Retrospektive Diskussion der Versuchsplanung für eine Holzqualitätsuntersuchung*)

Von J. SABOROWSKI und F. HAPLA

Abteilung für Forstliche Biometrie und Informatik und Institut für Forstnutzung der Universität Göttingen

1. Einführung

Die vorliegende Arbeit beschäftigt sich noch einmal rückblickend mit der Versuchsplanung für den von den Autoren beschriebenen Versuch in Abteilung 39 a des Staatsforstamtes Daun (Rheinland-Pfalz), in dem der Einfluß von Durchforstungsmaßnahmen auf verschiedene Holzeigenschaften geprüft werden sollte (HAPLA u. SABOROWSKI, 1984). Da die Gesamtzahl der Probestämme bereits festlag, konnte nur noch Einfluß auf die Anzahl der je Stamm auszuformenden Probekörper genommen werden. Aus der umfangreichen Voruntersuchung eines der insgesamt zehn aus beiden Flächen (Nullfläche, Durchforstungsfläche) entnommenen Probestämme konnten lediglich Erkenntnisse über die Varianzen der Qualitätsmerkmale **innerhalb** eines Stammes gewonnen werden, jedoch nicht über die Varianzen **zwischen** den Stämmen eines Bestandes. Deshalb war es damals nur möglich, die Stichprobenumfänge für die einzelnen Merkmale so zu wählen, daß die Einzelstammittelwerte mit möglichst geringem Aufwand hinreichend genau geschätzt werden könnten. Der ins Auge gefaßte Flächenvergleich spielte bei dieser Stichprobenplanung also noch keine Rolle.

Als Maß für die Genauigkeit der Mittelwertschätzung aus der Stichprobe wurde die halbe Breite $l = z\,(0.025) \cdot \sigma/\sqrt{n}$ des 95-%-Konfidenzintervalls für den Erwartungswert einer normalverteilten Grundgesamtheit gewählt, die in diesem Sinn mit 95 % Sicherheit eine obere Grenze für die mögliche Abweichung des arithmetischen Mittels der Stichprobenwerte vom wahren Merkmalsmittelwert darstellt. Für die Anzahl der den übrigen Stämmen zu entnehmenden Probekörper wurden aufgrund der in der Untersuchung von Stamm Nr. 05 ermittelten Varianzen und Mittelwerte Empfehlungen gegeben, die gewährleisten sollten, daß die geschätzten von den wahren Mittelwerten um nicht mehr als 5 % abweichen.

Dies schien uns nach den an diesem Probestamm beobachteten Varianzen bei allen Merkmalen (ausgenommen Bruchschlagarbeit im Kern 2 und Zugfestigkeit) in allen Stammabschnitten erreichbar zu sein. Das Hauptaugenmerk lag dabei auf den für den Vergleich von Durchforstungs- und Nullfläche interessanten Bereichen Splint und Kern 1 (Tabelle 4 in HAPLA u. SABOROWSKI, 1984).

Nun ist es an der Zeit zu überprüfen, in welchem Umfang die gesteckten Ziele erreicht wurden, und ob das hier gewählte Vorgehen für zukünftige Versuche zu empfehlen ist. Darüber hinaus stellt sich aber auch die Frage nach der Probestammzahl, die für eine angestrebte Testschärfe notwendig ist, bzw. ob insbesondere eine Reduktion der Stammzahl durch eine naheliegende Erhöhung der Anzahl der Probekörper je Stamm ausgeglichen werden kann. Mit diesen Fragen wollen wir uns in den nächsten Abschnitten beschäftigen.

2. Zur erreichten Genauigkeit der Mittelwertschätzung

Die aufgrund der Voruntersuchung geforderten Stichprobenumfänge wurden in der Hauptuntersuchung realisiert und nur in wenigen Fällen wegen zu geringer Stammdimension nur unwesentlich unterschritten. Da wir es nun fast immer mit zufälligen Stichproben und nicht mit einer maximalen Ausnutzung des vorhandenen Stammholzes zu tun haben, verwenden wir nach Abschluß der Hauptuntersuchung $l = t\,(n-1{,}0.025) \cdot s/\sqrt{n}$ in Prozent von der unteren Vertrauensbereichsgrenze $\bar{x} - l$ als Obergrenze für den Schätzfehler

2.1 Rohdichte ($\varrho\mu 0$)

Bei den insgesamt 80 zu schätzenden Mittelwerten (10 Stämme, Splint und Kern 1, 4 Stammabschnitte) lagen die so erhaltenen Obergrenzen für den Schätzfehler in nahezu allen Fällen (zum Teil weit) unter 5 %, nur viermal nicht nennenswert darüber (5.2 %, 5.4 %, 6.8 %, 7.2 %).

2.2 Statische Biegefestigkeit (β_B)

Die Fehlerabschätzungen lagen bei 13 von 40 Fällen über 5 %, davon 8mal noch unter 7 %, viermal unter 10 % und in einem Fall bei der „Rekordhöhe" von 14.7 %.

2.3 Statische Längsdruckfestigkeit (β_{DII})

Bei 22 von 80 Mittelwerten wurde die 5-%-Marke überschritten, davon lagen die Ergebnisse allerdings 16mal noch unter 7 %, nur 4mal unter 10 % und in zwei Fällen bei 12.8 % bzw. 12.7 %.

2.4 Dynamische Bruchschlagarbeit (ω)

Bei diesem Qualitätsmerkmal wurden wegen der relativ großen Streuung bei Stamm Nr. 05 schon n = 20 Probekörper je Stamm und Bereich für notwendig gehalten. Die Hauptuntersuchung zeigte jedoch, daß auch dies in sehr vielen Fällen nicht ausreichend war. In 12 von 20 Fällen liegt die Fehlerschranke zum Teil erheblich über 5 %, davon 9mal zwischen 7 % und 10 % und 3mal noch darüber (11 %, 12 % bzw. 14 %).

Die Ergebnisse bei den drei erstgenannten Holzeigenschaften können insgesamt als zufriedenstellend bezeichnet werden. Lediglich die Mittelwerte der dynamischen Bruchschlagarbeit konnten zu einem doch erheblichen Anteil nicht mit der geplanten Genauigkeit geschätzt werden. Die Ursache für die teilweise sehr starken Abweichungen von dem geplanten Wert ist in der bemerkenswerten Heterogenität der Varianzen **aller** Variablen zu sehen, die eine Stichprobenplanung natürlich sehr erschwert. So schwanken die 10 empirischen Standardabweichungen der Bruchschlagarbeit im Kern 1 des zweiten Mittelblocks zwischen 0.39 J/cm² und 1.01 J/cm², während für die Planung der Stichprobenumfänge aufgrund der Voruntersuchung 0.45 J/cm² zugrundegelegt wurde. Für die Rohdichte schwanken sie im Splintholz des ersten Mittelblocks zwischen 0.01 g/cm³ und 0.028 g/cm³ (Voruntersuchung: 0.028 g/cm³). Der Zufall hat dafür gesorgt, daß die Bruchschlagarbeit im Stamm Nr. 05 eine relativ geringe Streuung besaß. Im übrigen mußte die mit dem Bartlett-Box-Test geprüfte Homogenität der je zehn Varianzen in nahezu allen Bereichen und Variablen auf dem 5 %-Niveau verworfen werden. Die Verwendung des „Sicherheitszuschlages" (Faktor 1.5 für zu erwartende Varianzen der übrigen Stämme) war also durchaus sinnvoll, aber oft nicht ausreichend.

Andererseits müssen wir berücksichtigen, daß bei den Extremfällen eine bemerkenswerte Verbesserung der Schätzungen nur mit unverhältnismäßig größerem Meßaufwand zu erzielen ist, der häufig wegen der begrenzten Stammdimensionen gar nicht realisierbar wäre. So hatte sich für die Bruchschlagarbeit im Kern 1 des zweiten Mittelblocks von Stamm Nr. 10 mit $\bar{x}$ = 4.05 J/cm² und s = 1.01 J/cm² eine Schranke von ca. 14 % ergeben. Die gleichen Werte hätten mit n = 30 (statt n = 20) zu auch nur etwa 10 % Genauigkeit geführt und erst n = 60 zu knapp 7 %. Die angestrebte 5-%-Grenze liegt damit völlig außer Reichweite.

Darüber hinaus wird sich im folgenden zeigen, daß die Stichprobenumfänge innerhalb der Stämme für den Vergleich der beiden Probeflächen nur von untergeordneter Bedeutung sind. Die Streuung zwischen den Stämmen spielt dabei eine wesentlich wichtigere Rolle als die bisher diskutierte Streuung im Stamm.

Die durch die beobachtete Heterogenität der „Innerhalbvarian-

*) Diese Untersuchung wurde aus Forschungsmitteln der Deutschen Forschungsgemeinschaft gefördert.

zen" entstehenden Probleme für die Stichprobenplanung können nur durch eine größere Zahl von Probestämmen in der Voruntersuchung oder, wo immer dies möglich ist, durch Erfahrungen aus vergleichbaren Flächen angegangen werden. Anstelle des in der Voruntersuchung verwendeten recht unmotivierten Sicherheitszuschlags $1.5 \cdot \sigma^2$ bei einer angenommenen wahren Streuung σ^2 empfiehlt es sich, $l = t(n-1, \alpha/2) \cdot \sigma^*/\sqrt{n}$ zu verwenden. Dabei ist $\sigma^{*2} = \sigma^2 \cdot \chi^2(n-1, \alpha)/(n-1)$ das $100 \cdot (1-\alpha)$ Prozent-Quantil für s^2.

3. Ein statistisches Modell für den Flächenvergleich

Da es sich bei den Durchforstungsmaßnahmen um gezielte Eingriffe, bei der Auswahl der Stämme aus beiden Flächen um eine Zufallsauswahl handelt, haben wir ein hierarchisches Modell mit einem festen Faktor A „Bestandesbehandlung" (2 Stufen) und einem darin geschachtelten zufälligen Faktor B vor uns, dessen Stufenzahlen m_1 und m_2 durch die Zahl der Probestämme je Bestand gegeben sind (hier: $m_1 = m_2 = 5$). Auf jeder Stufe des zufälligen Faktors seien n Wiederholungen (n Probekörper je Stamm) vorhanden. Um die Fragen nach der geeignetsten Wahl von m_1, m_2 und n beantworten zu können, unterstellen wir der Einfachheit halber ein eindimensionales Modell gemäß Tabelle 1 (siehe auch RASCH (1976), S. 71 und S. 161 ff.).

$$y_{ijk} = \mu + a_i + b_{ij} + \varepsilon_{ijk} \quad i=1,2 \quad j=1,m_i \quad k=1,n$$
$$N = (m_1+m_2) \cdot n$$

Faktor	SQ	FQ	$\frac{1}{FG} \cdot E(SQ)$
A	$n \cdot \sum_{i=1}^{2} m_i (\bar{y}_{i..} - \bar{y}_{...})^2$	2-1	$\sigma_R^2 + n\sigma_B^2 + 2n \frac{m_1 m_2}{m_1+m_2} (a_1^2 + a_2^2)$
B(in A)	$n \cdot \sum_{i,j} (\bar{y}_{ij.} - \bar{y}_{i..})^2$	m_1+m_2-2	$\sigma_R^2 + n\sigma_B^2$
R	$\sum_{i,j,k} (y_{ijk} - \bar{y}_{ij.})^2$	$N-m_1-m_2$	σ_R^2
Gesamt	$\sum_{i,j,k} (y_{ijk} - \bar{y}_{...})^2$	N-1	σ_y^2

Aus den üblichen Voraussetzungen und Nebenbedingungen heben wir insbesondere die Unkorreliertheit von b_{ij} $(N(O, \sigma_B^2))$ und ε_{ijk} $(N[O, \sigma_R^2])$ hervor. $\mu + \alpha_i$ $(i = 1,2)$ sind die Bestandesmittelwerte, und für die Nullhypothese $\mu + \alpha_1 = \mu + \alpha_2$ bzw. $\alpha_1^2 + \alpha_2^2 = O$ ergibt sich aus der letzten Spalte der Varianztabelle sofort die mit den Freiheitsgraden 1 und $m_1 + m_2 - 2$ F-verteilte Teststatistik $F = (m_1 + m_2 - 2) \cdot SQ_A/SQ_B$. Verwenden wir noch statt der Einzelwerte y_{ijk} die Stammittelwerte $z_{ij} = \bar{y}_{ij.}$ mit der Verteilung $N(\mu + \alpha_i, \sigma^2)$ und $\sigma^2 = \sigma_B^2 + \sigma_R^2/n$, so ergibt sich

$$\sqrt{F} = \sqrt{\frac{m_1 m_2}{m_1 + m_2}} \frac{\bar{z}_{1.} - \bar{z}_{2.}}{\sqrt{\frac{1}{m_1 + m_2 - 2} \Sigma (z_{ij} - \bar{z}_{i.})^2}} = t(m_1 + m_2 - 2)$$

d. h., der übliche Zweistichproben-t-Test auf der Basis der Mittelwerte z_{ij}. σ_R^2 ist die bereits diskutierte Varianz innerhalb der Stämme, σ_B^2 die sogenannte Varianz zwischen den Stämmen.

4. Zur Stichprobenplanung für das vorliegende Testproblem

Der Experimentator hat die Aufgabe, die Anzahl n der Proben je Stamm, aber auch die Anzahlen m_1 und m_2 der Stämme aus den beiden Probeflächen zu bestimmen. Es ist naheliegend, daß sich ohne weitere Anforderungen an das Testverfahren eine optimale Lösung für diese Stichprobenumfänge nicht finden läßt. Deshalb ist es üblich, neben der Einhaltung des Fehlers 1. Art α (Signifikanzniveau) auch eine sogenannte Testschärfe $\beta = \beta$ (d) zu fordern, die definiert ist als die Wahrscheinlichkeit dafür, daß der Test die Nullhypothese verwirft, falls die Flächenmittelwerte sich um die Differenz $d = |\alpha_1 - \alpha_2|$ unterscheiden. Unsere Teststatistik $t(m_1 + m_2 - 2)$ hat aber unter der Hypothese $d \neq O$ eine sog. nichtzentrale $t(m_1 + m_2 - 2, \delta)$-Verteilung mit dem Nichtzentralitätsparameter $\delta = d/\sigma \cdot \sqrt{m_1 m_2 / (m_1 + m_2)}$. δ ist offensichtlich immer dann am größten, wenn $m_1 = m_2 = m$ ist.

Antwort 1:
Der t-Test für den Vergleich der Flächenmittelwerte ist immer dann am schärfsten, wenn $m_1 = m_2$ ist. Daher empfiehlt es sich, nach Möglichkeit gleiche Stammzahlen aus beiden Flächen zu entnehmen.

Man kann zeigen (HERRENDÖRFER u. a., 1973), daß für großes m die gewünschte Schärfe β bei vorgegebenem d erreicht wird, wenn m und n so gewählt werden, daß

$$m = 2 \cdot (t(2m-2, \alpha/2) + t(2m-2, 1-\beta))^2 \cdot \sigma^2/d^2 = \text{const.} \cdot (\sigma_B^2 + \sigma_R^2/n) \quad (*)$$

$t(k, \gamma)$ ist das $(1-\gamma)$-Quantil der zentralen t-Verteilung mit k Freiheitsgraden. Hieraus ergeben sich unmittelbar die folgenden bereits bei HERRENDÖRFER u. a. in anderem Zusammenhang gegebenen Antworten.

Antwort 2:
Ist die Zahl n der Proben je Stamm vorgegeben, so kann m für jede erforderliche Testschärfe β (z. B. $\beta = 0.9$) iterativ aus (*) bestimmt oder aus Tabellen bzw. Grafiken (z. B. OWEN, 1962; S. 41 ff.) abgelesen werden, wenn σ^2 bekannt ist.

Antwort 3:
Steht umgekehrt die Zahl m der Stämme pro Fläche fest, so müssen wir drei Fälle unterscheiden, denn die rechte Seite der Gleichung (*) kann durch Wahl von n nie größer als const. $\cdot (\sigma_B^2 + \sigma_R^2)$ und nie kleiner als const. $\cdot \sigma_B^2$ werden.

(i) *$m < \text{const.} \cdot \sigma_B^2$*
In diesem Fall kann die angestrebte Schärfe auch durch beliebig großes n niemals erreicht werden.

(ii) *$m \geqq \text{const.} \cdot (\sigma_B^2 + \sigma_R^2)$*
Es genügt n = 1, um β zu erreichen oder sogar zu überschreiten.

(iii) *$\text{const.} \cdot \sigma_B^2 \leqq m < \text{const.} \cdot (\sigma_B^2 + \sigma_R^2/n)$*
Dann ist $n = \sigma_R^2/(m/\text{const.} - \sigma_B^2)$ zu setzen.

Liegt weder n noch m, sondern der Gesamtstichprobenumfang $N = 2 \cdot m \cdot n$ als limitierende Größe fest, so sind n und m als optimal anzusehen, wenn d minimal, d. h. die Schärfe β für möglichst kleines d angenommen wird. Aus (*) folgt

$$d^2 = 2(t(2m-2, \alpha/2) + t(2m-2, 1-\beta))^2 \cdot (\sigma_B^2/m + \sigma_R^2/N \cdot 2)$$

und

Antwort 4:
Für gegebenen Gesamtumfang $N = 2 \cdot n \cdot m$ ist $m = N/2$ und damit n = 1 die optimale Lösung.
Wir wollen die Konsequenzen dieser Aussagen kurz an zwei Beispielen diskutieren.

5. Anwendungsbeispiele

Die im soeben beschriebenen Modell vorausgesetzte Varianzgleichheit (σ_R^2 unabhängig von i, j und k) mußte anhand der vorliegenden Daten meist verworfen werden. Diese Abweichung von den Modellvoraussetzungen ist jedoch vernachlässigbar, da nur der Bruchteil σ_R^2/n in die Varianz σ^2 der Stammittelwerte eingeht und σ_B^2 in der Regel höchstens von ungefähr gleicher Größenordnung war wie σ_R^2. Hierin ist auch der Grund für den sehr schnell verschwindenen Einfluß von n auf die Schärfe von β zu sehen, den wir weiter unten an den Beispielen demonstrieren werden.

Die Streuungen der Mittelwerte z_{ij} in den beiden Flächen konnten hingegen bei unserer Untersuchung immer als homogen angesehen werden.

Mit den Bezeichnungen von Tabelle 1 schätzen wir die im weiteren Verlauf benötigten Varianzkomponenten wie üblich durch

$$s_R^2 = SQ_R/(N - 2m)$$
$$s_B^2 = [SQ_B/(2m - 2) - s_R^2]/n$$

Beispiel 1: Druckfestigkeit, 3. Mittelblock, Kern 1
(n = 10, m = 5)

$s_R^2 = 1043.5 \qquad s_B^2 = 1686.0$

$s^2 = 1686.0 + 104.4 = 1790.4$

Unterstellen wir, daß s_R^2 und s_B^2 die wahren Varianzen sind, so folgt für d = 40 daN/cm², d. h. $\Delta = d/\sigma = 40/\sqrt{1790.4} = 0.95$, $\alpha = 0.05$ und $\beta = 0.90$ eine notwendige Stammzahl von m = 26 für jeden der beiden Bestände (OWEN, S. 44). Das bedeutet, in jedem Bestand hätten aus 26 Stämmen je 10 Probekörper ausgewertet werden müssen, um bei einer Bestandesmittelwertdifferenz von mindestens 40 daN/cm² (dies bedeutet grob gerechnet 10 % der wahren Bestandesmittelwerte) mit einer Wahrscheinlichkeit von wenigstens 90 % signifikante Unterschiede auf dem 5-%-Niveau erwarten zu können. Dies steht in keinem Verhältnis zu der tatsächlich verwendeten Anzahl m = 5.

Umgekehrt läßt sich bei gegebenen Stichprobenumfängen n und m die Schärfe β aus den Grafiken von OWEN (S. 31 ff.) entnehmen. Es ist zu beachten, daß dort mit β nicht die Schärfe selbst, sondern 1-Schärfe gemeint ist.

Mit welcher Schärfe ist nun bei den von uns verwendeten Umfängen n = 10 und m = 5 unter den genannten Bedingungen zu rechnen? Wegen $\delta = \Delta \cdot \sqrt{m/2} = 0.95 \cdot \sqrt{2.5} = 1.5$ ergibt sich nach OWEN, S. 35 (f = 2 m − 2 = 8) $\beta \approx 0.4$! Diese Testschärfe ist durch Erhöhung von n praktisch nicht mehr zu verbessern, denn σ^2 wird nie kleiner als 1686, ganz gleich, wie groß n gewählt wird, d. h. $\delta = d/\sigma \cdot \sqrt{m/2} = 1.54$ und damit $\beta \approx 0.42$, so daß die Testschärfe bei m = 5 nie größer als etwa 0.42 werden kann. Für n = 1 und m = 5, das andere Extrem, erhalten wir $\beta = 0.3$.

Bei der in unserem Fall gegebenen Stammzahl m = 5 je Bestand wurde mit n = 10 also nahezu das Maximum an Testschärfe erreicht. Eine Erhöhung der Probekörperzahl erweist sich damit als wenig ökonomisch.

Beispiel 2: Bruchschlagarbeit, 2. Mittelblock, Kern 1
(n = 20, m = 5)
Aus $s_R^2 = 0.62$, $s_B^2 = 0.45$, d = 0.5 J/cm² (ca. 10 % der wahren Mittelwerte) folgt wie im Beispiel 1 $\delta = 1.14$ und $\beta \approx 0.28$. Bei festgehaltener Stammzahl hätte β hier nie größer als etwa 0.29 werden können. Aber auch für n = 10 ergäbe sich bereits $\beta \approx 0.26$, so daß die für die Schätzung der Stammittelwerte für notwendig gehaltene Zahl von 20 Probekörpern je Stamm im Hinblick auf den Vergleich der beiden Versuchsflächen als stark überhöht angesehen werden muß. Statt m = 5 hätte hier wegen $\Delta = 0.72$, m = 44 gewählt werden müssen, um die Vorgaben einzuhalten.

5. Berücksichtigung von Kosten bzw. Arbeitszeiten

Bisher haben wir uns bei der Frage nach optimalen Stichprobenumfängen nur von einer wünschenswerten Testschärfe leiten lassen. Daß dabei andere wichtige Nebenbedingungen wie etwa Kosten oder statt dessen Arbeitszeiten völlig außer acht gelassen werden, wird besonders in Antwort 4 deutlich, in der bei vorgegebenem Gesamtstichprobenumfang die maximal mögliche Stammzahl mit nur je einer Probe pro Stamm als Optimum erkannt wird. Dies kann ja für die Praxis nur bedeuten, daß im Rahmen der zeitlichen und finanziellen Möglichkeiten eine möglichst große Stammzahl anzustreben ist. Um diese Nebenbedingungen besser in den Griff zu bekommen, nehmen wir an, daß sich die Gesamtarbeitszeit Z additiv zusammensetzt aus der Zeit je Stamm c_B multipliziert mit der Stammzahl und der Zeit je Probekörper c_R multipliziert mit der Anzahl aller Probekörper, also

$$Z = 2\,m \cdot c_B + 2\,m \cdot n \cdot c_R = 2\,m\,(c_B + nc_R)$$

Dieser Ansatz ist aus der Stichprobentheorie bekannt (COCHRAN, 1977; S. 280). Anders als dort, wo die Varianz als Optimalitätskriterium benutzt wird, wollen wir hier die Arbeitszeit Z minimieren unter der Nebenbedingung, daß bei gegebenem d eine Testschärfe β eingehalten wird. Es soll also wieder die Gleichung (*) für n und m erfüllt sein. Ersetzen wir m in Z durch (*), so folgt weiter

$$dZ/dn = 2 \cdot \text{const.}\,(-\,c_B\,(\sigma_R/n)^2 + c_R\sigma_B^2)$$

Demnach ist Z minimal, wenn

$$n = \sqrt{\frac{c_B}{c_R} \cdot \frac{\sigma_R^2}{\sigma_B^2}}$$

m läßt sich dann wieder aus (*) bzw. Tabellen oder Grafiken bestimmen. Wir wollen auch hier einige Beispiele durchrechnen. Aus Tabelle 2 entnehmen wir die Zeiten pro Baum und pro Probekörper (Baumzeit und Probenzeit). Die Zeit für Aufspannen, Anreißen usw. von 320 Minuten je Stamm muß verdoppelt werden, da zwei Arbeitskräfte beteiligt sind, und wir legen immer $\alpha = 0.05$ und $\beta = 0.9$ zugrunde.

Beispiel 1: Druckfestigkeit, 3. Mittelblock, Kern 1
Aus c_B = 730 Min, c_R = 9 Min und d = 40 daN/cm² folgt zunächst

$$n = \sqrt{\frac{730}{9} \cdot \frac{1043}{1686}} = 7 \text{ und } \sigma^2 = 1686 + 1043/7 = 1835$$

Mit $\Delta = d/\sigma = 0.93$ ergibt sich daraus nach OWEN m = 27. Wählt man statt dessen d = 50 daN/cm² (d. h. $\Delta = 1.17$), so kommt man schon mit m = 18 Stämmen je Fläche aus.

Beispiel 2: Bruchschlagarbeit, 2. Mittelblock, Kern 1
Hier ist c_R = 7 Min, $n = \sqrt{(730/7) \cdot (0.62/0.45)} = 12$ und somit $\sigma^2 = 0.5$. Aus $\Delta = 0.5/\sigma$ folgt dann m = 46.

Beispiel 3: Biegefestigkeit, 1. Mittelblock, Kern 1
In diesem Fall ist $\sigma_R^2 = 3069$ sehr klein im Vergleich zu $\sigma_B^2 = 12474$, so daß schon mit kleinem n gerechnet werden kann. Mit den Zeiten aus der Tabelle 2 erhält man n = 4 und damit $\sigma^2 = 13241$, so daß nach OWEN (mit $\Delta = 0.87$) m = 30 folgt. Dabei wurde d = 100 daN/cm² verwendet.

6. Diskussion

Für eine exakte Stichprobenplanung für künftige Versuche dieser Art wäre nach den vorangegangenen Abschnitten die genaue Kenntnis der Varianzkomponenten σ_R^2 und σ_B^2 aller Merkmale (alle Qualitätseigenschaften in allen berücksichtigten Stammbereichen) und darüber hinaus bei multivariater Betrachtungsweise auch der Korrelationen zwischen den einzelnen Stammittelwerten notwendig. Um wenigstens Schätzungen all dieser vor der Hauptuntersuchung i. A. unbekannten Größen zu erhalten, müßten umfangreiche Voruntersuchungen durchgeführt werden, die den üblichen finanziellen und zeitlichen Rahmen für derartige Versuche klar sprengen würden. Denkbar wäre es aber, einen Vorversuch und eine darauf aufbauende Stichprobenplanung für Variablen durchzuführen, die sich etwa aus Bohrkernen ermitteln lassen, so z. B. für die Rohdichte, und nur nachträglich eine Abschätzung für die erreichte Testschärfe bei den übrigen Merkmalen, falls notwendig auch in mehrdimensionalen Modellen, vorzunehmen (z. B. AHRENS u. LÄUTER, 1981, S. 163 u. S. 214 f.; LÄUTER, 1978). Auch eine nachträgliche Schärfebestimmung ist für die Interpretation der Testergebnisse sehr hilfreich, wie die hier dargestellten Beispiele gezeigt haben. So haben wir erkannt, daß m = 5 Stämme sicher nicht ausreichend sind, um eine einigermaßen hohe Testschärfe zu garantieren, auch wenn dabei der Fehler 1. Art α eingehalten wird. Es war aber auch nicht sinnvoll, mehr als etwa 10 Proben je Stamm und Merkmal zu untersuchen, da sich dies auf die Testschärfe praktisch nicht mehr ausgewirkt und somit nur eine sinnlose Mehrarbeit bedeutet hätte.

7. Zusammenfassung

Aufgrund der an 10 Douglasienstämmen aus zwei unterschiedlich durchforsteten Versuchsflächen eines Douglasien-Fichten-Mischbestandes untersuchten Holzeigenschaften (Darrdichte, statische Längszugfestigkeit, statische Biegefestigkeit, statische Längsdruckfestigkeit und dynamische Bruchschlagarbeit) werden die Streuungen der einzelnen Eigenschaften sowohl innerhalb als auch zwischen den Stämmen und deren Einfluß auf die Stichprobenumfänge diskutiert.

Die Heterogenität der Varianzen aller geprüften Eigenschaften stellt ein Problem für die Planung der Probenanzahl je Stamm dar. Trotz des induktiven Vorgehens erwiesen sich die Empfehlungen für die Stichprobenumfänge aufgrund der Voruntersuchung eines zufällig ausgewählten Stammes für die Genauigkeit der Stammittelwerte in den meisten Fällen als ausreichend. Darüber hinaus zeigte sich, daß die Streuung zwischen den Stämmen und damit die Zahl der zu untersuchenden Stämme je Fläche, für den statistischen Flächenvergleich von erheblich größerer Bedeutung war.
Anhand der in der Hauptuntersuchung geschätzten Varianzkomponenten wurden nachträglich an einigen Beispielen die für vorgegebene Testschärfen notwendigen Stichprobenumfänge ermittelt, wobei in einem zweiten Schritt auch die Arbeitszeiten je Stamm und je Probekörper in die Stichprobenplanung mit einflossen.

Tabelle 2: **Der mittlere Zeitaufwand für die Anfertigung, Vermessung und Prüfung der DIN-Holzprüfkörper**

Ort		statische Längszugfestigkeit DIN 52 188	statische Längsdruckfestigkeit DIN 52 185	statische Biegefestigkeit DIN 52 186	dynamische Bruchschlagarbeit DIN 52 189	Zeitart
Sägewerkshalle	– Transport eines Stammes mit Hängekran vom Polterplatz in die Sägewerkshalle; Stamm abladen und in 3 m lange Blöcke trennen	90 Min. (1 Baum)				Baumzeit
	– Aufspannen, anreißen, schneiden am Horizontalgatter eines 3 m langen Blockes; stapeln der Bohlen bzw. Bretter (2 Personen)	(80 Min. x 4 Blöcke) = 320 Min. (1 Baum)				
	– Bretter eines 3 m-Blockes begutachten und im Querschnitt in Teile trennen	3 Min. (1 Brettabschnitt)				Probenzeit
	– Ausformen von Probestäben aus den Brettstücken der 3 m-Blockware; Stäbe anhobeln, sägen auf 22 x 22 mm Querkantenlänge und beschriften – Probestäbe auf 20 x 20 mm Querkantenlänge hobeln und gemäß DIN-Vorschrift ablängen (70 Stück)	(140 Min. : 70 Stück) = 2 Min. (1 Probestab)				
	– Anleimer anfertigen (4 Stück pro Zugprobestab)	5 Min.				
	– Anleimer auf Probestab anleimen, Zugprobestab abrichten und auf vorgeschriebene Dicke hobeln	6 Min.				
	– Zugprobe ausfräsen und beschriften	1,5 Min.				
	mittlerer Zeitaufwand pro Probe	17,5 (+5) Min.	5 Min. *	5 Min.	5 Min.	
Labor	– Feuchtemessung pro Probe	18 sec.	18 sec.	18 sec.	18 sec.	
	– Wägen pro Probe	–	30 sec.	30 sec.	30 sec.	
	– Vermessen pro Probe (radial, tangential, longitudinal)	–	40 sec.	40 sec.	40 sec.	
	– Vermessen pro Probe (radial, tangential)	30 sec.	–	–	–	
	– Festigkeitsprüfung	2 Min.	2 Min.	3 Min.	30 sec.	
	mittlerer Zeitaufwand pro Probe	2 Min. u. 48 sec.	3 Min. u. 28 sec.	4 Min. u. 28 sec.	1 Min. u. 58 sec.	
Σ	mittlerer Gesamtzeitaufwand pro Probe	ca. 21 Min.	ca. 9 Min.	ca. 10 Min.	ca. 7 Min.	

* Ausformen der Druckproben aus Brettern

Literatur

AHRENS, H., LÄUTER, J.: Mehrdimensionale Varianzanalyse. Akademie-Verlag, Berlin 1981. – COCHRAN, W. G.: Sampling Techniques. Wiley, New York 1977. – DEUTSCHES INSTITUT FÜR NORMUNG e. V. (DIN): Normung über Holz. Veuth-Verlag GmbH. Berlin-Köln, Taschenbuch 31 (1982), S. 71–98 u. 205–206. – HAPLA, F., SABOROWSKI, J.: Überlegungen zur Wahl des Stichprobenumfangs bei Untersuchungen der physikalischen und technologischen Holzeigenschaften. Forstarchiv 55 (1984), Nr. 4, 135–138. – HERRENDÖRFER, G., BOCK, J., RASCH, D.: Beiträge zur Planung des Stichprobenumfanges II. Zur Planung des Stichprobenumfanges für den Vergleich von k Mittelwerten bei hierarchischer Klassifikation. Biom. Zeitschrift 15 (1973) Heft 6, 411–415. – LÄUTER, J.: Sample size Requirements for the t-Test of MANOVA (Tables for one-way Classification). Biom. Journal 20 (1978), Nr. 4, 389–406. – OWEN, D. B.: Handbook of Statistical Tables. Addison-Wesley, Reading 1962. – RASCH, D.: Einführung in die mathematische Statistik II. VEB Deutscher Verlag der Wissenschaften, Berlin 1976. – SAUTER, U.: Untersuchung über den Durchforstungseinfluß auf die Holzeigenschaften der Douglasie. Diplomarbeit Göttingen 1985.

Verfasser: Dr. J. SABOROWSKI, Abt. Forstl. Biometrie u. Informatik, und Dr. F. HAPLA, Institut für Forstnutzung der Universität, Büsgenweg 4, 3400 Göttingen.

Summary

Retrospective Discussion of the Experiment planning for an Investigation of Wood Properties

Based on the investigation of the wood properties (specific gravity ϱ_0, tensile strength parallel to fibre, bending strength, compressive strength parallel to fibre and impact bending strength) of 10 Douglas-fir stems taken from two experimental areas with different thinning grades in a mixed stand of Douglas-fir and spruce the dispersions of the different properties are discussed within the stems as well as among them and also their influence on the sample sizes.

Variance heterogenity of all examined properties represents a problem for the planning of sample size per stem. In spite of the inductive procedure the recommendations, made for the number of samples taken from a stem chosen at random during the preliminary examination, in most cases proved to be sufficient for the exactitude of the mean value of stem.

Moreover it turned out that the dispersion among the trees and therefore the number of stems to be examined per stand was of greatest importance for the statistic stand comparsion.

Based on the estimination of variance components in the main examination the number of samples necessary for required power was established in a second step, working time per stem and per object of examination being included in the planning of the number of samples to be taken.