl. Abb. 90).

Der Kurvenverlauf der echten Zeitreihendaten verläuft aufgrund der suboptimalen Wachstumsbedingungen des Baumes in seiner Jugendphase flacher, steigt mit ähnlichem Steigungsverhalten in der Vollkraftphase an und schwingt in der Abschwungphase leicht asymptotisch ein. Die Kurve für den Ausgleich der unechten Zeitreihendaten steigt hingegen in der Jugendphase etwas steiler an, verläuft in der

Vollkraftphase in etwa parallel zur Kurve der Längsschnittdaten und steigt in der Abschwungphase mit einer geringeren Steigung.

Grund für das beschriebene Steigungsverhalten der beiden Ausgleichskurven sind die verbesserten Wachstumsbedingungen. Die deutlich verbesserte Nährstoffversorgung sowie evtl. günstigere Temperatur- und Niederschlagsverhältnisse können in der Jugendphase von den Bäumen besser umgesetzt werden als von alten Bäumen und sind so maßgeblich dafür verantwortlich, dass die Kurvenverläufe wie beschrieben verlaufen.

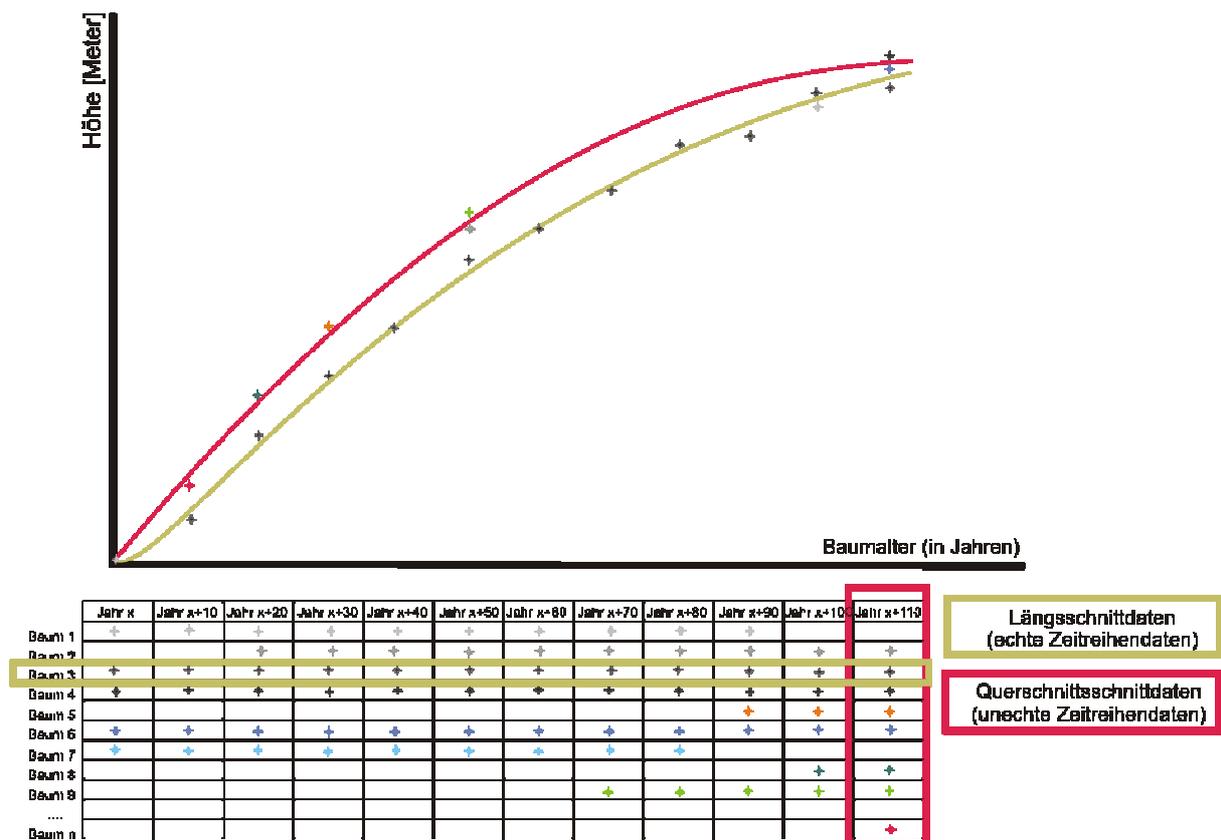


Abb. 90: Schematische Darstellung der Alters-Höhenbeziehung von Bäumen beim regressionsanalytischen Ausgleich von echten und unechten Zeitreihendaten

Der Kurvenverlauf des Ausgleichs unechter Zeitreihendaten wird insbesondere in höheren Baumaltern durch zwei weitere Faktoren beeinflusst: Zum einen wird in der Literatur immer wieder berichtet, dass insbesondere in höheren Baumaltern die Bäume auf Standorten, die sich für das Wachstum der Baumarten als günstig erwiesen haben, aufgrund des früheren Erreichens von Zieldimensionen eher geerntet werden als auf schlechten Standorten (Burschel und Huss, 1997). Dies führt zu einer geringeren Steigung im Vergleich zur Verwendung von Längsschnittdaten, da ab einem, durch die Leistungsfähigkeit des Standortes bestimmten Alter Bäume, die die Zieldimensionen erreicht haben, eher geerntet werden. Zum anderen ist

durch die unbekannte Geschichte der Bäume, insbesondere bei Bäumen höheren Alters, die Gefahr gegeben, dass sie in der Jugend eine längere, unbekannte Übershirmungsphase durchlebt haben und sie dadurch z. T. ihre maximal möglichen Baumhöhen noch nicht erreicht haben, was bei einem regressionsanalytischen Ausgleich ebenso zu einem stärkeren Abfallen der Kurve im Alter führt.

Bei dem Ausgleich der unechten Zeitreihendaten wird darauf hingewiesen, dass es sich hierbei definitionsgemäß nicht um eine Alters-Höhenkurve handelt (Kramer, 1988). Der Ausgleich mit der Wachstumsfunktion von Chapman-Richards wird dennoch als statthaft erachtet, da es sich bei der von beiden entwickelten Wachstumsfunktion nicht um eine Beschreibung oder Nachbildung eines real beobachteten Wachstumsverlaufes handelt, sondern lediglich um eine Verallgemeinerung der von Bertalanffy zur Beschreibung des Körpermassenwachstums entwickelten Wachstumsfunktion (Zeide, 1993).

Der vorgelegten Arbeit unterliegt die Annahme, dass es die Nutzung von Querschnittsdaten bzw. ggf. von Querschnitts- und kurzfristigen Längsschnittdaten ermöglicht, einzelbaumbasierte Wuchsmodelle regional und standörtlich zu justieren. Es wird angenommen, dass es möglich ist, aus den Höhen der Einzelbäume über dem jeweiligen Baumalter potenzielle, standortabhängige, funktionelle Beziehungen abzuleiten, die für derartige Zwecke geeignet sind. Die Gründe für diese Annahme liegen in der vergleichsweise geringen Beeinflussbarkeit der Höhenentwicklung von Einzelbäumen durch aktive wirtschaftliche Maßnahmen i. V. m. der vergleichsweise hohen Anzahl an Daten, die eine weitgehende Flächen- und Standortrepräsentativität gewährleisten sollen und die eine maximale Höhenentwicklung in einem bestimmten Alter aufzeigen sollen. Weiterhin folgt die Arbeit der Ansicht Bibers (1996), der die gleiche Aktualität der Baumdaten als Garant dafür sieht, aktuelle Entwicklungstendenzen von Bäumen abbilden zu können. Biber sieht, wie bereits in diesem Abschnitt dargestellt, eine Einschränkung darin, dass derartig eingestellte Modelle lediglich für kurzfristige Prognosen besonders geeignet sind, und dass die Annahme der Wuchsreiheneigenschaft als Zeitreihe eine wesentliche Voraussetzung für die Prognosegenauigkeit ist. Diese Einschränkung lässt sich sehr gut mit den Zielen dieser Arbeit vereinen, da im Rahmen der Arbeit eine Softwarelösung erstellt werden sollte, die eine regionale, standortbezogene Feinjustierung des Wuchsmodells SILVA für kurzfristige Prognosen ermöglicht. Das entwickelte Softwaresystem nimmt aktuell vorliegende Forstinventurdaten auf und nutzt damit die Wachstumstendenzen der letzten Forsteinrichtungsperiode bzw. -perioden, um möglichst treffsichere Prognosen für die nächste Forsteinrichtungsperiode zu erstellen. Die Annahme der Konstanz der Standortbedingungen innerhalb der letzten

Aufnahmeperiode ist dabei ebenso maßgeblich wie die Annahme, dass sich das Wachstum in den darauffolgenden Jahren nicht wesentlich anders entwickeln wird. Anders ausgedrückt soll das Wachstum der Waldbäume innerhalb der letzten 10 bis 20 Jahre in einer Region zum Maßstab gemacht und für Prognosen über die nächsten 10 bis maximal 20 Jahre nutzbar gemacht werden. Dies stellt allerdings eine Einschränkung der Extrapolierbarkeit der Modellergebnisse eines derartig justierten Modells dar.

Eine grundlegende Arbeit, die Forstinventurdaten der Österreichischen Nationalen Forstinventur heranzieht, um ein einzelbaumbasiertes Wachstumsmodell zu parametrisieren, stammt von Filla (1981). Diese Arbeit entstammt allerdings einer Zeit, in der die Anzahl der Datensätze aus Forstinventuren vergleichsweise gering war und auf der anderen Seite die datenverarbeitungstechnischen Möglichkeiten noch nicht sehr weit entwickelt waren. Filla nutzte die Daten eines österreichischen Forstbetriebes (Seilern-Aspang, Litschau), um das Wachstumsmodell FOREST von Ek und Monserud (1974) in einem frühen Entwicklungsstadium zu parametrisieren. Er versuchte, das programminterne Konkurrenzmodell, das Verjüngungsmodell sowie das Mortalitätsmodell mit diesen Inventurdaten sowie mit „waldwachstumskundlichen Gesetzmäßigkeiten“ zu parametrisieren. Aufgrund fehlender Daten zur Parametrisierung der Teilmodelle wurden im Rahmen dieser Arbeit zudem permanente Versuchsflächen angelegt und erstmalig aufgenommen. Wegen der geringen Anzahl der Datensätze in Verbindung mit zum Teil fehlenden Informationen und eingeschränkter Datenverarbeitungskapazitäten „zeigen die Ergebnisse der Testläufe (nach der Neuparametrisierung) zumindest theoretisch mögliche Entwicklungen auf“. Filla weist in seiner Arbeit abschließend darauf hin, dass „eine kritische Ergebniskontrolle erst nach den Folgeaufnahmen der angelegten Dauerversuchsflächen zu erwarten sei“.

Im Vergleich zu dieser ersten Arbeit, die frühzeitig versuchte Forstinventurdaten zur Einstellung grundlegender Modellbeziehungen zu nutzen, weist der Handlungsrahmen für diese Arbeit deutliche Verbesserungen auf. Aufgrund der Bereitschaft und des Interesses der ehemaligen Bayerischen Staatsforstverwaltung können für diese Arbeit flächen- und standortrepräsentative Forstinventurdaten aus ganz Bayern herangezogen werden. Auch modernere Datenverarbeitungstechniken erlauben bessere bzw. leichter anzuwendende Auswertungsmöglichkeiten.

Bezüglich der Eigenschaften haben sich die Inventurdaten allerdings gegenüber des frühen Ansatzes von Filla kaum verändert. Im Rahmen der Forstinventur werden Daten nach forstbetrieblichen Erfordernissen aufgenommen. Die Intensität der Datenaufnahme hängt dabei insbesondere vom Informationsbedürfnis der Entscheidungsträger bzw. von den festgelegten Sicherungsgrenzen der Ergebnisse

ab (Zöhrer, 1980). Die Datenaufnahmen erfolgen dabei im Gegensatz zur Aufnahme von langfristig angelegten Versuchsflächen vergleichsweise extensiv. Aufgenommen werden waldwachstumskundliche Kenngrößen, auf deren Grundlage sich über rechnerische Hilfsbeziehungen die für forstbetriebliche Entscheidungen benötigten Größen ermitteln lassen. Das Aufnahmeprogramm umfasst dabei Standards die regelmäßig Baumart, Durchmesser und stichprobenartig die Baumhöhen erfassen (vgl. z. B. Forstdirektion Oberfranken-Mittelfranken, 2001). Im Falle der Anlage oder Wiederholungsaufnahme permanenter Probeflächen sind zudem Informationen zu den Positionen der Bäume vorhanden. Neben dem Standardprogramm können weiterhin Zusatzinformationen wie z. B. Waldschäden oder Qualitätsmerkmale aufgenommen werden.

Bei der Betrachtung der dieser Arbeit zugrundeliegenden Daten muss allerdings berücksichtigt werden, dass es sich um Daten aus dem Wirtschaftswald handelt. Die einzelnen Probeflächen liegen fast ausnahmslos innerhalb von Waldbeständen des Wirtschaftswaldes. Eine Ausnahme hierbei stellen die Inventurdaten für das Gebiet des Nationalparks Bayerischer Wald dar, die sich ebenfalls in der Zentralen Forsteinrichtungsdatenbank der Bayerischen Staatsforsten befinden. Der Wirtschaftswald in Bayern wird traditionsgemäß nachhaltig bewirtschaftet (Bayerische Landesanstalt für Wald und Forstwirtschaft, 2003). Extreme Behandlungsprogramme, die waldwachstumskundliche „Grenzverhältnisse“ abbilden können, spiegeln sich daher kaum in den Inventurdaten wieder bzw. können aufgrund einer fehlenden Verknüpfung zu Behandlungsdaten nicht von z. B. historischen „Störereignissen“ wie Windwürfen o. ä. unterschieden werden. Vergleicht man z. B. für die in dieser Arbeit verwendeten Datensätze die maximalen Durchmesser der Bäume (Abb. 12 bis 14) auf den einzelnen Probepunkten mit den Ergebnissen von Untersuchungen zum Solitärwachstum von Bäumen in Bayern (Uhl et al., 2006) so sieht man, dass innerhalb der Forstinventurdatensätze praktisch keine Bäume mit vergleichbaren Dimensionen aufgefunden werden konnten.

Ein weiteres, diese Arbeit beeinflussendes Charakteristikum der Forstinventurdaten besteht darin, dass die Menge der Inventurdaten von Jahr zu Jahr stetig zunimmt. Demzufolge nimmt auch die Menge der verfügbaren Informationen zu den maßgeblichen Triebkräften des Waldwachstums für die einzelnen Baumarten zu.

Vorliegende Arbeit konzentriert sich darauf, aus den flächen- und standortrepräsentativen „Massendaten“ der Forstinventur für die Hauptbaumarten der Wirtschaftswälder einer Region den Standort-Leistungsbezug zu extrahieren. Verfolgt wird aufgrund der Eigenschaften der Forstinventurdaten ein empirisch-statistischer Ansatz zur Einschätzung eines regionalen, standörtlich justierten Leistungsbezuges für SILVA-Simulationen. Dies steht im Gegensatz zu dem bisherigen modellhaften Ansatz zur standortsensitiven Modellierung der Alters-

Höhenbeziehung von Waldbäumen, wie sie von Kahn (1994) erarbeitet wurde und in die aktuelle Version des Wachstmodells SILVA 2.2 eingegangen ist (Pretzsch und Kahn, 1998). Der in dieser Arbeit verfolgte Ansatz steht im Zusammenhang mit dem geplanten Einsatzbereich für kurz- und mittelfristige Prognosen insbesondere im Rahmen von Forsteinrichtungsverfahren. Er bringt allerdings folgende Einschränkungen im Einsatzbereich mit sich: Zum einen können nur Steuergrößen für Baumarten ermittelt werden, die bereits in ausreichender Anzahl auf den zu untersuchenden Standorten vorhanden sind. Eine direkte Übertragung auf andere Standorte ist nicht oder aufgrund örtlichem Erfahrungswissen nur unsicherheitsbehaftet möglich. Zum anderen basiert die Extraktion der Steuergrößen aus Querschnittsdaten bzw. z. T. aus Paneldaten auf der Überlegung, dass der Wachstumsgang der letzten Forsteinrichtungsperiode die Grundlage für die Abbildung von Wachstumsprozessen im Modell darstellt. Diese Einschränkung impliziert aufgrund der aktuellen Kenntnisse um die Veränderung der wachstumsbeeinflussenden Faktoren in den letzten Jahrzehnten (vgl. z.B. Pretzsch, 1985 oder Prietzel et al., 1997), dass die neuen Steuergrößen nur für kurzfristige Prognosen, die sich auf eine Laufzeit von maximal 15 bis 20 Jahren beschränken, zum Einsatz kommen sollten. Diese Einschränkung wird allerdings aufgrund des geplanten Einsatzzweckes in Kauf genommen.

Mit dem verfolgten Ansatz dieser Arbeit wird versucht, eine regionale und standörtliche Feinjustierung des Wachstmodells SILVA durch eine möglichst zutreffende Überführung der in der Realität vorgefundenen Inventurdaten in das Modell zu erreichen. Im Gegensatz zu der von Vanclay (1994) vorgeschlagenen und der z. B. von Sterba und Monserud (1997) durchgeführten Modellanpassung durch Einführung von Korrekturfaktoren zur regionalen Anpassung der maßgeblichen Wachstumsverhältnisse, verfolgt diese Arbeit die Überlegung, dass die Anpassung der modellinternen Triebgrößen mit Hilfe einer zweiten Datenquelle erfolgen sollte, die hierfür besser geeignet ist. Im Falle des Wachstmodells SILVA können die Daten des langfristigen ertragskundlichen Versuchswesens in Bayern als maßgebliche Grundlage benutzt werden, um modellinterne Wachstumsbeziehungen wie z. B. die Konkurrenzverhältnisse oder die Mortalitätsbeziehungen realitätskonform zu parametrisieren. Hierbei wird unterstellt, dass sich die Standortveränderungen nicht auf die Konkurrenzverhältnisse zwischen den Baumarten auswirken. Andernfalls wäre eine Neuparametrisierung der modellinternen Gleichungen mit Versuchsflächendaten der gleichen Zeitperiode erforderlich.

Demnach versucht der Ansatz dieser Arbeit alle verfügbaren Datengrundlagen mit ihren in Hinblick auf Datenqualität und Datenquantität unterschiedlichen Daten und Informationsgehalten kombiniert zu nutzen und an der Stelle in das Wachstmodell einfließen zu lassen, wo sie den größten Zugewinn an Prognosegenauigkeit und

Treffsicherheit erwarten lassen. Ein in diesem Zusammenhang ähnlicher Ansatz wird z. B. auch von Garcia (2005) beschrieben.

6.2 Diskussion der angewendeten Methoden

Vorliegende Arbeit versucht mit Hilfe eines zweistufigen Ansatzes, standortbezogene Leistungsparameter mit einem hohen Autonomiegrad aus den Inventurdaten zu extrahieren. Zum einen wird versucht, mit Hilfe von Techniken aus dem Bereich des Data-Mining standörtliche Muster aus den Inventurdaten zu isolieren. Hierfür wurde die Algorithmeklasse der Klassifikationsalgorithmen herangezogen. Diese Algorithmeklasse wird dem Bereich des supervised-learning, also dem sogenannten überwachten Lernen zugeordnet. Klassifikationsalgorithmen versuchen auf der Basis einer durch den Benutzer vorgenommenen Vorklassifikation des Datenmaterials mit Hilfe von rekursiven Partitionierungstechniken die maßgeblichen wachstumsbestimmenden Faktoren für die Zuordnung zu den einzelnen Klassen zu isolieren. Innerhalb des Data-Minings wäre auch die Anwendung von Clusteralgorithmen denkbar gewesen, die versuchen selbständig homogene Klassen innerhalb der Datensätze zu finden. Clusteralgorithmen werden daher dem Bereich des unsupervised learnings zugeordnet, da sie versuchen weitgehend selbständig Klassen zusammenzufassen. Nach Witten und Frank (2005) arbeiten diese Algorithmen grundlegend nach einem als „covering-principle“ bezeichnetem Ansatz. Sie versuchen auf Basis der vorhandenen Attribute Regeln zu definieren, die eine möglichst gute „Abdeckung“ einzelner homogener Gruppen innerhalb des Datenmaterials gewährleisten. Aus drei Gründen wurde den Klassifikationsalgorithmen der Vorzug gegeben: Erstens neigen die Clusteralgorithmen dazu Gruppen zu bilden, die möglichst geklumpt innerhalb der Daten vorliegen. Für eine Isolation von standörtlich möglichst homogenen Gruppen aus Alters-Höhenbeziehungen ist es dagegen notwendig, möglichst homogene Gruppen über dem gesamten Altersspektrum zusammenzufassen, um dann, auf Basis der regressionsanalytisch vorgenommenen Klasseneinteilung, ein in sich konsistentes und stabiles, standörtlich geprägtes System aus Wachstumskurven ableiten zu können. Zweitens existieren kaum Clusteralgorithmen die in der Lage sind, mehrdimensionale Triebkräfteinflüsse über der Zeit ausreichend gut zu berücksichtigen. Drittens wird zwar ein hoher Autonomiegrad bei der Extraktion standörtlicher Muster aus den Massendaten der Forsteinrichtung angestrebt, die weitgehend selbständige Zusammenfassung der Clusteralgorithmen von standörtlich möglichst homogenen Gruppen innerhalb des Datenmaterials widerstrebt allerdings

der Forderung, dass die aus dem Datenmaterial extrahierten standörtlichen Muster dem Benutzer in Form von Hypothesen zur Verfügung gestellt werden sollen. Hierfür bieten die Klassifikationstechniken bzw. deren visuelle Darstellungstechniken in Form von Entscheidungsbäumen bessere Möglichkeiten, um Benutzeränderungen in geeigneter Weise möglich zu machen.

Nach dem Klassifikationsschritt wird nach dem vorgestellten Ansatz für jede nach klein- und großstandörtlichen Aspekten zusammengefasste Gruppe, die durch einen Experten als solche anerkannt wurde, mit Hilfe von Techniken der nichtlinearen Regression ein neuer Satz von Parameterwerten für die Chapman-Richards-Wachstumsfunktion als Grundlage für die Einsteuerung in das Wuchsmodell SILVA ermittelt. Eine innere Konsistenz des entstehenden Wachstumskurvenfächers wird dabei dadurch erzwungen, dass die Anpassung einer nichtlinear regressionsanalytischen Wachstumskurve für mittlere standörtliche Höhenwachstumsgänge über dem (Inventur-)Alter für die einzelnen Standortgruppen lediglich durch eine Veränderung des Wertes des Parameters A der Wachstumsfunktion nach Chapman-Richards erfolgt. Dies hat zur Konsequenz, dass die aus den Wachstumsfunktionen abgeleiteten potenziellen Höhenzuwächse gemäß der standörtlichen Klassifikation in sich konsistent sind, d. h. dass die potenziellen Höhenzuwächse auf Standorten, die für die Baumart ungünstiger sind, niedriger sind als auf Standorten, die für diese Baumart besser geeignet sind. Ein Nachteil dieses Verfahrens besteht darin, dass durch die Variation des Wertes des Parameters A bei gleichzeitigem Konstanthalten der Werte für die Parameter k und P zwar der Steigungsverlauf und die Asymptote variiert werden, der Wendepunkt aller Kurven allerdings immer konstant gehalten wird. Nach Sloboda (1971) verschiebt sich der Wendepunkt der Höhenzuwachsentwicklung von Waldbäumen mit abnehmender Standortgüte hin zu höheren Altern, was sich letztendlich auch in der veränderten Lage der Wendepunkte der Wachstumskurven widerspiegelt. Nachdem allerdings eine nach Chapman-Richards hergeleitete Wachstumsfunktion die Grundlage für die Ermittlung der potenziellen Höhenzuwächse darstellt, wird dieser Mangel in Kauf genommen, da nach dieser Wachstumsfunktion der maximal erreichbare Wendepunkt der Kurve bei einem Drittel der maximal erreichbaren Höhe (Asymptote) liegt (Zeide, 1993).

Ergebnis des zweistufigen Prozesses aus der Anwendung von Klassifikationstechniken und regressionsanalytischer Verfahren ist also ein in sich schlüssiges System aus Wachstumskurven, das standörtliche Faktoren in der vorliegenden Verschlüsselung der Forstpraxis berücksichtigt. Damit kann einer weiteren eingangs dieser Arbeit gestellten Forderung für den Einsatz in der forstlichen Praxis, nämlich der Verwendung der in der Forstpraxis üblichen

Standortkodierung, Rechnung getragen werden. Zur Ableitung der maßgeblichen Triebkräfte des Wachstums einer Region wurde, wie in Abschnitt 3.1.6 dieser Arbeit beschrieben, ein „forstliches“ Data Warehouse entwickelt, welches für jeden Probekreis bzw. letztendlich für jeden Baum eines Probepunktes eine Verknüpfung mit den standörtlichen Wachstumsfaktoren vorsieht. Derartige Entwicklungen sind in anderen Wirtschaftszweigen bereits häufig die Grundlage zur Ermittlung potenziell neuer Informationen aus Massendaten. Im forstlichen Bereich ist dagegen diesbezüglich bisher nichts bekannt. Wie eine erste vorsichtige Einschätzung der im Rahmen dieser Arbeit durchgeführten Befragung von Forstbetrieben zur Datenhaltung der Forstinventurdaten in Deutschland gezeigt hat, befinden sich zahlreiche, insbesondere kleinere, Forstbetriebe derzeit erst in der Anfangsphase der Nutzung von Datenbanken und Datenbanksystemen. Größere Forstbetriebe dagegen weisen bereits eine längere Tradition in der EDV-gestützten Haltung von Forstinventurdaten auf. Neft (2005) hat, wie bereits erwähnt, darauf hingewiesen, dass die Bayerischen Staatsforsten sowie die Bayerische Staatsforstverwaltung in den letzten Jahren und Jahrzehnten eine Datengrundlage von unschätzbarem Wert für forstbetriebliche Fragestellungen aufgebaut haben. Mit dem vorgestellten zweistufigen Ansatz können mit Hilfe der eingeführten analyseorientierten Datenhaltung neue standörtliche Erkenntnisse hypothetisch aus den Forstinventurdaten abgeleitet werden, die einerseits eine Bestätigung von forstlichem Erfahrungswissen darstellen, die andererseits aber auch den Ansatzpunkt für neue Forschungsarbeiten darstellen können. Über den Aufbau eines forstlichen Data Warehouses mit Hilfe geeigneter Datenbanksysteme, die in der Lage sind, geographische Daten in geeigneter Weise zu verwalten, können diese neuen Erkenntnisse bereits frühzeitig an fachkundige Entscheidungsträger innerhalb der Forstbetriebe herangebracht werden. Dies kann z. B. in Form von thematischen Karten o. ä. erfolgen. Die Verschneidung derartiger thematischer Karten ermöglicht z. B. eine produktionsoptimierte Bestockungszielplanung. Hierdurch erhält die vorgestellte Methodik, neben der Herleitung von praxisgerechten Steuergrößen für die Simulation mit dem Wachstumsmodell SILVA, bereits einen eigenen Gebrauchswert. Beispiele für Datenbanksysteme mit derartigen Fähigkeiten sind das kommerzielle Produkt Oracle 10g mit der als „Spatial Database Engine“ bezeichneten Erweiterung für die Verwaltung von Geodaten oder die frei verfügbare Datenbanklösung PostgreSQL mit der Erweiterung „PostGIS“ (Bilke, 2006). Gerade PostgreSQL erscheint als ein Beispiel für eine Lösung, die es kostengünstig auch kleineren Forstbetrieben ermöglichen könnte, sowohl Forsteinrichtungs- als auch Geodaten in einer Form zu halten, die flexible Auswertungen ermöglichen. Derartige Lösungen würden in Zukunft den zeitaufwendigen Schritt der Datenvorverarbeitung für die Herleitung der standortsensitiven Steuergrößen aus den Forstinventurdaten eines

Forstbetriebes deutlich vereinfachen und damit eine leichtere regionale und standortbezogene Feinjustierung des einzelbaumbasierten Wachstumsmodells SILVA mit Hilfe der eigenen Forstinventurdaten ermöglichen.

Zahlreiche Autoren (z.B. Fayyad et al., 1996; Knobloch und Weidner, 2000; Ferber, 2003 oder Witten und Frank, 2005) weisen im Zusammenhang mit der Anwendung von Data-Mining-Techniken auf die große Bedeutung der Datenqualität des verwendeten Datenmaterials hin. Vor der Anwendung des zweistufigen Verfahrens zur Herleitung von standortsensitiven Startgrößen für Berechnungsläufe mit dem Wachstumsmodell SILVA wurde ein restriktives Auswahlverfahren zur Sicherstellung der Auswahl von qualitativ hochwertigen Daten durchgeführt. Dieses Verfahren ist verantwortlich dafür, dass für einzelne Wuchsregionen Bayerns keine Ableitung eines regionalen Standort-Leistungsbezuges nach dem beschriebenen Verfahren möglich war. Diese Einschränkung wurde bewusst in Kauf genommen, da bei erfolgreicher Entwicklung der Methodik eine benutzerfreundliche Anwendung erstellt werden sollte, die eine rasche Anwendbarkeit und Wiederholbarkeit des Verfahrens auf neue Daten ermöglichen soll. Ein Prototyp dieser Anwendung wurde in Abschnitt 5 dieser Arbeit vorgestellt und liegt dieser Arbeit bei.

6.3 Diskussion der vorgestellten Ergebnisse

In Abschnitt 4.3 wurden die Ergebnisse exemplarisch für das Wuchsgebiet 6.2 „Südliche Frankenalb und Südlicher Oberpfälzer Jura“ vorgestellt. Im Rahmen dieser Arbeit wurden mit Hilfe der in Abschnitt 3.2.1.5 beschriebenen Algorithmen CHAID, CART und C4.5 Klassifikationsanalysen durchgeführt. Die generierten Entscheidungsbäume unterschieden sich zwar im Baufbau deutlich, in der Wahl der angewendeten Prädiktoren zur Unterteilung der Daten in Untergruppen ließen sich allerdings kaum Unterschiede feststellen. Aufgrund der Möglichkeit zum Aufbau n-närer Bäume zeigte der CHAID-Algorithmus allerdings die Tendenz zum Aufbau von sehr breiten Entscheidungsbäumen mit einer Vielzahl von Einzelelementen in den Blättern. Bei kleineren Datenmengen hingegen war es gelegentlich mit diesem Algorithmus nicht möglich, eine weitergehende Klassifikation anhand des zugrundeliegenden Datenmaterials durchzuführen. Mit Hilfe des CART-Algorithmus wurden, aufgrund der Verwendung gleichartiger Trenngrößen zur Generierung von Untereinheiten des Baumes, ähnliche Baufbauformen erreicht wie mit dem C4.5-Algorithmus. Beide Algorithmen erzeugen binäre Bäume, die die Datensätze automatisch in einer Art und Weise unterteilen, die mit forstlichem Sachverstand leicht nachvollzogen werden kann.

Die automatisch vorgenommene Klassifikation der Datensätze bzw. die im zweiten Schritt ermittelten Werte für die Parameter A, k und P sollten allerdings nicht unmittelbar zur Durchführung von Berechnungen mit SILVA verwendet werden. Vielmehr sollte ein forstlich sachkundiger Experte die Ergebnisse der Klassifikation beurteilen und eventuell den generierten Baum nachträglich manuell aufgrund von lokalem Expertenwissen abändern. Bei der Betrachtung der Klassifikationsergebnisse müssen dabei stets die Anzahlen der Datensätze in den einzelnen Unterklassen berücksichtigt werden. Eine Stärke des angewendeten Verfahrens zur Klassifikation der Datensätze liegt in der Tatsache, dass die angewendeten Klassifikationstechniken versuchen, alle vorliegenden Datensätze Unterklassen zuzuordnen (Witten und Frank, 2005). Dieses angewendete „Prinzip der Vollständigkeit“ bedingt allerdings auch die Berücksichtigung des Expertenwissens bei der Beurteilung der Klassifikationsergebnisse. Bei den vorliegenden Daten muss ein Anwender der beschriebenen Methodik bzw. der in Abschnitt 5 beschriebenen Anwendung über das notwendige lokale Erfahrungswissen verfügen, um die Ergebnisse der Klassifikation im Kontext der regionalen Wuchsverhältnisse für die einzelnen Baumarten einschätzen zu können. Bei der Berücksichtigung von klimatischen Standortfaktoren als Prädiktoren der Klassifikationsläufe darf zudem nicht außer Acht gelassen werden, dass diese in vorliegender Arbeit einer Datenquelle entstammen, die über ganz Bayern modellhaft diese Werte ermittelt hat (BayFORKLIM, 1996). Dies stellt eine potenzielle Fehlerquelle für die Durchführung der weitergehenden Datenanalysen dar, die allerdings mangels genauerer Daten für die einzelnen Waldmesspunkte in Kauf genommen wurde. Ähnliches gilt für die Berücksichtigung der standörtlichen Faktoren. Bei der Durchführung einer Standortkartierung im Gelände werden die Standorte mit Hilfe einer visuell-sensorischen Standortsansprache eingewertet. Dies beinhaltet eine stark subjektive Komponente, die sich u. a. negativ auf die Ergebnisse der Klassifikation auswirken kann. Diese Fehlerquelle könnte in Zukunft, evtl. durch die Berücksichtigung neuer standortkundlicher Erkenntnisse wie beispielsweise des von Schultze et. al (2005) beschriebenen quantitativen Verfahrens zur Kennzeichnung des Wasserhaushaltes von Waldböden, ganz oder teilweise ausgeschaltet werden.

Bezüglich der verwendeten Klassifikationsalgorithmen konnten keine eindeutigen Vorteile des einen oder anderen Algorithmus bei der Suche nach standörtlichen Mustern innerhalb der Inventurdaten gefunden werden. Insgesamt erscheint es wichtig, dass die verwendeten Algorithmen über geeignete Mechanismen verfügen, mit deren Hilfe der Baumaufbau von Anfang an (Forward-Pruning) oder nachträglich (Backward-Pruning) gesteuert werden kann. In vorliegender Arbeit konnten die eingängigsten Baumaufbauformen mit einer Beschränkung der Baumtiefe auf maximal fünf Ebenen und einer Mindestanzahl von 5 % der Gesamtfälle, mindestens

aber 100 Fällen pro Blattelement ermittelt werden. Für die Entwicklung einer Anwendung zur semiautomatischen Feinjustierung des Wachstumsmodells SILVA nach der vorgestellten zweistufigen Methodik wird im Prototyp der C4.5-Algorithmus angewendet. Grund für diese Entscheidung ist die freie Verfügbarkeit einer getesteten JAVA-Implementierung innerhalb des Data-Mining-Frameworks WEKA der University of Waikato (Witten und Frank, 2005), die sich in der zu erstellenden Anwendung nutzen lässt.

Im zweiten Schritt der Herleitung der AkP-Steuergrößen wird der mittlere regressionsanalytische Ausgleich der Alters-Höhenentwicklung einer Baumart in einem speziellen Wuchsbezirk als Grundlage für die Bestimmung der Werte für den Parameter A für die neuen Klassen verwendet. Hierbei entsteht durch die Variation lediglich des Parameters A eine in sich schlüssige Schar von Wachstumskurven für die betrachtete Baumart der Wuchsregion auf den verschiedenen Standorten. Die Anhebung der Kurven im Parameter A basiert dabei auf der Annahme, dass die nach groß- und kleinstandörtlichen Aspekten zusammengefassten Datensätze eine ähnliche Varianz bezüglich der Streuung besitzen.

6.4 Ausblick

Mit vorgelegter Arbeit wurden erstmals Inventurdaten zur Ermittlung eines regionalen Standort-Leistungsbezuges genutzt. Die erarbeitete Methodik stellte die Grundlage für die Ermittlung von wachstumsbestimmenden Steuergrößen für die Baumarten Fichte, Tanne, Kiefer und Lärche sowie Eiche und Buche in zahlreichen Wuchsregionen Bayerns dar. Für diese Baumarten hat sich die Anwendung der neuen Methodik bewährt und hat beim Vergleich von Prognoseergebnissen und realem Waldwachstum auf Reinbestandsversuchsflächen gute Ergebnisse erbracht. Durch die vergleichsweise restriktive Vorauswahl der Daten konnten für mehrere Regionen in Bayern keine Steuergrößen ermittelt werden. Grund hierfür waren z. T. Datenmängel oder die mangelnde Verfügbarkeit von Informationen. Beispiele hierfür sind die mangelhafte Kennzeichnung von Höhenmessbäumen im Wuchsgebiet 10 („Oberpfälzer Wald“) oder das Nichtvorliegen einer digitalen Standortkartierung für die Wuchsbezirke 8.3 („Fichtelgebirge“) bis 8.7 („Selb-Wunsiedler Bucht“). Im Zuge neuer Inventurverfahren bzw. des Abschlusses von Digitalisierungsverfahren können derartige „Lücken“ mit der Zeit geschlossen werden. Einen Beitrag hierzu soll die im Rahmen dieser Arbeit erstellte Softwarelösung leisten, die eine leichte Anwendbarkeit des Verfahrens auf neue Daten ermöglicht. Weiterhin soll es mit

dieser Softwarelösung möglich werden, weitere Bauarten (z. B. Edellaubhölzer wie Bergahorn der Esche) rasch zu bearbeiten. Für die Verwendung der abgeleiteten Größen zur Einstellung des Standort-Leistungsbezuges in SILVA muss allerdings sichergestellt sein, dass für neue Baumarten wie z. B. Bergahorn oder Esche in SILVA eine Parametrisierung der Teilmodelle im Modellkern vorliegt.

Weiterhin liefert die vorgestellte Methodik neue Hypothesen zur Standortabhängigkeit des Waldwachstums in Bayern. Pretzsch und Utschig (1998) haben beispielsweise in einer großangelegten Untersuchung zu den Wachstumstrends der Baumart Fichte in Bayern darauf hingewiesen, dass das Wachstum durch den Faktorenkomplex Standort und Klima gesteuert wird, der Zusammenhang zwischen den einzelnen Wirkungsfaktoren allerdings noch nicht vollständig geklärt ist. Die im Rahmen dieser Arbeit vorgestellte Methodik bringt Ergebnisse, die entweder existierendes Praxiswissen bestätigen oder neue Hypothesen formulieren. Beispiel hierfür könnte die Formulierung einer Ober- und Untergrenze der Niederschlagsmenge sein, die für das Höhenwachstum der Fichte in Bayern optimal ist. Ausgangspunkt für diese Hypothese sind die Ergebnisse der Entscheidungsbaumklassifikation für die Wuchsbezirke 4.2, 6.2, 11.2 und 14.4/ 1 nach denen in der trockensten Region Bayerns (4.2) das Höhenwachstum auf guten Standorten durch weitere verfügbare Niederschläge steigt, während in anderen Regionen auf guten Standorten die Höhenwuchsleistung der Baumart Fichte mit steigender Niederschlagsmenge (6.2, 11.2 und 14.4/ 1) rückläufig ist.

Die vorgestellte Methodik hat die real gemessenen Höhen der Bäume über den Altersangaben der Inventur als Grundlage einer standortgruppenpezifischen Leistungseinwertung genutzt. Hiermit konnten, nach entsprechender Einsteuerung in SILVA, realitätskonforme Simulationsergebnisse beobachtet werden. Die Altersangaben der Inventur basieren dabei einerseits auf Altersfortschreibungen, andererseits auf Altersermittlungen (Jahrringszählungen, Auszählung der Jahrestriebe etc.) oder aber auf Altersschätzungen. Für vorliegende Arbeit wurden nur tatsächlich höhengemessene Bäume der obersten Bestandesschichten herangezogen. Für diese Bäume fanden sich vergleichsweise gute Altersangaben in den Inventurdaten. In Zukunft wird allerdings diese Altersansprache, aufgrund sich ändernder Waldaufbauformen, schwieriger bzw. z. T. für die forstliche Praxis unbedeutender. Aus diesem Grund sollte bereits rechtzeitig über die Weiterentwicklung des vorgestellten Verfahrens dahingehend nachgedacht werden, dass die Altersangabe durch eine zweite, real gemessene bzw. über allometrische Grundprinzipien ableitbare Größe ersetzt werden kann.

Darüber hinaus könnten die Ergebnisse beider Schritte (Klassifikation und Regression) als Ausgangspunkt dafür genutzt werden, die regional unterschiedliche

Standortkodierung über das Höhenwachstum der Bäume auf den (zusammengefassten) Standorten in ein Relationsniveau zu bringen.

Die grundlegende Validierung der vorgestellten Methodik war nur aufgrund der Existenz von Flächen des langfristigen ertragskundlichen Versuchswesens in Bayern möglich. Diese Flächen bzw. die aufstockenden Bäume werden kontinuierlich seit Jahrzehnten mit größtmöglicher Sorgfalt in zahlreichen Aufnahmegrößen aufgenommen. Hieraus entsteht der unschätzbare Wert dieser Datengrundlage für die Waldwachstumsforschung in Deutschland und in Europa. Die Validierungsbemühungen haben allerdings auch gezeigt, dass dieses Flächennetz z. T. eine uneinheitliche Ausstattung bezüglich der Baumarten und Standortsspektren aufweist. Es scheint daher notwendig, dieses Flächennetz auch in Zukunft durch Neuanlagen, Wiederanlagen oder den Ausbau von Versuchsflächen in seinem Wert zu erhalten.

7. Zusammenfassung und Summary

7.1 Zusammenfassung

In Zeiten sich verengender Handlungsspielräume stellen forstliche Wuchsmodelle immer wichtiger werdende Werkzeuge für die Planung und den Betriebsvollzug in Forstbetrieben dar. Die Akzeptanz dieser Modelle hängt dabei neben Aspekten der Softwareergonomie insbesondere von der Übereinstimmung der Prognoseergebnisse mit erwarteten oder beobachteten Wachstumsvorgängen der Realität ab. Mit dem einzelbaumbasierten, standortsensitiven Wachstumssimulator SILVA wurde am Lehrstuhl für Waldwachstumskunde der Technischen Universität München ein derartiges Wuchsmodell entwickelt, dessen interne Formelgleichungen bisher insbesondere auf der Grundlage von Daten des langfristigen ertragskundlichen Versuchswesens in Bayern eingestellt wurden.

Mit vorliegender Arbeit wird ein Verfahren vorgestellt, welches die Forstinventurdaten aus dem Bayerischen Staatswald nutzt, um eine regionale, baumartenspezifische Einschätzung des Standort-Leistungsbezugs auf Basis der Alters-Höhenentwicklung der Bäume vorzunehmen. Diese Einwertung stellt die Grundlage für eine regionale, standortbezogene Feinjustierung des Wuchsmodells SILVA dar.

Der vorgestellte empirisch-statistische Ansatz baut methodisch auf einem Klassifikations- und einem Regressionsschritt auf.

Im Klassifikationsschritt werden die gesamten Forstinventurdaten einer Baumart und Region über die wachstumsbestimmenden Triebkräfte in verschiedene Klassen aufgeteilt. Die Grundlage für diese Partitionierung stellt die reale Alters-Höhenentwicklung der höhengemessenen Bäume der Forstinventur in Abhängigkeit ihres Alters dar, die als Leitgröße für die Wuchsleistung einer Baumart auf einem Standort angesehen wird. Zur Durchführung der Klassifikation werden Klassifikationsalgorithmen aus dem Bereich des Data-Mining angewendet. Diese partitionieren rekursiv die zugrundeliegende Datenmenge so lange, bis bestimmte Abbruchregeln erfüllt sind bzw. fassen sie gleichartige Datensätze zusammen. Im konkreten Anwendungsfall werden Bäume mit einer gleichartigen Alters-Höhenentwicklung aufgrund ähnlicher standörtlicher Ansprüche zusammengefasst bzw. von anderen Gruppen abgetrennt. Die Ergebnisse der Klassifikation werden in Form von Entscheidungsbäumen dargestellt (vgl. Abb. 17, S. 75).

Der Klassifikationsschritt stellt die Grundlage für den darauf folgenden Regressionsschritt dar. Für jede isolierte Gruppe an Datensätzen werden die Koeffizienten der Parameter A , k und P der Wachstumsfunktion nach Chapman-

Richards ermittelt, wobei nicht standardmäßig für jede Gruppe die drei Koeffizienten neu ermittelt werden. Statt dessen wird lediglich der Wert des Parameters A einer im Vorfeld des Klassifikationsschrittes für die gesamten Forstinventurdaten durchgeführten nichtlinearen Regressionsrechnung solange variiert, bis 95 % der zu einer eigenen Gruppe zusammengefassten Datensätze unterhalb der Wachstumskurve zu liegen kommen. Dadurch, dass nur der Wert des Parameters A verändert wird und die Koeffizienten der Parameter k und P konstant gehalten werden, entsteht eine in sich konsistente Schar an Wachstumskurven, deren Anzahl mit der Anzahl der im Klassifikationsschritt erzeugten Gruppen übereinstimmt (vgl. Abb. 18, S. 75).

Die Koeffizienten der Parameter A, k und P können in SILVA zur Einsteuerung des standortbezogenen, baumartenspezifischen Leistungsbezuges benutzt werden.

Ein Vergleich der Prognoseergebnisse mit Daten des langfristigen ertragskundlichen Versuchswesens in Bayern hat eine gute Übereinstimmung zwischen den in der Realität beobachteten und den im Modell geschätzten Höhenzuwächsen nach der neuen Methodik zur Einsteuerung des Standort-Leistungsbezuges ergeben. Insbesondere für junge Reinbestände konnten für alle Baumarten größere Höhenzuwächse ermittelt werden als mit der bisherigen modellbasierten Einsteuerungsmöglichkeit, während für alte Reinbestände geringere Höhenzuwächse als bisher zu beobachten waren. Die Übereinstimmung der vom Modell erzeugten Höhenzuwächse mit den Ergebnissen der Versuchsfelder war dabei größer als mit dem bisher standardmäßig angewendeten modellhaften Ansatz.

Der hier vorgestellte statistisch-empirische Ansatz besitzt gegenüber dem modellhaften Ansatz den Hauptvorteil, dass real beobachtete Größen die Grundlage für die Generierung von Steuergrößen darstellen. Dadurch wird das Höhenwachstum im Modell auf die maximal in der Realität beobachteten Grenzen beschränkt. Gleichzeitig stellt diese Vorgehensweise allerdings auch eine Einschränkung im Anwendungsbereich dar, da sichere Prognosen nur für die Regionen oder Standorte möglich sind, auf denen bisher bereits eine ausreichende Zahl von Bäumen für die Herleitung der Grundbeziehungen vorhanden sind. Aufgrund einer mangelnden Extrapolierbarkeit der Ergebnisse wird der Einsatzbereich zudem auf Prognosen mit kurzem Prognosezeitraum beschränkt, da aufgrund der statistisch-empirischen Herangehensweise unterstellt wird, dass das Waldwachstum in der kommenden Prognoseperiode vergleichbar mit dem Wachstum der zur Herleitung der Steuergrößen verwendeten Inventurdaten der zurückliegenden Inventurperiode ist.

Im Rahmen dieser Arbeit wurde ein Prototyp einer als Silva-PEM bezeichneten Software erstellt, mit deren Hilfe für neue oder veränderte Forstinventurdaten eine Einschätzung des Standort-Leistungsbezuges automatisiert ermöglicht wird. Mit Hilfe dieser Software wird der Aufwand für die Ermittlung neuer, standortsensitiver

Steuergrößen für SILVA erheblich vermindert. Weiterhin können durch die leichte Anwendbarkeit und Wiederholbarkeit neue Forstinventurdaten oder Inventurdaten für andere Baumarten oder Regionen zur Generierung von SILVA-Steuergrößen herangezogen werden.

7.2 Summary

In times of narrow forestry scopes growth models become increasingly important planning and management tools within forest enterprises. The acceptance of such models depends as well on software ergonomic aspects as especially on prognosis accuracy of simulated forest growth processes. Such a growth model was realized with the site and distance-dependent single-tree forest growth simulator SILVA of the Chair of Forest Yield Science at TU München, which is mainly adjusted by data derived from long-term experimental plots of Bavarian Forest Research.

Within this research work a two-staged methodology is introduced, which uses forest inventory data from state forest in Bavaria as an alternative datasource to adjust the growth model regionally and site-specific.

The introduced empirical-statistical approach is separated into a classification-step and a regression-step.

Within the classification-step all regionally available inventory data for one tree species are subdivided in several subpartitions. Basis for this subdivision is the real observed age-height development of the trees, which is an indicator for the site-specific growth potential of this species. The classification is done by applying classification algorithms derived from Data Mining techniques. These algorithms subdivide datasets by recursive partitioning techniques until a preliminary defined stop-criterion is reached or by uniting similar datasets. Here tree datasets with a comparable age-height development because of similar site-specific demands are grouped or separated from other groups. The results of this classification procedure are presented in form of decision trees (cf. fig. 17, p. 75).

Based on the classification-results the regression-step follows. For each isolated group of datasets the coefficients of the parameters A , k and P of Chapman-Richards' growth equation are calculated. However, this calculation is not a new or independent calculation of each coefficient for each group. Rather it is based on the variation of the parameter A from a preliminary calculated set of coefficients derived from calculation for the whole available inventory data of a growth region. This variation is stopped when the growth curve forms the upper boundary of 95 % of the

data. By adjusting only the parameter A and keeping constant k and P, a conclusive group of growth curves is generated. The number of growth curves is equivalent to the number of leaves generated within the decision tree (cf. fig. 18, p. 75).

These calculated coefficients can be used for adjusting the site dependent growth behaviour of tree species within the forest growth model SILVA.

The comparison of the predicted results with data derived from long term experimental plots in Bavaria shows a high level of accordance between observed and simulated height increment values using the new adjusting technique. Especially for young pure stands for all tree species higher height increments could be observed while for old pure stands the estimated height increments are lower than by the application of the old method for adjusting site dependency within SILVA. In general the level of accordance between observed height increment proved better when applying the new techniques for adjusting site-dependency of height growth processes within SILVA.

The main advantage of the introduced statistical-empirical approach in comparison to the established model-based approach for adjusting site-dependency of height growth processes is that actually measured trees build the basis for the derivation of site-dependent parameters for calculation purposes. In this way simulated height growth is limited to maximum observed values. In this way also the site-coding of reality is easily introduced within the growth model SILVA

On the other hand this way of adjusting SILVA limits the application of the growth model because reliable prognosis runs can only be done for regions where enough tree data are available for justification purposes. Also the range of application is limited to short term prognoses because the statistical-empirical approach assumes that the simulated forest growth development is comparable to the real observed forest growth development within the last forest inventory period.

Within this research work a prototype of a software called Silva-PEM was developed. This software provides the automatized calculation of the site-dependent parameterisation for SILVA prognosis runs. Because it is easy to apply this software on new forest inventory data or inventory data of different growth regions, it should be possible to adjust SILVA site-specifically for different regions or tree species in the future.

8. Literatur

- Agrawal, R., Srikant, R., 1994: **Fast algorithms for mining association rules in large databases**. Research report RJ 9839, IBM Almaden Research Center, San Jose, California, 34 p.
- Amaro, A., 2003: **The SOP Model. The parameter estimation alternatives**. In: Amaro, A., Reed, D., Soares, P., 2003: *Modelling Forest Systems*, CABI Publishing, p. 213-224.
- Arbeitskreis Standortkartierung, 2003: **Forstliche Standortaufnahme**. 6. Auflage IHW-Verlag, Eching, 352 S.
- Assmann, F., 1961: **Waldetragskunde**. Organische Produktion, Struktur, Zuwachs und Ertrag von Waldbeständen. BLV Verlagsgesellschaft, München, Bonn, Wien, 490 S.
- Balzert, H., 2005: **Lehrbuch Grundlagen der Informatik. Konzepte und Notationen in UML 2, Java 5, C++ und C# Algorithmik und Software-Technik Anwendungen**. Elsevier, Spektrum Akademischer Verlag, München, 925 S.
- Bauer, A., Günzel, H., 2004: **Data Warehouse Systeme**. dPunkt-Verlag. 608 S.
- Bayerische Landesanstalt für Wald und Forstwirtschaft, 2001: **Die regionale natürliche Waldzusammensetzung Bayerns**. LWF Wissen Nr. 32, 107 S.
- Bayerische Landesanstalt für Wald und Forstwirtschaft, 2003: **Der Wald für morgen – Eine Naturalbilanz über 25 Jahre naturnahe Forstwirtschaft im Bayerischen Staatswald**. LWF Wissen Nr. 39, 79 S.
- Bayerische Landesanstalt für Wald und Forstwirtschaft, 2005: **Die zweite Bundeswaldinventur 2002. Ergebnisse für Bayern**. LWF Wissen Nr. 49, 103 S.
- Bayerisches Staatsministerium für Ernährung, Landwirtschaft und Forsten, 1986: **Der Wald in Unterfranken**. Bayerisches Staatsministerium für Ernährung, Landwirtschaft und Forsten (Hrsg.), 34 S.

- Bayerisches Staatsministerium für Landwirtschaft und Forsten, 1999: **Datenbankbeschreibungen und Tabellenabfragen.** (Stand Sept. 1999), unveröffentlicht, 31 S.
- Bayerisches Staatsministerium für Landwirtschaft und Forsten, 2004: **Jahresbericht der Bayerischen Staatsforstverwaltung. Band II (Statistikteil).** 68 S.
- BayFORKLIM, 1996: **Klimaatlas für Bayern.** Bayerischer Klimaforschungsverbund (Herstellungsleitung Gerhard Enders). 114 S.
- Biber, P., 1996: **Konstruktion eines einzelbaumorientierten Wachstumssimulators für Fichten-Buchen-Mischbestände im Solling.** Dissertation. Berichte des Forschungszentrums Waldökosysteme Göttingen, Reihe A, Bd. 142, 1996, 252 S.
- Bilke, J., 2006: **Das Ende der Inseln – PostgreSQL als Geodaten-Server.** iX (Magazin für professionelle Informationstechnik) 11/ 2006, S. 160-163.
- Böckmann, T., 2004: **Möglichkeiten und Grenzen von Waldwachstumsmodellen aus Sicht der forstlichen Praxis.** In Waldwachstumsmodelle für Prognosen in der Forsteinrichtung – Stand der Wissenschaft in Mitteleuropa. Berichte Freiburger forstliche Forschung der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg, Freiburg i. Br., Heft 50, S. 95-102.
- Borgelt, Ch., 2004: **Intelligente Datenanalyse.** Unveröffentlichtes Vorlesungsskript des Instituts für Wissens- und Sprachverarbeitung der Otto-von-Guericke Universität Magdeburg, 126 S.
- Borgelt, Ch., Kruse, R., 1998: **Attributsauswahl für die Induktion von Entscheidungsbäumen.** In Nakhaeizadeh, G. (Hrsg.), 1998: Data Mining – Theoretische Aspekte und Anwendungen. Physika-Verlag Heidelberg, erschienen in der Reihe: Beiträge zur Wirtschaftsinformatik, Band 27, S. 23- 55.
- Bortz, J., 2005: **Statistik für Human- und Sozialwissenschaftler.** Springer Medizin Verlag Heidelberg, 6. Auflage, 882 S.

- Borrie, W., Christensen, N., Watson, A. E., Miller, Th. A., McCollum, 2002: **Public Purpose Recreation Marketing: A Focus on the relationship between the public and public lands.** Journal of Park and Recreation Administration. Vol. 20, No. 2. p. 49-68.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, Ch., J., 1984: **Classification and Regression Trees.** CRC Press LLC, Boca Raton, Florida, 358 p.
- Bundeswaldinventur, 2005: **Die zweite Bundeswaldinventur – BWI².** Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft (Hrsg.), **Inventurbericht**, Berlin, 230 S.
- Burschel, P., Huss, J., 1997: **Grundriss des Waldbaus.** Parey Buchverlag im Blackwell Wissenschafts-Verlag, Berlin, 487 S.
- Clasen, M., Stricker, S., 2004: **Programmierung von Online-Befragungen mit HTML und Perl.** Zeitschrift für Agrarinformatik 3/ 04. S. 57-62.
- Connolly, T., Begg, C. und Strachan, A., 2002: **Datenbanksysteme. Eine praktische Anleitung zu Design, Implementierung und Management.** Addison-Wesley Verlag, Boston, München, 1077 S.
- Corona, P., Dettorio, S., Filigheddu, M. R., Maetzke, F., Scotti, R., 2005: **Site quality evaluation by classification tree: an application to cork quality in Sardinia.** European Journal of Forest Research (2005), 124, p. 37-46.
- Döbbeler, H., 2004: **Simulation und Bewertung von Nutzungsstrategien unter heutigen und veränderten Klimabedingungen mit dem Wachstumsmodell SILVA 2.2.** Dissertation an der Fakultät für Forstwissenschaften und Waldökologie an der Georg-August-Universität Göttingen. (unveröffentlicht). 494 S.
- Draper, N. R., Smith, H., 1981: **Applied regression analysis.** 2. ed., Wiley, New York, 709 p.
- Duden, 1996: **Duden – Die deutsche Rechtschreibung.** Bibliographisches Institut Mannheim, 21. Auflage. 1136 S.

- Dursky, J., 1999: **Modellvalidierung durch Vergleich von Prognose und Wirklichkeit.** Tagungsbericht zur Jahrestagung der Sektion Ertragskunde im DVFFA in Vollpriehausen, S. 33-43.
- Dzeroski, S., Kobler, A., Gjorgjosky, V., Panov, P., 2006: **Using Decision Trees to Predict Forest Stand Height and Canopy Cover from LANDSAT and LIDAR Data.** In: Tochtermann, K., Scharl, A. (eds.), 2006: *EnvirolInfo 2006. Managing Environmental Knowledge. Proceedings of the 20th International Conference "Informatics for Environmental Protection (Graz)*, Shaker Verlag, Aachen, S. 125-135.
- Ek, A. R., Monserud, R. A., 1974: **FOREST: A computer based model for simulating the growth and reproduction of mixed species forest stands.** School of Natural Resources (ed.), College of Agricultural and Life Sciences, 107 p.
- Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., 1996: **From Data Mining to Knowledge Discovery in Databases.** *AI Magazine* (<http://www.aaai.org/Magazine/magazine.php>), Fall 1996, S. 37-54.
- Fei, S., 2005: **Hotspot Detection for Chestnut Regeneration.** In *Proceedings of The 6th Annual National Conference on Digital Government Research*. Atlanta, Georgia, p. 243-244.
- Ferber, Reginald, 2003: **Information Retrieval.** Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web, dpunkt Verlag, Heidelberg, 400 S.
- Filla, K., 1981: **Die Parametrisierung von Einzelstammwachstumsmodellen über die bei Forstinventuren erhobenen Daten.** Dissertation an der Universität für Bodenkultur Wien, 81 S.
- Flanagan, D., 2004: **Java examples in a nutshell.** O'Reilley. 3rd edition, 701 p.
- Forstdirektion Oberfranken-Mittelfranken, 2001: **Aufnahmeanweisung für die Wiederholungsaufnahme im Forstamt Weißenstadt bzw. für die Neuanlage in den Bereichen des ehemaligen Forstamtes Wunsiedel.** (unveröffentlicht). 44 S.

- Fuchs, A., 1993: **Wiederholungsaufnahme und Auswertung einer permanenten Betriebsinventur im Bayerischen Forstamt Ebrach.** Forstliche Forschungsberichte München. Nr. 131, 157 S.
- Fukuda, K., Pearson Ph., 2006: **Data mining and image segmentation approaches for classifying defoliation in aerial forest imagery.** Paper presented on the iEMSS (International Environmental Modelling and Software summit 2006, Vermont, USA), Workshop 7, http://www.iemss.org/iemss2006/papers/w7/83_Fukuda_4.pdf (erneuter Abruf vom 07.02.2007)
- FVS, 2006: **“Forest Vegetation Simulator”-Homepage.** <http://www.fs.fed.us/fmse/fvs> (Abruf vom 14.06.2006)
- Garcia, O., 2005: **Comparing and combining stem analysis and permanent sample plot data in site index models.** Forest Science 51(4) 2005. p. 277-283
- Göttlein, A., 1994: **Die Wuchseistung von Furniereichen in Abhängigkeit vom Standort.** Allg. Forst- u. J.-Ztg., 165. Jg., 5-6, S. 81-87.
- Grob, H. L., Bensberg, F., 1999: **Das Data-Mining-Konzept.** Arbeitsberichte der Reihe „Computergestütztes Controlling“ des Instituts für Wirtschaftsinformatik der Westfälischen-Wilhelms-Universität Münster, Nr. 8, 33 S.
- Haas, R. und Schreiner, U., 2002: **Java-Technologien für Unternehmensanwendungen. Konzepte, Methoden und Anwendungen der Java 2 Enterprise Plattform.** Carl Hanser Verlag, München, Wien, 513 S.
- Häckel, H., 1993: **Meteorologie,** 3. verb. Auflage, Ulmer, Stuttgart, 402 S.
- Hanewinkel, M., 2005: **Neural networks for assessing the risk of windthrow on the forest division level: a case study in southwest Germany.** Eur. J. Forest Research (2005) 124: S. 243-249.

- Harrer, E., 2004: **Wachstumstrends der Buche in Bayern und Untersuchungen zur Auswirkung der Bestandesbehandlung auf das Wachstum der Buche**, (unveröffentlichter) Abschlußbericht zum gleichnamigen Projekt des Kuratoriums der Bayerischen Landesanstalt für Wald und Forstwirtschaft (W 33), 124 S.
- Hasenauer, H., 1997: **Dimensional relationships of open-grown trees in Austria**. Forest Ecology and Management 96, p. 197-206.
- Hasenauer, H., 1999: **Methodische Aspekte bei der Evaluierung von Baummodellen**. Tagungsbericht zur Jahrestagung der Sektion Ertragskunde im DVFFA in Vollpriehausen, S. 45-53.
- Hasenauer, H., 2004: **Waldwachstumsmodelle als Grundlage einer nachhaltigen Waldwirtschaft**. In: Waldwachstumsmodelle für Prognosen in der Forsteinrichtung – Stand der Wissenschaft in Mitteleuropa. Berichte „Freiburger Forstliche Forschung“ der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg, Freiburg i. Br., Heft 50, S. 15-31.
- Hasenauer, H., 2006: **Concepts within tree growth modelling**. In Hasenauer, H. (Ed.): Sustainable Forest Management. Growth Models for Europe. Springer Verlag, Berlin, Heidelberg, New York, 398 p. (hier: p: 3-19).
- Hasenauer, H., Merkl, D., Weingartner, M., 2001: **Estimating Tree Mortality of Norway Spruce Stands with Neural Networks**. Advances in Environmental Research, Vol. 5, Elsevier, 2001. p. 341-356.
- Hein, M., Zeller, H., 2003: **Java Web Services. Entwicklung plattformübergreifender Dienste mit J2EE, XML und SOAP**. Addison-Wesley Verlag, Boston, München, 353 S.
- Herbolsheimer, A., 2002: **Datenbank-Programmierung**. Beispiellösungen mit Access, SQL Server und PostgreSQL. Addison-Wesley Verlag, Boston, München, 463 S.
- Herling, H., 2005: **Wachstumstrends der Eiche in Bayern**, Abschlussbericht (unveröffentlicht) des gleichnamigen Projektes des Kuratoriums der Bayerischen Landesanstalt für Wald und Forstwirtschaft (W 35). 88 S.

- Hill, M. C., 1998: **Methods and guidelines for effective model calibration**, US Geological Survey Water-Resources Investigations Report 98-4005, <http://water.usgs.gov/nrp/gwsoftware/modflow2000/WRIR98-4005.pdf> (Abruf vom 13.01.2005)
- Hippner, H., Wilde, K. U., 2001: **Der Prozess des Data Mining im Marketing**. In: Handbuch Data Mining im Marketing. Vieweg. S. 22-84
- IBusiness, 2004: Newsletter der HighText Verlag Graf und Treplin OHG, Wilhelm-Riehl-Straße 13, 80687 München. <http://www.ibusiness.de/members/knowledgemanager/db/knowledgemanager.973166796dt.2006jg.html> (Abruf vom 04.05.2006)
- IPCC, 2007: **Fourth Assessment Report** of Working Group I of the Intergovernmental Panel on Climate Change (**Summary for Policymakers**), 18 p., <http://www.ipcc.ch/spm2feb07.pdf> (Abruf vom 07.02.2007)
- Iratxeta, C., Bork, P., Bork, P., Andrade, M.A., 2002: **Association of genes to genetically inherited diseases using data mining**. Nature Genetics. Vol. 31. S. 316-319
- Jobst, F., 2002: **Programmieren in Java**. Hanser Wirtschaftsverlag, 4. Auflage, 528 S.
- Kahn, M., 1994: **Modellierung der Höhenentwicklung ausgewählter Baumarten in Abhängigkeit vom Standort**. Forstl. Forschungsber. München, Vol. 141, 221 S.
- Kass, G. V., 1980: **An exploratory technique for investigating large quantities of categorical data**. Appl. Staist. (1980), 29, No. 2, pp. 119-127
- Kennel, R., 1973: **Die Bestimmung des Ertragsniveaus bei Buche**. Forstwiss. Cbl. 92(5): 226-234.
- King, R. D., Karwath, A., Clare, A., Dehapse, L., 2000: **Genome scale prediction of protein functional class from sequence using Data Mining**. Proceedings KDD 2000, Boston, USA, S. 384-389 (<http://www.informatik.uni-freiburg.de/~karwath/publications/kdd2000.pdf>) (Abruf vom 04.05.2006)

- Knaufft, F. J., 2003: **Landscape Visualization with three different Forest Growth Simulators**. In Amaro, A., Reed, D., Soares, P., 2003 Modelling Forest Systems. CABI Publishing, p. 347-359.
- Knobloch, B., 2000: **Der Data-Mining-Ansatz zur Analyse betriebswirtschaftlicher Daten**. Bamberger Beiträge zur Wirtschaftsinformatik. Nr. 58, Bamberg, 111 S.
- Knobloch, B., Weidner, J., 2000: **Eine kritische Betrachtung von Data-Mining-Prozessen**. In: Data Warehousing 2000. Methoden, Anwendungen, Strategien, Heidelberg (Physica) 2000, S. 345-365.
- Krahl, D., Windheuser, U., Zick, K.-F., 1998: **Data Mining – Einsatz in der Praxis**, Addison-Wesley Verlag, Bonn, 176 S.
- Kramer, H., 1988: **Waldwachstumslehre. Ökologische und anthropogene Einflüsse auf das Wachstum des Waldes, seine Massen- und Wertleistung und die Bestandessicherheit**. Verlag Paul Parey, Hamburg, Berlin, 374 S.
- Kramer, H. und Akça, A., 1995: **Leitfaden zur Waldmesslehre**. J. D. Sauerländer's Verlag, Frankfurt a. Main, 266 S.
- Krüger, G., 2006: **Handbuch der Java-Programmierung**. Addison-Wesley (München), 6. Auflage, 1280 S.
- Kublin, E., Scharnagel, G., 1988: **Verfahrens- und Programmbeschreibung zum BWI-Unterprogramm BDAT**. Forstl. Versuchs- und Forschungsanstalt Baden-Württemberg, Abschlußbericht, 87 S.
- Küsters, E., 2001: **Wachstumstrends der Kiefer in Bayern**. (unveröffentlichte) Dissertation am Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt. 197 S.
- Küsters, U., 2001: **Data Mining Methoden: Einordnung und Überblick**. In Hippner, H., Küsters, U., Meyer, U., Wilde, K. D., 2001: Handbuch Data Mining im Marketing. Vieweg Verlag, S. 95 - 123
- Kurth, H, 1994: **Forsteinrichtung – Nachhaltige Regelung des Waldes**. Deutscher Landschaftsverlag Berlin GmbH, 592 S.

- Laatsch, W., 1963: **Bodenfruchtbarkeit und Nadelholzanbau**. Bayerischer Landwirtschaftsverlag, München, 237 S.
- Lemm, R., 2004: **Ein Modell für alle Fälle – einfache, alternative Simulationsmodelle**. In: Berichte „Freiburger Forstliche Forschung“ der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg, Freiburg i. Br., Heft 50, S. 32-48.
- Lessard, V. C., McRoberts R. E., Holdaway M. R., 2001: **Diameter Growth Models Using Minnesota Forest Inventory and Analysis Data**. Forest Science 47 (3) 2001, p. 301-310.
- Lindner, M., Badeck, F.-W., Bartelsheimer, P., Bonk, S., Cramer, W., Dieter, M., Döbbeler, H., Dursky, J., Duschl, Ch., Fabrika, M., Frömdling, D., Gundermann, E., Henning, P., Hölzer, W., Lasch, P., Leischner, B., Liesebach, M., Pommerening, A., Pott, M., Pretzsch, H., Schlott, W., Scholz, E., Spellmann, H., Suckow, E., Suda, M., Wolff, B., 2002: **Integrating Forest Growth Dynamics, Forest Economics and Decision Making to Assess the Sensitivity of the German Forest Sector to Climate Change**. Forstw. Cbl. 121, Supplement 1 (2002), S. 191-208.
- Ludwig, L., Flies, D., Wilson, A., 2001: **Data Mining Techniques applied to the relationship of Latitude and the lifespan of Aspen Trees**. http://www.cs.uni.edu/~fienup/mics_2001/MICS2001_Proceedings/papers/ludwig.pdf (Abruf vom 20.06.2006)
- Lusti, M., 2002: **Data Warehousing und Data Mining. Eine Einführung in entscheidungsunterstützende Systeme**. Springer-Verlag, Berlin, Heidelberg, 444 S.
- Mantel, St., 2000: **Einführung in das Data Warehouse-Konzept**. 14 S. (unveröffentlichte Informationsschrift der MIK AG). [http://www.mik.com/WebSite/MIKWebArchiv.NSF/PDF/DataWarehouse/\\$File/DataWarehouse.pdf](http://www.mik.com/WebSite/MIKWebArchiv.NSF/PDF/DataWarehouse/$File/DataWarehouse.pdf) (erneuter Abruf vom 26.09.2006)
- Moosmayer, H. U., Schöpfer, W., 1972: **Beziehungen zwischen Standortfaktoren und Wuchsleistung der Fichte**. Allg. Forst- und Jagdzeitung 143, S. 203-215.

- Moosmayer, H. U., 1987: **Regionale Vergleiche zum Wachstum der Buche in Baden-Württemberg**, Mitteilungen des Vereins für Forstliche Standorterkundung und Forstpflanzenzüchtung 33, S. 36-40.
- mySAP CRM, 2006: **Wertschöpfung durch mySAP CRM** (Benchmarking Studie) <http://www.sap.com/germany/media/50623000.pdf> . 24 S. (Abruf vom 19.06.2006)
- Nagel, J., Schmidt, M., 2006: **The silvicultural decision support system BWINPro**. In Hasenauer, H. (Ed.): Sustainable Forest Management. Growth Models for Europe. Springer Verlag. Berlin, Heidelberg, New York. 398 p., hier: p. 59-63.
- Naisbitt, J., 2000: **Megatrends 2000**. Ten new directions for the 1990s. William Morrow –Verlag. 384 S.
- Neft, R., 2005: **Bestätigung und Ansporn für die Bewirtschaftung des Staatswaldes – Stellenwert und Konsequenzen der Inventurergebnisse für das Unternehmen**. LWF aktuell 51/ 2005, S. 17-19.
- Newnham, R.M., 1964: **The development of a stand model for Douglas-fir**. Ph.D. thesis, Fac. Forestry, Univ. B.C., Vancouver, 201 p.
- Precht, M., Kraft, R., 1993: **Biostatistik II**. Oldenburg Wissensch. Verlag. 5. völlig überarbeitete Auflage, 457 S.
- Pretzsch, H., 1985: **Der Wachstumsmerkmale Oberpfälzer Kieferbestände in den letzten 30 Jahren**. Vitalitätszustand – Strukturverhältnisse – Zuwachsgang. Allg. Forst Zeitschrift. 42. S. 1122-1126.
- Pretzsch, H., 1993: **Analyse und Reproduktion räumlicher Bestandesstrukturen. Versuche mit dem Strukturgenerator STRUGEN**. Schriften aus der Forstlichen Fakultät der Universität Göttingen und der Nieders. Forstl. Versuchsanstalt, Band 114, J. D. Sauerländer's Verlag, Frankfurt a. Main, 87 S.
- Pretzsch, H., 2001: **Modellierung des Waldwachstums**. Blackwell Wissenschafts-Verlag, Berlin, Wien, 336 S.

- Pretzsch, H., 2002: **Grundlagen der Waldwachstumsforschung**. Blackwell Verlag GmbH, Berlin, Wien, 414 S.
- Pretzsch, H., 2003a: **Strategische Planung der Nachhaltigkeit auf Forstbetriebsebene. Beitrag der Waldwachstumsforschung**. Forstw. Cbl. 122 (2003), S. 231-249.
- Pretzsch, H., 2003b: **Diversität und Produktivität von Wäldern**. Allg. Forst- und Jagdzeitung. 174 (5-6). S. 88-98.
- Pretzsch, H., 2004: **Wuchsmodell für die Planung auf Betriebsebene**. In: Waldwachstumsmodelle für Prognosen in der Forsteinrichtung – Stand der Wissenschaft in Mitteleuropa. Berichte „Freiburger Forstliche Forschung“ der Forstlichen Versuchs- und Forschungsanstalt Baden-Württemberg, Freiburg i. Br., Heft 50, S. 49-73.
- Pretzsch, H., Kahn, M., 1998: **Konzeption und Konstruktion des Wachstumsmodells SILVA 2.2**. (unveröffentlichter) Abschlußbericht zum Projekt W 28, Teil 2 (Methodische Grundlagen). 257 S.
- Pretzsch, H., Seifert, S., 2000: **Methoden zur Visualisierung des Waldwachstums**. Forstw. Cbl., 119. Jg., S. 100-113.
- Pretzsch, H., Utschig, H., 2000: **Wachstumstrends der Fichte in Bayern**. Mitteilungen der Bayerischen Staatsforstverwaltung, Nr. 49, 169 S.
- Pretzsch, H., Utschig, H., Bachmann, M., 2001: **Innovation durch Kontinuität – das Ertragskundliche Versuchswesen in Bayern**. Mitteilungen aus der Bayerischen Staatsforstverwaltung. 51 (2), S. 425-443
- Pretzsch, H., Biber, P., Dursky, J., von Gadow, K., Hasenauer, H., Kändler, G., Kenk, G., Kublin, E., Nagel, J., Pukkala, T., Skovsgaard, J. P., Sotke, R., Sterba, H., 2006: **Standardizing and categorizing growth models**. In Hasenauer, H. (Ed.): Sustainable Forest Management. Growth Models for Europe. Springer Verlag. Berlin, Heidelberg, New York. 398., hier: p. 39-58.
- Prietzl, J., Kölling, Ch., 1998: **Trajectory analysis of long-term changes in the nutritional status of a Scots Pine stand**, Forstwiss. Centralblatt. 101, S. 388-407.

- Prietzl, J., Kolb, E., Rehfuss, K.E., 1997: **Langzeituntersuchungen ehemals streugennutzter Kiefernökosysteme in der Oberpfalz: Veränderungen von bodenchemischen Eigenschaften und der Nährstoffversorgung der Bestände**, Forstwiss. Cbl. 116, S. 269-290
- PROGNOSIS, 2006: **PROGNOSIS^{BC} –Homepage**, <http://www.for.gov.bc.ca/hre/gymodels/progbc/index.htm> (Abruf vom 14.06.2006)
- Quinlan, J. R., 1986: **Induction of decision trees**, Machine Learning 1: p. 81-106.
- Quinlan, J. R., 1993: **C 4.5 Programs for machine learning**. Morgan Kaufman Publisher, 302 p.
- R Development Core Team, 2006: **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Rajanowa, W., 1997: **Zur Verwaltung und Exploration waldökologischer Daten**. (unveröffentlichte) Dissertation der Fakultät für Forst-, Geo- und Hydrowissenschaften der Technischen Universität Dresden. 113 S.
- Reed, D., Tome, M., Araujo, M.C., Jones, E., 2001: **A Re-examination of Potential Modifier Dimensional Growth Models**. 11 p. <http://cms1.gre.ac.uk/conferences/iufro/proceedings/Reed2potmod.pdf> (Abruf vom 11.06.2006)
- Rehfuss, K.-E., 1990: **Waldböden – Entwicklung, Eigenschaften und Nutzung**. Verlag Paul Parey, Hamburg und Berlin, 2. Auflage, 294 S.
- Richter, M., (1998): **Online-Befragung als neues Instrument zur Beurteilung der Benutzerfreundlichkeit interaktiver Software am Beispiel einer Internet-Anwendung**. Unveröffentlichte Lizentiatsarbeit, Physiologisches Institut der Universität Zürich, 112 S.
- Ritchie, M. W., 1999: **A Compendium of Forest Growth and Yield Simulators for the Pacific Coast States**. USDA Forest Service (eds.). General Technical Report PSW-GTR-174, 68 p.

- Röhe, P., 1985: **Untersuchungen über das Wachstum der Buche in Baden-Württemberg**. Schriftenreihe der Landesforstverwaltung Baden-Württemberg, Nr. 61, 113 S.
- Röhle, H., 1995: **Zum Wachstum der Fichte auf Hochleistungsstandorten in Südbayern. Ertragskundliche Auswertung langfristig beobachteter Versuchsreihen unter besonderer Berücksichtigung von Trendänderungen im Wuchsverhalten**. Habilitationsschrift, Forstwissenschaftliche Fakultät, Universität München, 249 S.
- Rudolph, A., 2006: **Data Mining Verfahren**. (unveröffentlichtes) Skript der Universität der Bundeswehr München. 276 S.
(<http://www.bwi.unibw-muenchen.de/home/rudolph/dm.html>)
- Savnik, I., Lausen, G., Kahle, H.-P., Spiecker, H., Hein, S., 2000: **Algorithm for matching sets of time series**. In: D.A. Zighed, J. Komorowsky and J. Zytkov (Eds.): PKDD 2000, LNAI 1910. S. 277-288.
- Schlittgen, R., Streitberg, B. H. J., 2001: **Zeitreihenanalyse**. R. Oldenburg-Verlag, München, Wien, 571 S.
- Schmidt, A., 1971: **Wachstum und Ertrag der Kiefer auf wirtschaftlich wichtigen Standorteinheiten in der Oberpfalz**. Forstliche Forschungsberichte München, 187 S.
- Schmidt, M., Böckmann, Th., Nagel, J., 2006: **The use of tree models for silvicultural decision making**. In Hasenauer, H. (Ed.): Sustainable Forest Management. Growth Models for Europe. Springer Verlag. Berlin, Heidelberg, New York. 398 S.
- Schultze, B., Kölling, Ch., Dittmar, Ch., Rötzer, Th., Elling, W., 2005: **Konzept für ein neues quantitatives Verfahren zur Kennzeichnung des Wasserhaushalts von Waldböden in Bayern: Modellierung – Regression – Regionalisierung**. Forstarchiv 76 (2005), S. 155-163.
- Seifert, E., Seifert, Th., Biber, P., Pommerening, A., Dursky, J., Pretzsch, H., 1999: **SILVA 2.2 Benutzerhandbuch**. 119 S. (unveröffentlicht) (liegt jeder SILVA 2.2 Installationsversion in digitaler Form als PDF-Dokument bei).

- Seifert, St., Seifert, E., Biber, P., Seifert, Th., Klemmt, H.-J., Pretzsch, H., 2005: **SILVA 3 – Eine modulare Softwarelösung**. AFZ – Der Wald 21/ 2005, S. 1144-1145.
- Sloboda, 1971: **Zur Darstellung von Wachstumsprozessen mit Hilfe von Differentialgleichungen erster Ordnung**, Mitteilungen der Baden-Württembergischen Forstlichen Versuchsanstalt, 32, 87 S.
- Spellmann, H., 2003: **Sicherung einer nachhaltigen Waldentwicklung auf überbetrieblichen Ebenen**. Forstwiss. Cbl., 122 Jg., S. 250-257.
- Spelsberg, G., 1994: **Zum Höhenwachstum der Fichte in Nordrhein-Westfalen**. Allg. Forst- u. J.-Ztg., 165. Jg., 4, S. 77-80.
- Statistisches Bundesamt, 2006: **Betriebsgrößenstruktur von Forstbetrieben**. <http://www.destatis/basis/d/forst/forsttab26.php> (Abruf vom 24.10.2006).
- Sterba, H., 2002: **Forest inventories and growth models to examine management strategies for forests in transition**. Forestry, Vol. 75, No. 4, p. 411-418.
- Sterba, H., Monserud, R.A., 1996: **Applicability of the forest stand growth simulator PROGNAUS for the Austrian part of the Bohemian Massif**, Ecological Modelling 98 (1997), p. 23-34.
- Stoyan, D., Stoyan, H. und Jansen, U., 1997: **Umweltstatistik. Statistische Verarbeitung und Analyse von Umweltdaten**. B. G. Teubner Verlagsgesellschaft, Stuttgart, Leipzig, 348 S.
- Tillmans, Ch., 2003: **Data Mining zur Unterstützung betrieblicher Entscheidungen**. Dissertation der Universität Dortmund, 371 S. [<http://deposit.ddb.de/cgi-bin/dokserv?idn=969324456>] (Abruf vom 19.06.2006).
- Uhl, E., Metzger, H.-G., Seifert, Th., 2006: **Dimension und Wachstum von solitären Buchen und Eichen**. Tagungsbericht zur Jahrestagung 2006 der Sektion Ertragskunde im Deutschen Verband der Forstlichen Versuchs- und Forschungsanstalten in Stauffen. S. 47-54. (<http://www.nw->

- fva.de/~nagel/SektionErtragskunde/band2006/Tag2006_06.pdf ,
Abruf vom 03.10.2006)
- van de Weyer, D., 2004 : **Online-Fragebogen**. c't (Magazin für Computertechnik) 16/
2004, heise Verlag. S. 62.
- Vanclay, J., 1994: **Modelling Forest Growth and Yield**. CABI Publishing,
Wallingford, UK, 312 p.
- v. Gadow, K., 2003: **Waldstruktur und Wachstum**, Institut für Waldinventur und
Waldwachstum (Hrsg.), Universitätsdrucke Göttingen, 2003, 246 S.
- v. Mammen, A., Bachmann, M., Prietzel, J., Pretzsch, H., Rehfuss, K.E., 2003:
**Bodenzustand, Ernährungszustand und Wachstum von Fichten
(Picea abies KARST) auf Probeflächen des Friedenfelser
Verfahrens in der Oberpfalz**. Forstwiss. Centralblatt. 122. S. 99-
114.
- v. Teuffel, K., Hein, S., Kotar, M., Preuhsler, E. P., Puumalainen, J., Weinfurter, P.,
2006: **End User Needs and Requirements**. In Hasenauer, H. (Ed.):
Sustainable Forest Management. Growth Models for Europe.
Springer Verlag. Berlin, Heidelberg, New York, 398 p., hier: p. 19-
38.
- WIKIPEDIA, 2004a: Kalibrierung.
<http://de.wikipedia.org/wiki/Kalibrieren> (Abruf vom 25.10.2004)
- WIKIPEDIA, 2004b: Justierung.
<http://de.wikipedia.org/wiki/Justieren> (Abruf vom 25.10.2004)
- WIKIPEDIA, 2004c: Calibration
<http://en.wikipedia.org/wiki/Calibration> (Abruf vom 25.10.2004)
- WIKIPEDIA, 2004d: Calibration in statistics.
http://en.wikipedia.org/wiki/Calibration_%28statistics%29 (Abruf
vom 25.10.2004)
- WIKIPEDIA, 2006e: Befragung.
<http://de.wikipedia.org/wiki/Befragung> (Abruf vom 17.10.2006)

-
- Wilde, K. D., 2001: **Data Warehouse, OLAP und Data Mining im Marketing –** Moderne Informationstechnologien im Zusammenspiel. In: Handbuch Data Mining im Marketing, Vieweg Verlag, S. 1-13.
- Wilde, K. D., 2005: **Data Mining im Marketing**, (unveröffentlichtes) Vorlesungsskript des Lehrstuhls für Allgemeine Betriebswirtschaftslehre und Wirtschaftsinformatik der Katholischen Universität Eichstätt-Ingolstadt, 230 S.
- Witten, I., Frank, E., 2005: **Data Mining – Practical machine learning tools and techniques**. 2nd. Ed. Morgan Kaufmann Publishers (an imprints of Elsevier), 524 p.
- Wenk, G., Vaidotas, A, Smelko, S., 1990: **Waldertragslehre**. Deutscher Landwirtschaftsverlag. 448 S.
- Zeide, B., 1993: **Analysis of Growth Equations**. Forest Science, Vol. 39, No. 3, p. 594-616.
- Zöhner, F., 1980: **Forstinventur. Ein Leitfaden für Studium und Praxis**. Verlag Paul Parey, Hamburg, Berlin, 207 S.

9. Anhang

9.1 Tabelle: Staatswaldfläche Bayerns nach Wuchsregionen

Wuchs- gebiet	Wuchs- bezirk	Teil- wuchs- bezirk	Name	Staatswald (Fläche in ha)
1	1		Unterrainebene	5,26
2	1		Grundgebirgsspessart	1416,43
2	2	1	Buntsandsteinspessart/ Hochspessart	13228,48
2	2	2	Buntsandsteinspessart/ Nordspessart	20409,36
2	2	3	Buntsandsteinspessart/ Mainspessart	4268,43
2	3		Bayerischer Odenwald	0
3	1		Vorrhön	6471,61
3	1	1	Saale-und Sinn-Vorrhön	5284,85
3	1	2	Östliche Vorrhön	8046,53
3	2		Hohe Rhön	1907,61
4	1		Nördliche Fränkische Platte	5018,98
4	1	1	Nördliche Gipskeuperplatte und Grabfeld	552,42
4	2		Südliche Fränkische Platte	8368,37
4	2	1	Südliche Gipskeuperplatte	323,8
4	2	2	Kitzinger Sandgebiet	975,82
5	1		Haßberge	10365,04
5	2		Steigerwald	17921,02
5	2	1	Südlicher Steigerwald	638,93
5	3		Frankenhöhe	9958,77
5	3	1	Rangau-Hochfläche	3301,74
5	4		Itz-Baunach-Hügelland	8351,03
5	5		Nördliche Keuperabdachung	7195,99
5	5	1	Regnitzsenke	1676,68
5	6		Südliche Keuperabdachung	9485,71
5	6	1	Rezat-Rednitzsenke	18047,83
5	7		Nördliches Albvorland	3456,62
5	8		Südliches Albvorland	18727,63
5	9		Ries	70,32
6	1		Nördliche Frankenalb und Nördlicher Oberpfälzer Jura	17653,24
6	2		Südliche Frankenalb und Südlicher Oberpfälzer Jura	38203,92
6	2	1	Ingolstädter Donaualb	1336,67
6	3		Schwäbische Riesalb	4530,74
6	3	1	Egualb	5739,36
6	4		Oberfränkisches Braunjuragebiet	4920,21
6	5		Oberpfälzer Jurarand	11392,03
7	1		Bruchschollenland	4423,94
7	2		Obermainhügelland	10256,08
7	3		Stedtlinger Gebiet	707,22

Fortsetzung: Tabelle: Staatswaldfläche Bayerns nach Wuchsregionen

Wuchs- gebiet	Wuchs- bezirk	Teil- wuchs- bezirk	Name	Staatswald (Fläche in ha)
8	1		Frankenwald	20417,29
8	2		Münchberger Sattel	236,66
8	3		Fichtelgebirge	21118,37
8	4		Brand-Neusorger-Becken	1797,52
8	5		Steinwald	2842,11
8	6		Bayerisches Vogtland	2771,52
8	7		Selb-Wunsiedler-Bucht	6102,01
9	1		Oberpfälzer Becken- und Hügelland	29880,73
9	1	1	Hessenreuther Wald	821,65
10	1		Mitterteicher Basaltgebiet	3205,07
10	2	1	Waldsassener Schiefergebiet	6417,44
10	2	2	Wiesauer Senke	3520,04
10	3		Vorderer Oberpfälzer Wald	11358
10	4		Innerer Oberpfälzer Wald	17928,24
10	5		Cham-Further-Senke	87,2
11	1		Westlicher Vorderer Bayerischer Wald	668,86
11	1	1	Vorwaldrand	2074,85
11	2		Östlicher Vorderer Bayerischer Wald	9546,95
11	2	1	Lallinger Winkel	144,63
11	2	2	Ilzvorland	944,03
11	3		Innerer Bayerischer Wald	19388,68
12	1		Donauried	2589,43
12	2	1	Ingolstädter Donauniederung/ Donauau	77,16
12	2	2	Ingolstädter Donauniederung/ Donaumoos	846,78
12	3	1	Ostbayerische Donauniederung/ Donauau	78,42
12	3	2	Ostbayerische Donauniederung/ Gäuland	0
12	4		Unteres Lechtal	21,57
12	5		Unteres Isartal	992,85
12	6		Unteres Inntal	1080,29
12	7		Mittelschwäbisches Schotterridel- und Hügelland	22459,85
12	7	1	Biburger Hügelland	3540,12
12	7	2	Illerau	493,98
12	7	3	Donauschwäbisches Hügelland	2399,18
12	8		Oberbayerisches Tertiärhügelland	7942,37
12	8	1	Aindlinger Terrassentreppe	2837,58
12	9	1	Westliches Niederbayerisches Tertiärhügelland	6013,24
12	9	2	Östliches Niederbayerisches Tertiärhügelland	2510,07
12	9	3	Neuburger Wald	3550,04
13	1		Lechfeld	77,49
13	2	1	Südliche Münchner Schotterebene	22514,34
13	2	2	Nördliche Münchner Schotterebene	1127,53
13	2	3	Moose und Auen nördlich Münchens	3406,62
13	3		Mühldorfer und Öttinger Schotterfelder	5401,83
13	4		Vorallgäu	10710,38
13	5		Landberger Altmoräne	844,48
13	6		Isener Altmoräne und Hochterasse	2002,04
13	7		Trostberger Altmoräne und Hochterasse	303,52

Fortsetzung: Tabelle: Staatswaldfläche Bayerns nach Wuchsregionen

Wuchs- gebiet	Wuchs- bezirk	Teil- wuchs- bezirk	Name	Staatswald (Fläche in ha)
14	1		Bayerische Bodenseelandschaft	338,06
14	2		Westallgäuer Hügelland	30,73
14	3		Schwäbische Jungmoräne und Molassevorberge	6659,08
14	4	1	Westliche kalkalpine Jungmoräne	18929,93
14	4	2	Inn-Jungmoräne	5897,63
14	4	3	Östliche Kalkalpine Jungmoräne	4280,78
15	1		Kürnacher Molassebergland	3212,58
15	2		Allgäuer Molassevoralpen	2118,46
15	3		Allgäuer Flysch- und Helvetikumvoralpen	4172,45
15	4	1	Ammergauer Flyschberge	3473,22
15	4	2	Tegernseer Flyschberge	1993,12
15	4	3	Teisendorfer Flyschberge	5109,5
15	5		Mittlere Bayerische Kalkalpen	77034,22
15	6		Chiemgauer Alpen und Saalforstamt St. Martin	43241,13
15	7		Allgäuer Hochalpen	2793,84
15	8		Karwendel und Wettersteinmassiv	18762,51
15	9		Berchtesgadener Hochalpen und Saalforstamt St. Martin	25085,42
15	9	1	Leoganger Schiefergebirge	2284,72
SUMME				775149,06

9.2 Tabelle: Aufstellung der Forsteinrichtungen, deren Datensätze für diese Arbeit Verwendung finden

fe_nr	fe_osfoa	Forstamt	p/ t	Daten in FEDB	Lage PK zu FBK	Verwendung
1123	307	Fichtelberg	p	nicht vorhanden	---	nein
1122	411	Freising (Zugang aus München)	t	vorhanden	nicht deckungsgleich	nein
1121	621	Mittelsinn	p	nicht vorhanden	---	nein
1120	626	Schoellkrippen	p	nicht vorhanden	---	nein
1119	525	Pfreimdt	--	nicht vorhanden	---	nein
1118	530	Siegenburg	--	nicht vorhanden	---	nein
1117	402	Altoetting	p/ t	nicht vorhanden	---	nein
1116	413	Garmisch-Partenkirchen	t	nicht vorhanden	---	nein
1115	214	Ottobeuren	t	nicht vorhanden	---	nein
1114	435	Wolfratshausen, Zugang Starnberg	t	nicht vorhanden	---	nein
1113	404	Bad Toelz, Zugang Wolfratshausen	t	nicht vorhanden	---	nein
1112	435	Wolfratshausen, Zugang Weilheim	t	nicht vorhanden	---	nein
1111	417	Landsberg a. Lech, Zugang Weilheim	t	nicht vorhanden	---	nein
1110	617	Kleinwallstadt (Begang ohne Inventur)	--	nicht vorhanden	---	nein
1109	113	Nürnberg (Zugang aus Erlangen)	--	nicht vorhanden	---	nein
1108	445	Nationalpark Berchtesgaden (nur Begang)	--	nicht vorhanden	---	nein
1107	617	Kleinwallstadt, ET-Schätzung und Relaskopierung	--	nicht vorhanden	---	nein
1106	321	Weißenstein	p	vorhanden	deckungsgleich	nein
1105	511	Koetzing	t	vorhanden	---	nein
1104	507	Freyung	p	vorhanden	---	nein
1103	521	Passau	t	vorhanden	deckungsgleich m.E.	ja
1102	531	Simbach am Inn	t	vorhanden	deckungsgleich m.E.	ja
1101	421	München	t	vorhanden	deckungsgleich	ja
1100	404	Bad Toelz	t	vorhanden	---	nein
1099	401	Anzing	t	vorhanden	deckungsgleich	ja
1098	612	Gemünden am Main	p	vorhanden	---	nein
1097	607	Marktheidenfeld	t	vorhanden	---	---
1096	513	Landshut	--	nicht vorhanden	---	---
1095	504	Burglengenfeld	p	vorhanden	---	---
1094	618	Lohr am Main	t	vorhanden	---	---
1093	319	Selb (Zugang aus Rehau)	p	vorhanden	---	---
1092	432	Starnberg	t	vorhanden	deckungsgleich	ja
1091	428	Schongau, Zugang Weilheim	t	vorhanden	---	---
1090	435	Wolfratshausen	t	vorhanden	deckungsgleich	ja
1089	601	Altenbuch	p	vorhanden	---	---
1088	---	Schälschadensinventur Weissenstadt	--	---	---	---
1087	---	Schälschadensinventur Fichtelberg	--	---	---	---

Fortsetzung: Tabelle: Aufstellung der Forsteinrichtungen, deren Datensätze für diese Arbeit Verwendung finden

fe_nr	fe_osfoa	Forstamt	p/ t	Daten in FEDB	Lage PK zu FBK	Verwendung
1086	541	Nationalpark Bayerischer Wald	p	vorhanden	---	---
1084	109	Heideck, Zugang Schwabach	p	vorhanden	---	---
1083	615	Heigenbrücken	p	vorhanden	---	---
1082	515	Mitterfels	t	vorhanden	---	---
1081	501	Amberg	p	vorhanden	---	---
1080	508	Bad Griesbach	p	vorhanden	deckungsgleich	ja
1079	406	Berchtesgaden	t	vorhanden	---	---
1078	319	Selb	t	vorhanden	---	---
1077	308	Forchheim	p	vorhanden	---	---
1076	210	Kaufbeuren	t	vorhanden	deckungsgleich	ja
1075	203	Dillingen	p	vorhanden	deckungsgleich	ja
1074	118	Uffenheim	p	vorhanden	deckungsgleich	nein
1073	105	Erlangen	p	vorhanden	deckungsgleich	ja
1072	---	Test-FE (?)	--	---	---	---
1071	101	Allersberg	p	vorhanden	deckungsgleich	ja
1070	304	Betzenstein	t	vorhanden	deckungsgleich	ja
1069	315	Pegnitz	p	vorhanden	deckungsgleich	nein
1068	520	Parsberg	t	vorhanden	deckungsgleich	ja
1067	527	Riedenburg	p	vorhanden	deckungsgleich	ja
1066	509	Kelheim	p	vorhanden	nicht deckungsgleich	nein
1065	418	Marquartstein	p	vorhanden	---	nein
1064	411	Freising	p	vorhanden	deckungsgleich	ja
1063	209	Kaisheim	t	vorhanden	deckungsgleich	ja
1062	212	Mindelheim	t	vorhanden	deckungsgleich	ja
1061	625	Rothenbuch	p	vorhanden	deckungsgleich	ja
1060	532	Sulzbach-Rosenberg (Zugang Hersbruck)	p	vorhanden	deckungsgleich	ja
1059	102	Altdorf (Zugang Hersbruck)	p	vorhanden	deckungsgleich	nein
1058	402	Altoetting (Zugang Mühldorf)	t	vorhanden	deckungsgleich	nein
1057	305	Burgebrach	t	vorhanden	nicht deckungsgleich	nein
1056	318	Scheßlitz	p	vorhanden	deckungsgleich m.E.	ja
1055	523	Pielenhofen	t	vorhanden	deckungsgleich	ja
1054	534	Tirschenreuth	p	vorhanden	---	nein
1053	434	Wasserburg	p	vorhanden	deckungsgleich	ja
1052	410	Schliersee	p	vorhanden	deckungsgleich	nein
1051	606	Neustadt/ Saale	t	vorhanden	deckungsleich	ja
1050	605	Bad Kissingen	p	vorhanden	deckungsgleich	ja
1049	217	Zusmarshausen, Ankauf Stauffenberg	p	vorhanden	nicht deckungsgleich	nein
1048	109	Heideck	p	vorhanden	deckungsgleich	ja
1047	211	Krumbach	p	vorhanden	deckungsgleich	ja
1046	423	Oberammergau (Teil Ebenwald)	p	vorhanden	---	nein
1045	426	Sauerlach	t	vorhanden	deckungsgleich	ja
1044	423	Oberammergau (Zugang Murnau)	t	vorhanden	---	nein
1043	528	Roding	p	vorhanden	nicht deckungsgleich	nein

Fortsetzung: Tabelle: Aufstellung der Forsteinrichtungen, deren Datensätze für diese Arbeit Verwendung finden

fe_nr	fe_osfoa	Forstamt	p/ t	Daten in FEDB	Lage PK zu FBK	Verwendung
1041	417	Landsberg a. Lech	p	vorhanden	deckungsgleich	ja
1042	519	Neureichenau	p	vorhanden	deckungsgleich	ja
1040	111	Hersbruck	p	vorhanden	deckungsgleich	ja
1039	429	Neuburg a.d. Donau	t	vorhanden	deckungsgleich	ja
1038	628	Steinach	p	vorhanden	deckungsgleich	ja
1037	433	Traunstein	t	vorhanden	deckungsgleich	ja
1036	661	Juliuspital; Tausch mit Hammelburg	p	---	---	nein
1035	314	Nordhalben; (Zugang Kronach)	t	---	---	nein
1034	306	Coburg	t	nicht vorhanden	---	---
1033	306	Coburg	p	vorhanden	nicht deckungsgleich	nein
1032	202	Biburg	t	vorhanden	deckungsgleich	ja
1031	204	Aichach	p	vorhanden	deckungsgleich	ja
1030	312	Lichtenfels (Zugang Coburg)	t	---	---	nein
1029	537	Waldmünchen	t	vorhanden	deckungsgleich m.E.	nein
1028	412	Fürstenfeldbruck	t	vorhanden	deckungsgleich	ja
1027	310	Hollfeld	t	vorhanden	deckungsgleich	ja
1026	428	Schongau	t	vorhanden	deckungsgleich	ja
1025	506	Flossenbürg	p	vorhanden	nicht deckungsgleich	nein
1024	303	Bayreuth (Zugang Fichtelberg)	p	---	---	---
1023	317	Rothenkirchen	p	vorhanden	deckungsgleich	ja
1022	206	Schwabmünchen	t	vorhanden	deckungsgleich	ja
1021	217	Zusmarshausen	p	vorhanden	deckungsgleich	ja
1020	533	Tännesberg	p	vorhanden	deckungsgleich m.E.	ja
1019	532	Sulzbach-Rosenberg	p	vorhanden	deckungsgleich	ja
1018	533	Tännesberg (doppelter Eintrag!!!)	---	---	---	nein
1017	117	Treuchtlingen	p	vorhanden	deckungsgleich m.E.	ja
1016	619	Mellrichstadt	t	vorhanden	deckungsgleich	ja
1015	627	Schweinfurt	t	vorhanden	nicht deckungsgleich	nein
1014	408	Eichstätt	p	vorhanden	deckungsgleich	ja
1013	415	Kipfenberg	t	vorhanden	deckungsgleich m.E.	ja
1012	323	Ebrach	p	vorhanden	nicht deckungsgleich	nein
1011	405	Beilngries	t	vorhanden	deckungsgleich m.E.	ja
1010	201	Kempten	p	vorhanden	deckungsgleich	ja
1009	312	Lichtenfels	p	vorhanden	nicht deckungsgleich	nein
1008	602	Arnstein	p	vorhanden	deckungsgleich	ja
1007	502	Bodenmais	p	vorhanden	deckungsgleich	ja
1006	900	Sonderinventur Hochlagen	---	---	---	nein
1005	107	Feuchtwangen	t	vorhanden	deckungsgleich	ja
1004	536	Vohenstrauß	t	vorhanden	nicht deckungsgleich	nein
1003	301	Bad Steben	t	vorhanden	deckungsgleich	ja
1002	614	Hammelburg	p	vorhanden	deckungsgleich	ja
1001	216	Weißenhorn	p	vorhanden	deckungsgleich m.E.	ja
1000	407	Ebersberg	p	vorhanden	deckungsgleich	ja

Fortsetzung: Tabelle: Aufstellung der Forsteinrichtungen, deren Datensätze für diese Arbeit Verwendung finden

fe_nr	fe_osfoa	Forstamt	p/ t	Daten in FEDB	Lage PK zu FBK	Verwendung
384	317	Rothenkirchen	t	vorhanden	---	nein
383	525	Regen	t	vorhanden	deckungsgleich m.E.	nein
382	525	Regen	p	vorhanden	deckungsgleich m.E.	nein
381	525	Regen	p	vorhanden	deckungsgleich m.E.	nein
380	525	Regen	---	nicht vorhanden	---	---
379	423	Oberammergau	t	vorhanden	nicht deckungsgleich	nein
378	423	Oberammergau	t	vorhanden	nicht deckungsgleich	nein
377	423	Oberammergau	t	vorhanden	nicht deckungsgleich	nein
376	113	Nürnberg	p	vorhanden	deckungsgleich m.E.	nein
375	110	Heilsbronn	t	vorhanden	deckungsgleich	nein
374	104	Dinkelsbühl	t	vorhanden	deckungsgleich	nein
373	102	Altdorf b. Nürnberg	p	vorhanden	deckungsgleich	nein
372	610	Ebrach	---	---	---	---
371	517	Neumarkt i.d. Opf.	p	vorhanden	deckungsgleich	ja
370	630	Würzburg	t	vorhanden	deckungsgleich	ja
369	629	Wiesentheid	t	vorhanden	deckungsgleich m.E.	nein
368	604	Bad Brückenau	t	vorhanden	deckungsgleich	ja
367	530	Siegenburg	p	vorhanden	deckungsgleich	ja
366	510	Kemnath	p	vorhanden	nicht deckungsgleich	nein
365	505	Deggendorf	p	vorhanden	deckungsgleich m.E.	nein
364	421	München	t	vorhanden	nicht deckungsgleich	nein
363	419	Mittenwald	t	vorhanden	deckungsgleich	nein
362	402	Altötting	t	vorhanden	nicht deckungsgleich	nein
361	317	Rothenkirchen	p	vorhanden	deckungsgleich	ja
360	305	Burgebrach	t	vorhanden	deckungsgleich	ja
359	214	Ottobeuren	t	vorhanden	deckungsgleich	ja
358	207	Illertissen	t	vorhanden	deckungsgleich	ja
357	610	Ebrach	p	vorhanden	deckungsgleich	nein
356	610	Ebrach	---	nicht vorhanden	---	---
355	622	Münnerstadt	t	vorhanden	deckungsgleich	ja
354	616	Gerolzhofen	t	vorhanden	deckungsgleich	ja
353	609	Ebern	t	vorhanden	deckungsgleich	ja
352	539	Weiden i.d. Opf.	p	vorhanden	deckungsgleich	ja
351	529	Schnaittenbach	p	vorhanden	deckungsgleich	ja
350	524	Pressath	p	vorhanden	nicht deckungsgleich	nein
349	514	Mauth	p	vorhanden	nicht deckungsgleich	nein
348	424	Rosenheim	t	vorhanden	nicht deckungsgleich	nein
347	414	Geisenfeld	t	vorhanden	nicht deckungsgleich	nein
346	401	Anzing	t	vorhanden	deckungsgleich m.E.	nein
345	318	Scheßlitz	t	vorhanden	deckungsgleich	ja
344	314	Nordhalben	t	vorhanden	deckungsgleich	ja
343	215	Sonthofen	t	vorhanden	deckungsgleich m.E.	nein
342	112	Neustadt a.d. Aisch	t	vorhanden	deckungsgleich	ja
341	611	Eltmann	p	vorhanden	deckungsgleich	ja
340	608	Bad Königshofen	t	vorhanden	deckungsgleich	ja
339	541	Nationalpark Bayerischer Wald	p	vorhanden	---	nein
338	531	Simbach am Inn	---	nicht vorhanden	---	---
337	524	Pressath	p	vorhanden	deckungsgleich	nein

Fortsetzung: Tabelle: Aufstellung der Forsteinrichtungen, deren Datensätze für diese Arbeit Verwendung finden

fe_nr	fe_osfoa	Forstamt	p/ t	Daten in FEDB	Lage PK zu FBK	Verwendung
336	522	Pfreimdt	p	vorhanden	nicht deckungsgleich	nein
335	521	Passau	p	vorhanden	deckungsgleich	ja
334	513	Landshut	p	vorhanden	nicht deckungsgleich	nein
333	503	Bodenwöhr	p	vorhanden	nicht deckungsgleich	---
332	436	Saalforstamt Leogang	t	---	---	nein
331	409	Saalforstamt Fall	t	---	---	nein
330	320	Stadtsteinach	t	vorhanden	nicht deckungsgleich	nein
329	307	Fichtelberg	t	vorhanden	deckungsgleich	nein
328	208	Immenstadt i. Allgäu	t	vorhanden	nicht deckungsgleich	nein
327	205	Füssen	t	vorhanden	deckungsgleich	ja
326	105	Erlangen	p	vorhanden	deckungsgleich	nein
325	103	Ansbach	t	vorhanden	---	nein
324	621	Mittelsinn	t	vorhanden	---	nein
323	615	Heigenbrücken	p	vorhanden	---	nein
322	538	Waldsassen	p	vorhanden	deckungsgleich	ja
321	516	Mitterteich	p	vorhanden	deckungsgleich	ja
320	515	Mitterfels	p	vorhanden	---	nein
319	511	Kötzting	p	vorhanden	---	nein
318	507	Freyung	p	vorhanden	---	nein
317	435	Wolfratshausen	t	vorhanden	deckungsgleich m.E.	nein
316	432	Starnberg	t	vorhanden	deckungsgleich	ja
315	431	Siegsdorf	t	vorhanden	deckungsgleich	nein
314	403	Bad Reichenhall	t	vorhanden	deckungsgleich	nein
313	210	Kaufbeuren	t	vorhanden	deckungsgleich	ja
312	203	Dillingen	p	vorhanden	deckungsgleich	ja
311	114	Rothenburg o.d. Tauber	p	vorhanden	deckungsgleich	ja
310	102	Altdorf b. Nürnberg	p	vorhanden	nicht deckungsgleich	nein
309	625	Rothenbuch	p	vorhanden	deckungsgleich	ja
308	612	Gemünden am Main	t	vorhanden	---	nein
307	607	Marktheidenfeld	t	vorhanden	---	nein
306	535	Viechtach	---	---	---	---
305	522	Pfreimdt	p	vorhanden	nicht deckungsgleich	nein
304	520	Parsberg	t	vorhanden	deckungsgleich	nein
303	518	Neunburg v. Wald	p	vorhanden	---	nein
302	509	Kelheim	p	vorhanden	deckungsgleich m.E.	ja
301	508	Bad Griesbach	p	vorhanden	nicht deckungsgleich	nein
300	430	Weilheim in Oberbayern	t	vorhanden	deckungsgleich m.E.	nein
299	425	Ruhpolding	p	vorhanden	---	nein
298	411	Freising	t	vorhanden	deckungsgleich	ja
297	322	Wunsiedel	p	vorhanden	---	nein
296	309	Goldkronach	t	vorhanden	---	nein
295	213	Monheim	t	vorhanden	deckungsgleich	ja
294	212	Mindelheim	t	vorhanden	deckungsgleich	ja
293	209	Kaisheim	t	vorhanden	deckungsgleich m.E.	ja
292	624	Rohrbrunn	p	vorhanden	deckungsgleich m.E.	ja
291	618	Lohr am Main	t	vorhanden	---	nein
290	618	---	---	---	---	---

Fortsetzung: Tabelle: Aufstellung der Forsteinrichtungen, deren Datensätze für diese Arbeit Verwendung finden

fe_nr	fe_osfoa	Forstamt	p/ t	Daten in FEDB	Lage PK zu FBK	Verwendung
289	601	Altenbuch	t	vorhanden	---	nein
288	534	Tirschenreuth	t	vorhanden	deckungsgleich m.E.	nein
287	526	Pielenhofen	t	vorhanden	deckungsgleich m.E.	ja
286	523	Pielenhofen	t	vorhanden	deckungsgleich	ja
285	416	Kreuth	p	vorhanden	---	nein
284	413	Garmisch-Partenkirchen	t	vorhanden	---	nein
283	316	Rehau	p	vorhanden	---	nein
282	211	Krumbach (Schwabern)	p	vorhanden	deckungsgleich	ja
281	113	Nürnberg	p	vorhanden	nicht deckungsgleich	nein
280	627	Schweinfurt	t	vorhanden	nicht deckungsgleich	nein
279	626	Schöllkrippen	p	vorhanden	---	nein
278	527	Riedenburg	p	vorhanden	deckungsgleich	ja
277	504	Burglengenfeld	t	vorhanden	---	nein
276	501	Amberg	p	vorhanden	---	nein
275	434	Wasserburg a. Inn	t	vorhanden	deckungsgleich	ja
274	433	Traunstein	t	vorhanden	deckungsgleich	ja
273	422	Murnau a. Staffelsee	t	vorhanden	deckungsgleich	nein
272	420	Mühdorf a. Inn	t	vorhanden	deckungsgleich	nein
271	417	Landsberg a. Lech	p	vorhanden	deckungsgleich	nein
270	404	Bad Tölz	t	vorhanden	---	nein
269	321	Weißensstadt	p	vorhanden	---	nein
268	319	Selb	t	vorhanden	---	nein
267	204	Aichach	p	vorhanden	deckungsgleich	ja
266	202	Biburg	t	vorhanden	deckungsgleich m.E.	nein
265	108	Gunzenhausen	p	vorhanden	---	nein
264	623	Partenstein	p	vorhanden	---	nein
263	537	Waldmünchen	t	vorhanden	deckungsgleich	ja
262	528	Roding	p	vorhanden	deckungsgleich m.E.	nein
261	506	Flossenbürg	p	vorhanden	nicht deckungsgleich	nein
260	426	Sauerlach	t	vorhanden	nicht deckungsgleich	nein
259	412	Fürstenfeldbruck	t	vorhanden	nicht deckungsgleich	nein
258	408	Eichstätt	t	vorhanden	nicht deckungsgleich	nein
257	402	Altötting	t	vorhanden	nicht deckungsgleich	nein
256	401	Anzing	t	vorhanden	deckungsgleich	ja
255	312	Lichtenfels	p	vorhanden	deckungsgleich	nein
254	310	Hollfeld	t	vorhanden	nicht deckungsgleich	nein
253	217	Zusmarshausen	p	vorhanden	nicht deckungsgleich	nein
252	206	Schwabmünchen	t	vorhanden	deckungsgleich	ja
251	118	Uffenheim	p	vorhanden	deckungsgleich m.E.	nein
250	116	Schwabach	t	vorhanden	---	nein
249	115	Höchstadt a.d. Aisch	t	vorhanden	---	nein
248	628	Steinach	p	vorhanden	deckungsgleich	ja
247	619	Mellrichstadt	t	vorhanden	nicht deckungsgleich	nein
246	536	Vohenstrauß	t	vorhanden	deckungsgleich	nein
245	533	Tännesberg	p	vorhanden	deckungsgleich	ja
244	519	Neureichenau	p	vorhanden	deckungsgleich	ja
243	429	Neuburg a.d. Donau	t	vorhanden	deckungsgleich	nein
242	427	Schernfeld	p	vorhanden	---	nein
241	421	München	t	vorhanden	nicht deckungsgleich	nein

Fortsetzung: Tabelle: Aufstellung der Forsteinrichtungen, deren Datensätze für diese Arbeit Verwendung finden

fe_nr	fe_osfoa	Forstamt	p/ t	Daten in FEDB	Lage PK zu FBK	Verwendung
240	415	Kipfenberg	t	vorhanden	nicht deckungsgleich	nein
239	303	Bayreuth	p	vorhanden	nicht deckungsgleich	nein
238	214	Ottobeuren	t	vorhanden	deckungsgleich m.E.	nein
237	201	Kempten (Allgäu)	p	vorhanden	nicht deckungsgleich	nein
236	112	Neustadt a.d. Aisch	t	vorhanden	deckungsgleich	nein
235	101	Allersberg	p	vorhanden	deckungsgleich m.E.	nein
234	606	Bad Neustadt a.d. Saale	t	vorhanden	nicht deckungsgleich	nein
233	605	Bad Kissingen	p	vorhanden	deckungsgleich	ja
232	532	Sulzbach-Rosenberg	p	vorhanden	deckungsgleich	ja
231	525	Regen	t	vorhanden	deckungsgleich	ja
230	517	Neumarkt i.d. Opf.	t	vorhanden	deckungsgleich	ja
229	502	Bodenmais	p	vorhanden	deckungsgleich	ja
228	445	Nationalpark Berchtesgaden	t	vorhanden	---	nein
227	418	Marquartstein	t	vorhanden	deckungsgleich m.E.	nein
226	410	Schliersee	t	vorhanden	deckungsgleich m.E.	nein
225	407	Ebersberg	p	vorhanden	deckungsgleich	ja
224	406	Berchtesgaden	t	vorhanden	---	nein
223	405	Beilngries	t	vorhanden	deckungsgleich m.E.	nein
222	313	Neustadt b. Coburg	p	vorhanden	nicht deckungsgleich	nein
221	306	Coburg	t	vorhanden	nicht deckungsgleich	nein
220	205	Füssen	t	vorhanden	deckungsgleich	ja
219	117	Treuchtlingen	t	vorhanden	deckungsgleich	nein
218	109	Heideck	p	vorhanden	deckungsgleich	ja
217	614	Hammelburg	p	vorhanden	deckungsgleich	ja
216	602	Arnstein	p	vorhanden	deckungsgleich m.E.	nein
215	540	Zwiesel	p	vorhanden	---	nein
214	530	Siegenburg	t	vorhanden	deckungsgleich	ja
213	510	Kemnath	t	vorhanden	deckungsgleich	nein
212	308	Forchheim	p	vorhanden	---	nein
211	305	Burgebrach	t	vorhanden	deckungsgleich	ja
210	304	Betzenstein	t	vorhanden	deckungsgleich m.E.	ja
209	216	Weißenhorn	p	vorhanden	nicht deckungsgleich	nein
208	207	Illertissen	t	vorhanden	deckungsgleich	ja
207	111	Hersbruck	p	vorhanden	nicht deckungsgleich	nein
206	110	Heilsbronn	t	vorhanden	deckungsgleich	nein
205	630	Würzburg	t	vorhanden	deckungsgleich m.E.	ja
204	323	Ebrach	t	nicht vorhanden	---	nein
203	539	Weiden i.d. Opf.	t	vorhanden	---	nein
202	529	Schnaittenbach	t	vorhanden	---	nein
201	524	Pressath	t	vorhanden	---	nein
200	514	Mauth	t	vorhanden	deckungsgleich	ja
199	432	Starnberg	t	vorhanden	deckungsgleich	ja
198	424	Rosenheim	t	vorhanden	nicht deckungsgleich	nein
197	423	Oberammergau	t	vorhanden	deckungsgleich m.E.	nein
196	302	Bamberg	t	vorhanden	---	nein
195	215	Sonthofen	t	vorhanden	deckungsgleich m.E.	nein
194	114	Rothenburg o.d. Tauber	t	vorhanden	deckungsgleich m.E.	nein
193	103	Ansbach	t	vorhanden	---	nein

Fortsetzung: Tabelle: Aufstellung der Forsteinrichtungen, deren Datensätze für diese Arbeit Verwendung finden

fe_nr	fe_osfoa	Forstamt	p/ t	Daten in FEDB	Lage PK zu FBK	Verwendung
192	621	Mittelsinn	t	vorhanden	---	nein
191	604	Bad Brückenau	t	vorhanden	deckungsgleich	ja
190	538	Waldsassen	t	vorhanden	nicht deckungsgleich	nein
189	522	Preimdt	t	vorhanden	nicht deckungsgleich	nein
188	508	Bad Griesbach i. Rottal	t	vorhanden	deckungsgleich m.E.	ja
187	503	Bodenwöhr	t	vorhanden	deckungsgleich	nein
186	428	Schongau	t	vorhanden	deckungsgleich	nein
185	419	Mittenwald	t	vorhanden	deckungsgleich	nein
184	320	Stadtsteinach	t	vorhanden	deckungsgleich m.E.	nein
183	317	Rothenkirchen	t	vorhanden	deckungsgleich m.E.	nein
182	315	Pegnitz	t	vorhanden	deckungsgleich	nein
181	314	Nordhalben	t	vorhanden	deckungsgleich m.E.	ja
180	311	Kronach	t	vorhanden	deckungsgleich m.E.	nein
179	301	Bad Steben	t	vorhanden	deckungsgleich m.E.	ja
178	208	Immenstadt i. Allgäu	t	vorhanden	deckungsgleich m.E.	ja
177	203	Dillingen a. d. Donau	t	vorhanden	deckungsgleich m.E.	ja
176	107	Feuchtwangen	t	vorhanden	deckungsgleich	nein
175	104	Dinkelsbühl	t	vorhanden	deckungsgleich m.E.	nein
174	629	Wiesentheid	t	nicht vorhanden	---	nein
173	616	Gerolzhofen	t	vorhanden	deckungsgleich	ja
172	609	Ebern	t	vorhanden	deckungsgleich	ja
171	531	Simbach a. Inn	t	nicht vorhanden	---	---
170	521	Passau	t	vorhanden	deckungsgleich	ja
169	518	Neunburg v. Wald	t	vorhanden	---	nein
168	516	Mitterteich	t	vorhanden	---	nein
167	513	Landshut	t	vorhanden	deckungsgleich	ja
166	505	Deggendorf	t	vorhanden	---	nein
165	438	Bayer. Saalforstverwaltung-Unken	t	vorhanden	---	nein
164	437	Bayer. Saalforstverwaltung-St. Martin	t	vorhanden	---	nein
163	436	Bayerische Saalforstverwaltung Leogang	t	vorhanden	---	nein
162	435	Wolfratshausen	t	vorhanden	deckungsgleich m.E.	ja
161	424	Rosenheim	t	nicht vorhanden	---	---
160	414	Geisenfeld	t	vorhanden	deckungsgleich m.E.	ja
159	318	Scheßlitz	t	vorhanden	deckungsgleich m.E.	ja
158	210	Kaufbeuren	t	vorhanden	nicht deckungsgleich	nein
157	106	Feucht	t	vorhanden	---	nein
156	105	Erlangen	t	vorhanden	nicht deckungsgleich	nein
155	622	Münnerstadt	t	nicht vorhanden	---	---
154	611	Eltmann	t	vorhanden	deckungsgleich m.E.	ja
153	541	Nationalpark Bayer. Wald	t	vorhanden	---	nein
152	535	Viechtach	t	nicht vorhanden	---	---
151	520	Parsberg	t	vorhanden	deckungsgleich m.E.	ja
150	515	Mitterfels	t	vorhanden	---	nein
149	511	Kötzting	t	vorhanden	---	nein
148	507	Freyung	t	vorhanden	---	nein
147	430	Weilheim i. OB	t	vorhanden	deckungsgleich m.E.	nein

Fortsetzung: Tabelle: Aufstellung der Forsteinrichtungen, deren Datensätze für diese Arbeit Verwendung finden

fe_nr	fe_osfoa	Forstamt	p/ t	Daten in FEDB	Lage PK zu FBK	Verwendung
146	411	Freising	t	vorhanden	deckungsgleich m.E.	ja
145	403	Bad Reichenhall	t	vorhanden	nicht deckungsgleich	nein
144	309	Goldkronach	t	vorhanden	---	nein
143	307	Fichtelberg	t	vorhanden	---	nein
142	212	Mindelheim	t	vorhanden	deckungsgleich	ja
141	209	Kaisheim	t	vorhanden	deckungsgleich m.E.	ja
140	102	Altdorf b. Nürnberg	t	vorhanden	nicht deckungsgleich	nein
139	608	Bad Königshofen i. Grabfeld	t	vorhanden	deckungsgleich	ja
138	534	Tirschenreuth	t	vorhanden	deckungsgleich	nein
137	526	Pielenhofen	t	vorhanden	deckungsgleich m.E.	ja
136	523	Pielenhofen	t	vorhanden	deckungsgleich m.E.	ja
135	509	Kelheim	t	vorhanden	deckungsgleich m.E.	ja
134	409	Fall	t	vorhanden	---	nein
133	322	Wunsiedel	t	vorhanden	---	nein
132	321	Weißensstadt	t	vorhanden	---	nein
131	316	Rehau	t	vorhanden	---	nein
130	213	Monheim	t	vorhanden	nicht deckungsgleich	nein
129	211	Krumbach (Schwabern)	t	vorhanden	nicht deckungsgleich	nein
128	113	Nürnberg	t	vorhanden	nicht deckungsgleich	nein
127	627	Schweinfurt	t	nicht vorhanden	---	---
126	601	Altenbuch	t	nicht vorhanden	---	---
125	527	Riedenburg	t	vorhanden	nicht deckungsgleich	nein
124	504	Burglengenfeld	t	vorhanden	---	---
123	501	Amberg	t	vorhanden	---	---
122	433	Traunstein	---	nicht vorhanden	---	---
121	431	Siegsdorf	---	nicht vorhanden	---	---
120	422	Murnau a. Staffelsee	---	nicht vorhanden	---	---
119	417	Landsberg a. Lech	---	nicht vorhanden	---	---
118	319	Rehau	---	nicht vorhanden	---	---
117	204	Aichach	---	nicht vorhanden	---	---
116	202	Biburg	---	nicht vorhanden	---	---
115	108	Gunzenhausen	---	nicht vorhanden	---	---
114	624	Rohrbrunn	---	nicht vorhanden	---	---
113	617	Kleinwallstadt	---	nicht vorhanden	---	---
112	612	Gemünden am Main	---	nicht vorhanden	---	---
111	434	Wasserburg a. Inn	---	nicht vorhanden	---	---
110	425	Ruhpolding	---	nicht vorhanden	---	---
109	420	Mühdorf a. Inn	---	nicht vorhanden	---	---
108	413	Garmisch-Partenkirchen	---	nicht vorhanden	---	---
107	411	Freising	---	nicht vorhanden	---	---
106	618	Lohr am Main	---	nicht vorhanden	---	---
105	615	Gerolzhofen	---	nicht vorhanden	---	---
104	607	Marktheidenfeld	---	nicht vorhanden	---	---
103	404	Bad Tölz	---	nicht vorhanden	---	---
102	625	Rothenbuch	---	nicht vorhanden	---	---
101	416	Kreuth	---	nicht vorhanden	---	---
100	626	Schöllkrippen	---	nicht vorhanden	---	---

9.3 Tabellen zur numerischen Charakterisierung des verwendeten Datenmaterials

Wuchsgebiet 2 „Spessart-Odenwald“.

Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
2.1	Buche	252	3	204	98	0	78,5	35,7	15,85	0,3	38	24,23	7,97
	Lärche	117	12	189	75	1,5	63,5	35,47	12,78	3	43	25,32	7,36
2.2	Buche	13391	2	374	107	0	119,5	39,08	17,22	0,2	48	26,71	8,26
	Douglasie	1048	7	150	43	0	92,5	31,92	16,67	1,2	48	21,67	9,4
	Eiche	5498	2	538	137	0	115,5	37,89	21,03	0,2	42,5	24,44	8,68
	Fichte	3674	3	182	69	0	87,5	37,34	15,58	0,2	44,5	26,1	8,81
	Kiefer	849	5	181	82	0	74,5	36,85	13,12	0,2	38,5	24,82	6,97
	Lärche	2446	4	192	72	0	87,5	39,71	13,02	1	42	26,98	7,27

Wuchsgebiet 3 „Rhön“

Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
3.1	Buche	15045	3	235	91	0	120,5	30,47	14,67	0,2	39	22,36	7,25
	Douglasie	892	5	139	38	0	98,5	25,86	13,9	0,2	45	18,55	8,14
	Eiche	7546	4	279	118	0	106,5	34	15,19	0,2	39,1	22,31	6,53
	Fichte	10425	3	176	72	0	92,5	31,11	13,48	0,2	44	23,05	7,99
	Kiefer	12148	2	193	87	0	73,5	33,19	12	0,2	40,4	23,33	6,17
	Lärche	5958	3	203	68	0	93,5	33,3	13,44	0,2	42	23,48	6,66
	Tanne	174	3	117	73	0	57,5	34,17	12,15	0,2	33,5	24,14	6,39
3.2	Buche	1202	7	214	95	0	130,5	36,27	16,18	0,3	38,5	24,03	7,24
	Fichte	1022	5	205	49	0	73,5	30,09	11,02	0,5	35	17,7	6,18
	Lärche	108	26	120	64	13,5	73,5	39,2	12,59	10,5	36,5	22,64	6,7

Wuchsgebiet 4 „Fränkische Platte

Wuchs			Alter (Jahre)			BHD (cm)				Höhe (m)			
bezirk	Baumart	Datensätze	von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
4.1	Buche	609	3	189	93	0	84,5	32,25	17,36	0,2	34	20,7	7,81
	Eiche	688	3	213	89	0	80,5	30,03	18,07	0,2	30,5	18,48	7,38
	Fichte	199	3	147	62	0	58,5	30,73	15,49	0,3	36	22,39	9,76
	Kiefer	233	5	163	65	0	58,5	24,88	11,92	0,8	32	18,33	6,63
	Lärche	101	3	131	60	0	63,5	34,17	14,37	0,3	35	21,9	6,83
4.2	Buche	2291	3	171	82	0	95,5	33,29	15,96	0,2	47	24,92	8,86
	Douglasie	106	4	127	43	1,5	74,5	28,05	16,62	1,8	40,5	20,19	8,77
	Eiche	1602	3	260	94	0	114,5	32,82	16,26	0,3	38	22,98	7,33
	Fichte	604	3	150	69	0	71,5	33,39	12,51	0,2	40	25,26	7,38
	Kiefer	540	6	134	71	2,5	74,5	29,12	10,84	2	38	22,69	6,42
	Lärche	473	25	129	74	7,5	74,5	37,66	10,67	11	41	27,11	5,55

Wuchsgebiet 5 „Fränkischer Keuper und Albvorland“

Wuchs			Alter (Jahre)			BHD (cm)				Höhe (m)			
bezirk	Baumart	Datensätze	von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
5.1	Buche	1460	15	200	91	12,5	89,5	34,49	14,2	8	41	25,19	6,17
	Eiche	778	19	230	107	12,5	94,5	33,57	11,83	10	38	24,52	4,72
	Fichte	550	17	133	71	12,5	64,5	35,28	10,98	10,5	42,5	27,1	6,23
	Kiefer	574	21	160	79	12,5	64,5	32,91	11,37	8,5	41	24,89	6,33
	Lärche	296	17	155	69	12,5	61,5	33,57	11,4	7	42	25,4	6,56
5.2	Buche	2793	17	180	91	12,5	84,5	33,82	14,81	8,5	43	24,87	6,4
	Eiche	1500	17	240	122	12,5	82,5	37,86	10,09	9	39	25,94	4,07
	Fichte	541	15	158	65	12,5	61,5	30,11	12,08	7	39,5	22,73	7,15
	Kiefer	1051	15	246	93	12,5	70,5	36,16	11,97	6,2	42	24,89	6,04
	Lärche	458	17	159	69	12,5	78,5	37,04	13,29	8,5	43	25,98	6,65
5.3	Fichte	144	12	137	71	0	61,5	28,15	10,09	0,7	37	31,76	6,53
	Kiefer	216	12	160	89	0	63,5	32,76	8,91	1	36,5	25,38	5,35
5.4	Buche	312	6	231	104	0	69,5	35,61	14,42	0,4	41,2	25,03	7,4
	Eiche	357	8	231	100	0	85,5	32	15,04	1,2	33,9	22,97	6,55
	Fichte	132	7	152	68	0	57,5	30,09	11,82	0,6	37,7	24,02	7,12
	Kiefer	257	4	161	90	0	66,5	32,41	10,72	0,4	36,8	25,34	5,79
5.5	Eiche	81	10	174	79	1,5	69,5	29,26	17,6	1,5	31,5	20,43	8,78
	Fichte	95	10	129	78	0	65,5	31,53	11,42	1	35	24,74	6,74
	Kiefer	440	10	189	96	0	60,5	33,79	8,33	1	60,5	26,08	4,75

Fortsetzung WUG 5													
Wuchs			Alter (Jahre)			BHD (cm)				Höhe (m)			
bezirk	Baumart	Datensätze	von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
5.6	Eiche	161	3	160	64	0	65,5	22,62	18,05	0,3	36	15,63	10,19
	Fichte	954	3	164	64	0	61,5	25,58	14,88	0,2	41,5	19,63	9,75
	Kiefer	3734	3	200	81	0	61,5	29,54	9,36	0,3	37	21,75	5,58
	Lärche	114	15	143	52	10,5	57,5	26,22	9,82	10	31,5	20,26	5,3
5.7	Buche	164	8	192	77	0	69,5	31,01	15,21	0,7	41,9	23,23	8,71
	Eiche	180	4	192	97	0	78,5	35,13	15,96	0,2	36,1	24,54	7,74
	Fichte	159	6	137	72	0	65,5	35,69	13,03	0,3	36,1	26,6	8,2
	Kiefer	420	20	144	87	10,5	62,5	36,26	9,53	10,3	38,1	26,03	5,09
	Lärche	159	20	141	74	4,5	67,5	35,08	11,8	6,6	38,7	25,39	5,36
5.8	Buche	231	6	183	84	0	68,5	29,48	15,73	0,4	37,4	21,84	8,42
	Eiche	286	10	250	98	0	110,5	34,32	17,9	0,3	39,6	23,43	8,55
	Fichte	1006	3	173	75	0	68,5	30,96	13,1	0,2	44	23,98	8,03
	Kiefer	903	4	173	88	0	61,5	34,02	10,64	0,4	36,5	24,25	5,94
	Lärche	160	15	163	67	10,5	53,5	29,43	10,19	8,5	36	23,2	6,52
	Tanne	176	20	163	103	1,5	65,5	38,72	11,71	1,8	40	27,85	6,32

Wuchsgebiet 6 „Frankenalb und Oberpfälzer Jura“

Wuchs-			Alter (Jahre)			BHD (cm)				Höhe (m)			
bezirk	Baumart	Datensätze	von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
6.1	Buche	4716	1	192	86	0	88,5	30,56	14,52	0,2	42	22,73	8,34
	Douglasie	179	14	118	54	1,5	81,5	34,63	18,89	1,8	43,7	24,32	9,78
	Eiche	182	7	159	55	0	65,5	20,6	14,27	0,4	33,9	16,47	8,64
	Fichte	6884	3	179	76	0	81,5	32,79	13,42	0,2	43,5	24,18	8,2
	Kiefer	3992	7	220	95	0	65,5	33,22	9,2	0,3	43	24,37	6,03
	Lärche	949	5	185	64	0	87,5	30,8	12,74	0,5	41,5	23,29	7,12
	Tanne	201	5	164	87	0	65,5	33,3	14,12	0,2	36	24,75	8,4
6.2	Buche	9675	1	215	78	0	88,5	30,27	15,61	0,2	44,1	22,51	8,74
	Douglasie	359	13	144	53	2,5	99,5	34,13	20,29	3	47	23,53	10,5
	Eiche	2087	2	292	76	0	90,5	29,02	17,28	0,2	40	21,25	9,92
	Fichte	16824	3	200	69	0	89,5	32,74	14,17	0,2	45,5	24,28	8,45
	Kiefer	3409	3	190	86	0	77,5	33,68	10,52	0,3	38,5	23,87	6,25
	Lärche	1654	3	201	70	0	87,5	33,34	13,98	0,2	46,5	24,31	7,95

Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
	Tanne	883	5	175	93	0	75,5	36,02	13,66	0,2	41,5	25,3	7,94
6.3	Buche	4236	4	210	81	0	94,5	30,75	13,86	0,2	40	23,09	7,28
	Eiche	2388	3	260	91	0	120,5	36,14	20,23	0,2	41,5	22,48	8,28
	Fichte	3889	3	154	55	0	81,5	29,31	13,09	0,2	40,1	22,45	8,47
	Kiefer	566	5	167	75	0	71,5	33,84	12,75	0,3	41,5	24,28	7,55
	Lärche	379	23	144	69	12,5	83,5	39,36	14,06	7,8	39,4	26,23	6,46
6.4	Buche	125	5	167	56	0	90,5	24,15	16,39	0,2	31,5	17,32	8,17
	Eiche	114	5	169	39	0	73,5	27,69	14,59	0,3	32	10,55	9,66
	Fichte	1046	4	169	67	0	70,5	29,99	14,78	0,3	40,2	21,78	9,08
	Kiefer	960	4	167	81	0	65,5	32,01	10,58	0,4	39,2	22,23	6,05
	Lärche	268	5	167	52	0	63,5	29,1	10,31	0,3	33,5	21,53	5,36
6.5	Buche	691	6	325	78	0	85,5	28,17	16,1	0,2	38	20	8,92
	Eiche	242	5	158	57	0	68,5	22,41	15,37	0,4	36	16,21	8,73
	Fichte	2549	5	194	75	0	67,5	30,52	12,53	0,2	39,8	23,24	7,55
	Kiefer	2405	5	270	88	0	64,5	32,62	9,86	0,3	38,6	24,06	5,8
	Lärche	366	6	158	63	0	77,5	31,22	12,03	1	37,5	22,8	6,62

Wuchsgebiet 8 „Frankenwald, Fichtelgebirge, Steinwald“

Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
8.1	Buche	1341	3	180	68	0	94,5	30,46	16,31	0,2	43,5	19,51	8,06
	Douglasie	143	11	100	43	1,5	86,5	31,29	13,68	1,9	42,1	20,05	6,63
	Fichte	8732	4	165	65	0	86,5	33,68	13,17	0,3	46,9	23,72	7,2
	Kiefer	156	12	165	82	4,5	55,5	32,2	9,05	3,2	31,7	21,85	4,97
	Lärche	310	4	100	54	0	52,5	31,98	9,18	0,3	33,9	21,36	4,8
	Tanne	185	12	150	84	0	63,5	35,83	11,82	1,2	38,2	25,58	6,8

Wuchsgebiet 9 „Oberpfälzer Becken- und Hügelland“

Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
9.1	Kiefer	133	35	140	92	12,5	48,5	32,5	8,49	11,9	33,4	23,78	5,42

Wuchsgebiet 11 „Bayerischer Wald“

Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
11.2	Buche	468	8	209	70	0	72,5	31,24	16,18	0,4	37	21,39	8,67
	Eiche	118	7	164	73,8	0	74,5	30,6	16,28	0,8	37	19,9	8,26
	Fichte	1851	5	209	77	0	92,5	40,19	16,22	0,2	44,5	26,98	9,2
	Kiefer	120	5	154	78	0	69,5	36,73	11,55	0,7	31,5	22,42	5,44
	Tanne	366	7	209	81	0	75,5	37,7	15,27	0,8	41	24,61	8,35
11.3	Buche	3491	2	240	88	0	91,5	28,15	11,86	0,2	38,5	20,87	5,96
	Fichte	11691	3	309	91	0	102,5	40,9	15,09	0,2	44	24,75	7,72
	Tanne	1565	5	300	104	0	85,5	42,07	14,06	0,3	42	25,46	6,77

Wuchsgebiet 12 „Tertiäres Hügelland“.

Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
12.1	Eiche	170	2	180	81	0	143,5	40,79	26,71	0,2	32	20,13	8,74
	Fichte	178	3	88	54	0	57,5	28,46	11,18	0,2	30,4	20,94	6,55
	Kiefer	96	32	100	59	13,5	62,5	28,33	8,78	13	27,8	22	3,78
12.5	Kiefer	107	6	70	33	0	30,5	17,36	4,47	0,7	19	13,77	2,66
12.6	Kiefer	185	20	141	91	11,5	54,5	36,48	8,17	9,4	32,8	24,51	4,1
12.7	Buche	6346	3	170	71	0	106,5	31,65	15,34	0,2	43,9	23,94	9,02
	Douglasie	503	5	120	46	0	85,5	32,89	16,44	0,3	49	23,46	9
	Eiche	2013	2	309	61	0	146,5	27,2	18,35	0,2	40,5	19,32	10,22
	Fichte	16910	2	168	62	0	92,5	32,38	13,55	0,2	45	25,49	9,41
	Kiefer	1514	2	168	75	0	66,5	35,26	10,2	0,2	43,3	27,31	6,65
	Lärche	740	3	188	74	0	86,5	43,48	13,48	0,3	46,2	30,63	6,72
	Tanne	763	6	107	46	0	66,5	26,68	13,45	0,3	39,5	18,76	8,63
12.8	Buche	1377	6	180	72	0	115,5	34,31	15,72	0,2	42	23,64	7,71
	Douglasie	131	8	118	50	0	110,5	38,11	21,09	0,4	50	25,17	10,54
	Eiche	479	4	180	62	0	100,5	29,89	16,74	0,2	38,4	20,11	8,45
	Fichte	5215	4	191	66	0	81,5	36,41	12,94	0,2	45,2	27,13	7,89
	Kiefer	1209	2	209	72	0	74,5	34,83	10,07	0,3	39,9	26,08	5,84
	Lärche	138	6	145	77	1,5	70,5	40,79	12,76	1,9	41,5	29,43	6,89
	Tanne	245	4	128	67	0	76,5	34,42	16,3	0,3	44,5	23,69	10,17

Fortsetzung WUG 12													
Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
12.9	Buche	1745	4	198	90	0	100,5	37,72	18,25	0,4	46	26,17	9,61
	Eiche	351	5	190	90	1,5	96,5	36,9	17,23	1,5	40,3	23,7	8,34
	Fichte	2541	3	181	84	0	104,5	40,06	14,89	0,2	48,8	27,26	8,48
	Kiefer	388	4	181	82	0	75,5	35,96	9,91	0,3	39,3	22,26	5,13
	Lärche	299	4	151	77	0	70,5	38,69	12,03	0,3	43,3	26,36	6,62
	Tanne	548	4	170	95	0	84,5	39,83	17,21	0,2	42,5	26,09	9,77

Wuchsgebiet 13 „Schwäbisch-Bayerische Schotterplatten- und Altmoränenlandschaft“

Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
13.2	Buche	3670	2	178	41	0	120,5	19,46	18,5	0,2	45	13,04	10,33
	Douglasie	233	6	100	28	0	79,5	19,91	13,48	0,9	41	14,22	7,52
	Eiche	2204	2	260	33	0	88,5	13,62	15,86	0,2	35	9,73	8,46
	Fichte	19390	2	220	56	0	92,5	31,7	18,21	0,2	47	22,15	11,45
	Kiefer	2193	2	270	61	0	75,5	30,75	14,1	0,2	40,5	21,32	8,5
	Lärche	249	4	115	55	0	74,5	32,2	17,98	0,3	40,5	22,47	11,01
13.3	Fichte	470	5	112	60	0	64,5	33,45	14,42	0,2	37	24,09	9,21
	Kiefer	289	3	145	70	0	54,5	32,02	11,69	0,3	35	24,3	7,42
13.4	Buche	1041	4	145	66	0	124,5	30,51	15,91	0,2	43,2	23,48	9,75
	Eiche	136	4	149	27	0	59,5	11,29	13,66	0,2	36,1	9,24	9,46
	Fichte	2578	3	143	59	0	89,5	31,19	15,68	0,2	43,7	23,56	11,01
13.5	Fichte	218	4	102	51	0	70,5	32,64	14,28	0,3	41	23,57	9,75
13.6	Buche	228	6	160	44	0	103,5	20,66	19,46	0,4	35	14,71	9,99
	Eiche	429	7	161	25	0	87,5	9,96	14,64	0,4	33	7,71	7,7
	Fichte	1344	3	145	41	0	77,5	22,97	18,41	0,2	40	16,45	11,95
	Kiefer	204	2	135	43	0	58,5	22,5	16,52	0,3	35	16,37	10,42
	Tanne	227	4	145	31	0	76,5	13,51	17,04	0,2	37,5	9,27	10,24

Wuchsgebiet 14 „Schwäbisch-Bayerische Jungmoräne und Molassevorberge“

Wuchs- bezirk	Baumart	Datensätze	Alter			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
14.3	Buche	452	5	165	63	0	70,5	25,33	17,25	0,2	43,5	18,88	10,94
	Fichte	2552	1	235	64	0	86,5	30,1	17,29	0,2	46,6	21,28	11,22
	Kiefer	95	5	156	80	0	63,5	33,04	11,9	0,3	36	22,43	7,9
	Tanne	153	5	186	63	0	78,5	25,76	20,29	0,2	41,9	16,81	12,38
14.4	Buche	2514	4	280	62	0	164,5	26,91	20,17	0,2	40	18,77	10,96
	Eiche	322	3	198	48	0	99,5	21,37	21,85	0,2	37,5	13,74	10,22
	Fichte	6885	4	232	62	0	97,5	31,97	17,22	0,2	45	22,21	10,7
	Kiefer	1295	3	250	80	0	68,5	29,35	14,11	0,2	36	16,83	9,03
	Lärche	242	4	155	70	0	82,5	37,69	17,5	0,5	42	25,97	9,49
	Tanne	348	4	165	48	0	85,5	23,64	21,21	0,2	38,5	15,15	12,11

Wuchsgebiet 15 „Bayerische Alpen“

Wuchs- bezirk	Baumart	Datensätze	Alter (Jahre)			BHD (cm)				Höhe (m)			
			von	bis	Mittel	von	bis	Mittel	Stabw	von	bis	Mittel	Stabw
15.1	Buche	804	7	219	80	0	93,5	29	14,2	0,3	39	21,72	9,18
	Fichte	1500	3	209	78	0	94,5	41,72	17,56	0,2	47	27,23	10,48

9.4 Installationsanleitung für die Softwarelösung „SilvaPEM“

Der schriftlichen Version dieser Arbeit liegt ein Prototyp der Softwarelösung Silva-PEM bei. Diese entspricht dem Entwicklungsstand zum 01.12.2006. Silva-PEM befindet sich auf beiliegendem Datenträger im Verzeichnis /SilvaPEM/. Weiterhin befinden sich im Verzeichnis /Testdatensatz/ XML-Dateien, an denen die grundlegende Funktionalität getestet werden kann. Bei den Testdaten handelt es sich um vergleichsweise alte Forsteinrichtungsdaten für die Baumarten Fichte, Kiefer und Lärche des Wuchsbezirkes 4.2 „Südliche Fränkische Platte“ (vgl. Abschnitt 3.2.3).

Um die Software testen zu können, müssen folgende Voraussetzungen erfüllt sein:

1. Es muss eine Java-Virtual-Machine installiert sein. Falls nicht vorhanden, kann diese kostenfrei heruntergeladen werden unter: <http://www.sun.com>
2. Zur Durchführung neuer regressionsanalytischer Berechnungen muss die freie statistische Programmiersprache R installiert sein. Hierbei muss mindestens die Version 2.3.5 vorliegen, um einen Daten- und Ergebnisaustausch zwischen SilvaPEM und R zu ermöglichen über die JRI-Schnittstelle zu ermöglichen (vgl. Abschnitt 5.2). Im R-Installationsordner muß das Paket rJava verfügbar sein. Sowohl R als auch rJava sind erhältlich unter <http://www.r-project.org>.
3. Zur Generierung von Entscheidungsbäumen muss die Java-basierte Data-Mining-Suite WEKA installiert sein. Diese ist erhältlich unter <http://www.cs.waikato.ac.nz/ml/weka>

Zum Programmstart sind folgende Schritte erforderlich (Beispiele für MS-Windows-Systeme):

1. Entweder Pfade für die DLL's (R.dll, RBas.dll, JRI.dll) setzen oder diese DLL's aus dem Silva-PEM-Unterverzeichnis /SilvaPEM/DLL in das Systemverzeichnis kopieren (einmaliger Arbeitsschritt).
2. Z. B. von MS-DOS-Eingabeaufforderung aus in das Verzeichnis wechseln, in dem sich JAVA.exe befindet:
Bsp: `cd C:\Sun\AppServer\jdk\bin\`
3. Z. B. von MS-DOS-Eingabeaufforderung aus SilvaPEM starten:
Bsp: `java -cp "c:/Temp/SilvaPEM/SilvaPEM/" SilvaPEM`

9.5 Quellcode des SilvaPEM-Moduls „NonlinearRegression.java“

```

/**
 * Die Klasse NonlinearRegression.java nimmt die eingelesenen Baumdaten-
 * sätze auf, bereitet sie auf und übergibt sie über das JavatoRInterface (JRI)
 * des R-Packages "rJava" an das Statistikpaket R. Ebenso werden die
 * Steueranweisungen zur Durchführung einer nichtlinearen Regression an
 * R übergeben.
 * Bei erfolgreicher Abarbeitung wird von R das Ergebnisobjekt geholt und
 * in ein Ergebnisobjekt der Klasse AKPValues uebertragen
 *
 * @author Hans-Joachim Klemmt
 * @param TreeList -> wegen Aufruf aus SilvaPEM Zugriff über getObj()
 * @return Objekt der Klasse AKPValues
 *
 * TODO Startparameter der nichtlinearen Regression vom Benutzer abfragen
 * TODO Exception-Handling implementieren
 */

import java.awt.FileDialog;
import java.awt.Frame;
import java.io.BufferedReader;
import java.io.InputStreamReader;

import org.rosuda.JRI.REXP;
import org.rosuda.JRI.RMainLoopCallbacks;
import org.rosuda.JRI.Rengine;

public class NonlinearRegression
{
    //Constructor
    public NonlinearRegression()
    {
        //macht nix
    } //eo Constructor;

    //Methode execute()
    public AKPValues execute()
    {

        ObjectManagement TreeList = ObjectManagement.getObj();
        String[] Baumliste = TreeList.return_DMDDataSets();

        System.out.println(TreeList.returnAnzahlDatensaetze());

        //überprüfen, ob die R-Version > 2.31 ist
        if (!Rengine.versionCheck()) {
            System.err.println("** Version mismatch - Java files don't match library version.");
            System.exit(1);
        }

        //keine Kommandozeilen-Argumente übergeben
        String[] args = null;

        //REngine starten
        System.out.println("Creating Rengine (with arguments)");
        Rengine re=new Rengine(args, false, new TextConsole());
        System.out.println("Rengine created, waiting for R");

        //Alters- und Höhenangaben aus Baumliste holen und umwandeln
        int[] Alter = new int[TreeList.returnAnzahlDatensaetze()];
        double[] Hoehe = new double[TreeList.returnAnzahlDatensaetze()];

```

```
for(int i=0; i<TreeList.returnAnzahlDatensaetze()-1; i++)
{
    Alter[i] = TreeList.returnInventuralter(i);
    Hoehe[i] = (double)TreeList.returnHoeheMessbaum(i);
}

long xp1 = re.miPutIntArray(Alter);
long xp2 = re.miPutDoubleArray(Hoehe);

//gleiche Namen in R wie in der Inventurdatenbank setzen
re.miAssign("inv_alt",xp1,0);
re.miAssign("bhd_h",xp2,0);

REXP x = re.eval("inv_alt");
int d[]=x.asIntArray();

//Wir basteln uns ein DataFrame;-)
long la[] = {xp1,xp2};
long xp3 = re.miPutVector(la);

//R-Commands ausführen
re.eval("sh1<-data.frame(inv_alt,bhd_h)");
//nachfolgende Werte ermöglichen Berechnung für Testdaten Fichte und Kiefer
re.eval("nlCR.start=c(A=35, k=-0.02, P=2)");
re.eval("erg<-nls(bhd_h~A*((1-exp(k*inv_alt))^P), data=sh1, start=nlCR.start, trace=TRUE)");

long e=re.miParse("coef(erg)", 1);
long r=re.miEval(e, 0);
REXP x_new=new REXP(re, r);

//Ergebnisobjekt in JAVA verfügbar machen
double Ergebnis[] = x_new.asDoubleArray(); //Ergebnis[0]=A usw.
AKPValues Ergebnisobjekt = new AKPValues(Ergebnis[0], Ergebnis[1], Ergebnis[2]);

//R beenden
re.end();
System.out.println("REngine beendet");

//Ergebnisobjekt an aufrufende Funktion zurückgeben
return(Ergebnisobjekt);

} //eoM execute()

} //eoC NonlinearRegression
```

9.6 Lebenslauf

Name: Hans-Joachim K l e m m t

Geburtsdatum: 04.01.1973 (in Amberg)

Schulbildung: 1979-1983: Grundschule Neuhaus/ Pegnitz
1983-1992: mathematisch-naturwissenschaftliches Gymnasium Pegnitz

Universität 1992-1997: Studium der Forstwissenschaften an der Ludwigs-Maximilians-Universität München (LMU)

Abschluss nach 8 Vorlesungs- und 1
Praxissemester: Diplom-Forstwirt (Univ.) (Note: 1,7)

Diplomarbeit zum Thema: Untersuchungen zum
Auftreten des Buchenfarbkerns in unterfränkischen
Beständen

Auszeichnung: Leo-Schörghuber-Preis 1997 zur Förderung der
Holzforschung (dotiert mit 5.000 DM)

Vorbereitungsdienst:

1997-1999: Referendarsausbildung bei der Bayerischen
Staatsforstverwaltung (FE-Zeit: Kempten/ Allgäu,
FoA-Zeit: Weiden/ Oberpfalz, FoD-Zeit: FoD
Oberfranken (Bayreuth), Forsteinrichtung: Forst-
und Domänenamt Coburg)

Juni 1999: Große Forstliche Staatsprüfung in Bayern;
Abschluss mit Rangziffer 1 von 30
Prüfungsteilnehmern.

Weiterbildung (parallel zum Referendariat): Teilnahme am
Vorbereitungskurs „Organisationsprogrammierer C/ C++“ der
Studiengemeinschaft Darmstadt als Vorbereitung auf die IHK-
Prüfung von 1997 bis 1999.

IHK-Prüfung vor der IHK Offenbach am Main vom 28.05.2000.
Abschluss: „Organisationsprogrammierer C/ C++ (IHK)“

Berufstätigkeit:	1999:	Übernahme in die Bayerische Staatsforstverwaltung
	1999-2001:	Geschäftsaushilfe im Leitungsdienst am Bayerischen Forstamt Weißenstadt (Fichtelgebirge)
	2001-2003:	Mitarbeiter an der Forstdirektion Oberfranken/ Mittelfranken (Organisation der Referendarsausbildung, Mitarbeit „Verwaltungs-Controlling“, Mitarbeit „Öffentlichkeitsarbeit“)
	seit 2003:	Mitarbeiter am Lehrstuhl für Waldwachstumskunde der TU München (abgeordnet von der Bayerischen Landesanstalt für Wald und Forstwirtschaft)

9.7 Erklärung der selbständigen Anfertigung

Ich erkläre an Eides statt, dass ich die der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Promotionsprüfung vorgelegte Arbeit mit dem Titel:

Standortabhängige Ableitung der Höhenwuchsleistung aus Forstinventurdaten mit Hilfe von
Data-Mining-Methoden.

Grundlage für die regionale, standortbezogene Feinjustierung des forstlichen Wachstumsmodells
SILVA

unter Anleitung und Betreuung durch Herrn Prof. Dr. Hans Pretzsch ohne sonstige Hilfe erstellt und bei der Abfassung nur die gemäß § 6 Abs. 5 angegebenen Hilfsmittel benutzt habe.

Ich habe die Dissertation in keinem anderen Prüfungsverfahren als Prüfungsleistung vorgelegt.

Ich habe den angestrebten Doktorgrad noch nicht erworben und bin nicht in einem früheren Promotionsverfahren für den angestrebten Doktorgrad endgültig gescheitert.

Dir Promotionsordnung der Technischen Universität München ist mir bekannt.

Freising, 04.04.2007

Hans-Joachim Klemmt