



Disclaimer: unless otherwise agreed by the Council of UPOV, only documents that have been adopted by the Council of UPOV and that have not been superseded can represent UPOV policies or guidance.

This document has been scanned from a paper copy and may have some discrepancies from the original document.

---

Avertissement: sauf si le Conseil de l'UPOV en décide autrement, seuls les documents adoptés par le Conseil de l'UPOV n'ayant pas été remplacés peuvent représenter les principes ou les orientations de l'UPOV.

Ce document a été numérisé à partir d'une copie papier et peut contenir des différences avec le document original.

---

Allgemeiner Haftungsausschluß: Sofern nicht anders vom Rat der UPOV vereinbart, geben nur Dokumente, die vom Rat der UPOV angenommen und nicht ersetzt wurden, Grundsätze oder eine Anleitung der UPOV wieder.

Dieses Dokument wurde von einer Papierkopie gescannt und könnte Abweichungen vom Originaldokument aufweisen.

---

Descargo de responsabilidad: salvo que el Consejo de la UPOV decida de otro modo, solo se considerarán documentos de políticas u orientaciones de la UPOV los que hayan sido aprobados por el Consejo de la UPOV y no hayan sido reemplazados.

Este documento ha sido escaneado a partir de una copia en papel y puede que existan divergencias en relación con el documento original.



TC/XX/5

ORIGINAL: englisch

DATUM: 1. August 1984

## INTERNATIONALER VERBAND ZUM SCHUTZ VON PFLANZENZÜCHTUNGEN

GENEVE

## TECHNISCHER AUSSCHUSS

Zwanzigste Tagung  
Genf, 6. und 7. November 1984

PLÄDOYER FÜR DIE ANNAHME DES KRITERIUMS DER KOMBINIERTEN ANALYSE  
ÜBER MEHRERE JAHRE FÜR DIE PRÜFUNG AUF UNTERSCHIEDBARKEIT

auf Wunsch der Technischen Arbeitsgruppe für  
Automatisierung und Datenverarbeitungsprogramme  
von ihrer Vorsitzenden ausgearbeitetes Dokument

Hintergrund

Auf ihrer ersten Tagung im Mai 1983 kam die Technische Arbeitsgruppe für Automatisierung und Datenverarbeitungsprogramme überein, dass das Kriterium der kombinierten Analyse über mehrere Jahre für die Bestimmung der Unterscheidbarkeit äusserst zufriedenstellend sei, zu einer besseren Unterscheidung führen würde und das Risiko der Begründung eines nicht-existenten Unterschieds vermindern würde. Der Technische Ausschuss bat um eine detaillierte Beschreibung der kombinierten Analyse über mehrere Jahre (siehe Absatz 34 des Berichts der Tagung vom November 1983). Dr. Weatherups Papier (siehe die Anlage zu diesem Dokument) enthält eine solche Beschreibung dieser Analyse und bewertet die Verwendung der Kriterien über mehrere Jahre im Vergleich mit dem gegenwärtigen UPOV-Unterscheidbarkeitskriterium und dem t-score. Die folgenden Bemerkungen fassen die Argumente für die Annahme der kombinierten Analyse über mehrere Jahre zusammen. Tabellen, auf die in diesem Dokument hingewiesen wird, finden sich in der Anlage zu diesem Dokument (identisch mit Dokument TWC/II/5).

Gegenwärtiges Kriterium (2 x 1 %-Regel)

Das gegenwärtige Kriterium für die Unterscheidbarkeit verlangt, dass der Unterschied zwischen zwei Sorten bei einem Signifikanzniveau von 1 % statistisch signifikant ist und in wenigstens zwei von drei Jahren bei einem oder mehreren wichtigen Merkmalen auftritt. Dies wird die 2 x 1 %-Regel genannt, und die technischen Sachverständigen, die sie angenommen haben, waren sich sehr wohl bewusst, dass es wichtig ist, dass sich die Wiederholbarkeit von signifikanten Sortenunterschieden über mehrere Jahre feststellen lässt.

An der 2 x 1 %-Regel wurde kritisiert, dass ein Unterschied zwischen Sorten, der das 1 %-Signifikanzniveau nicht ganz erreicht, nicht in einer grossen Masse zur Feststellung der Unterscheidbarkeit beiträgt als zum Beispiel ein fehlender Unterschied in einem einzigen Jahr. Wenn die Unterschiede zwischen Sorten in jedem der drei Jahre gleichgerichtet sind, einer von ihnen

aber ein Signifikanzniveau von 1 %, die zwei anderen ein Signifikanzniveau von 5 % aufweisen, so würde dies unter der gegenwärtigen 2 x 1 %-Regel nicht als schlüssiger Beweis für die Unterscheidbarkeit ausreichen. Eine intuitive Interpretation von drei Ergebnisse dieser Art würde es nahelegen, dass die Sorten unterscheidbar sind und dass die Informationen aus allen drei Jahren berücksichtigt werden sollten.

#### t-score-Methode

Die t-score-Methode berücksichtigt bis zu einem gewissen Grad die obenerwähnte Kritik, indem sie ein System für die Berechnung verwendet, das Ergebnisse aus allen drei Jahren verwendet, erlaubt jedoch nicht, dass ein in einem Jahr sehr extremes Ergebnis die Beweise der anderen beiden Jahre aufwiegen kann. Die Tabelle 3 zeigt ein Beispiel eines Falles, in dem die t-score-Methode den intuitiven Eindruck der obenerwähnten Ergebnisse bestätigt.

#### Risiko falscher Entscheidungen

Beide, das 2 x 1 %-System sowie das t-score-System, bestimmen Unterschiede zwischen Sorten auf der Grundlage der Variationen des Parzellenfehlers innerhalb der Prüfungen. Dies berücksichtigt nicht die Variationen der Sorten über mehrere Jahre. Es kann dargelegt werden, dass einige Merkmale der Pflanzensorten über längere Zeiträume sehr viel weniger beständig sind als andere. Dort, wo Entscheidungen auf Unterscheidbarkeit sich auf Merkmale gründen, die von Jahr zu Jahr beständig sind (niedriger Wert von  $\lambda$  in Tabelle 2), ist das Risiko, dass wiederholte Prüfungen in späteren Jahren zu einer unterschiedlichen Entscheidung führen könnten, gering. Entscheidungen, die sich auf weniger beständige Merkmale gründen, bringen ein grösseres Risiko mit sich, dass sie sich später nicht bestätigen lassen. Sowohl bei der Verwendung der 2 x 1 %-Regel als auch bei der t-score-Methode ist das Risiko der Entscheidung unterschiedlich gross, je nach Beständigkeit des Merkmals über die Jahre, auf das sich die Unterscheidbarkeitsentscheidung gründet. Das ist unbefriedigend sowohl für die prüfende Behörde als auch für den Züchter.

Die Verwendung des Kriteriums der Analyse über mehrere Jahre, das die Unterschiede auf der Grundlage der Variationen über die Jahre misst, führt zu einem gleichbleibenden Risiko, unabhängig davon, welches Merkmal die Grundlage für die Entscheidungen auf Unterscheidbarkeit darstellt. Bei Verwendung der vorgeschlagenen Kriterien über die Jahre würde die prüfende Behörde unabhängig von den Merkmalen, die zur Unterscheidung herangezogen werden, ein 1 %-Risiko eingehen, zwei Sorten fälschlicherweise als unterscheidbar zu erklären, die in der Tat nicht unterscheidbar sind. Bei Verwendung der 2 x 1 %-Regel könnte dieses Risiko weniger als 1 % für Merkmale sein, die über die Jahre beständig sind, und bis zu 9 % betragen, wenn die Entscheidung sich auf das am wenigstens beständige Merkmal gründet (siehe Prüfers Risiko in Tabelle 2).

Eine Statistik der F-Quotientenprüfung ist mit dem Kriterium über mehrere Jahre verbunden worden, um jegliche extreme Variation in den Unterschieden zwischen einer Kandidatensorte und einer Kontrollsorte aufzudecken, wobei die allgemeine Variation Sorte x Jahre über die drei Jahre der Prüfung verglichen wird. Ein signifikanter F-Quotient kann auf eine nicht beständige Kandidatensorte hinweisen.

#### Schlussfolgerung

Das Kriterium der kombinierten Analyse über mehrere Jahre hat die folgenden erwünschten Vorzüge, die die getroffenen Entscheidungen verlässlicher machen würden:

- von ihrer Definition her sieht es eine Prüfung auf Wiederholbarkeit von Sortenunterschieden über die Jahre vor
- es verwendet alle verfügbaren Informationen, die in den Daten aller Jahre enthalten ist
- das Risiko, falsche Entscheidungen zu treffen, ist für alle Merkmale konstant.

TC/XX/5  
Seite 3

Die Technische Arbeitsgruppe für Automatisierung und Datenverarbeitungsprogramme kam auf ihrer Tagung in La Minière in diesem Jahr im Prinzip überein, dass das Kriterium der kombinierten Analyse über mehrere Jahre das am besten geeignete Kriterium für die Unterscheidbarkeit darstellt, das zur Verfügung steht. Die Technische Arbeitsgruppe empfahl, dass die Verbandsstaaten das Kriterium während des kommenden Jahres weiter bewerten, damit detaillierte Vorschläge für seine Einführung auf der nächsten Tagung für die Vorlage an den Technischen Ausschuss beschlossen werden können.

[Anlage folgt]

Beschreibung und Bewertung der Kriterien der kombinierten Analyse  
über mehrere Jahre auf Unterscheidbarkeit zwischen Futterpflanzensorten

von S.T.C. Weatherup

Biometrische Abteilung, Department für Landwirtschaft,  
Nordirland, Vereinigtes Königreich

Einführung

Auf ihrer ersten Tagung kam die Technische Arbeitsgruppe für Automatisierung und Datenverarbeitungsprogramme der UPOV überein, dass "die kombinierte Analyse über mehrere Jahre äusserst zufriedenstellend sei, zu einer besseren Unterscheidung führen und das Risiko der Erstellung von Unterschieden, die nicht beständen, reduzieren würde". In dem vorliegenden Papier werden die Unterscheidungskriterien der kombinierten Analyse über mehrere Jahre beschrieben und anhand von tatsächlichen Prüfungsergebnissen verglichen, um zum Verständnis der Kriterien der kombinierten Analyse über mehrere Jahre beizutragen und um aufzuzeigen, wie sie die gegenwärtigen Kriterien verbessern. Die Methoden werden unter Verwendung von Daten aus einer Serie von Prüfungen von Deutschem Weidelgras (diploid) beschrieben, die in Crossnacreevy im Vereinigten Königreich während der Jahre 1979 bis 1981 durchgeführt worden sind. Die Tabelle 1 enthält eine Liste von Merkmalen, die in diesen Prüfungsserien gemessen wurden. Weitere Informationen über Einzelheiten der Kriterien der kombinierten Analyse über mehrere Jahre können bei Patterson H.D. und Weatherup S.T.C. (1984) nachgelesen werden.

Prüfungskriterien

(a) Die 2 x 1 %-Methode

Das gegenwärtige Kriterium für die Unterscheidbarkeit von Sorten beruht auf getrennten t-Tests in drei Jahren zwischen der Kandidatensorte und jeder anderen Sorte. Diese Tests verwenden die t-Werte, die durch

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{2} \text{ SE } (\bar{x})}$$

definiert werden, wobei  $\bar{x}_1$  und  $\bar{x}_2$  die Mittelwerte eines gemessenen Merkmals der beiden zu vergleichenden Sorten darstellen und  $\text{SE}(\bar{x})$  den Standardfehler eines Sortenmittelwerts, der aus dem Mittelquadrat des Parzellenfehlers errechnet wird. Das Signifikanzniveau eines errechneten t-Wertes kann aus den t-Tabellen mit Hilfe der Freiheitsgrade des Mittelquadrats des Parzellenfehlers erhalten werden. Für jeden Vergleich zwischen einer Kandidatensorten und einer anderen Sorte sind drei solcher Prüfungen aus drei Jahren verfügbar, in denen die Kandidatensorte in den Prüfungen angebaut wurde. Nach diesem Kriterium wird die Kandidatensorte als von einer anderen Sorte unterscheidbar angesehen, wenn wenigstens in zwei von drei Jahren der t-Wert bei gleichem Vorzeichen des Unterschieds mit einem Signifikanzniveau von 1 % signifikant ist.

(b) Die "t-score"-Methode

Die 2 x 1 %-Methode kann aufgrund der Tatsache kritisiert werden, dass ein Unterschied innerhalb eines Jahres, der das Signifikanzniveau von 1 % nicht erreicht, zu der Trennung eines Sortenpaares nicht mehr beiträgt als ein nicht bestehender Unterschied oder sogar ein nicht signifikanter Unterschied mit gegenteiligem Vorzeichen. Zum Beispiel würden drei Unterschiede mit gleichem Vorzeichen, von denen einer signifikant ist mit einem Signifikanzniveau von 1 % und die anderen mit einem Signifikanzniveau von 5 % nicht für die Unterscheidbarkeit ausreichen. Die "t-score"-Methode wurde daher eingeführt, um diesen Mangel auszugleichen. Bei dieser Methode wird der nach der Formel (1) errechnete t-Wert in einen "t-score" umgerechnet unter Verwendung der Konstanten  $k_1$  und  $k_2$ , wobei diese jeweils die tabellierten t-Werte zum Signifikanzniveau von 5 % und 0,1 % darstellen. Die Umrechnung der t-Werte zu den "t-scores" wird dann wie folgt definiert:

- (i) "t-score" = 0 wenn  $-k_1 < t < k_1$   
(ii) "t-score" = t-Wert wenn  $k_1 \leq t \leq k_2$  oder  $-k_2 \leq t \leq -k_1$   
(iii) "t-score" =  $k_2$  wenn  $t > k_2$   
(iv) "t-score" =  $-k_2$  wenn  $t < -k_2$

Diese Umwandlungsregeln sind in Abbildung 1 erläutert. Unter Verwendung der "t-scores" sind zwei Sorten unterscheidbar, wenn die absolute Summe ihrer "t-scores" über drei Jahre den kritischen Wert 5,2 übersteigt, der gleich dem zweifachen t-Wert aus der Tabelle bei 1 %igem Signifikanzniveau mit einer grossen Anzahl von Freiheitsgraden ist. Wie bei dem Kriterium  $2 \times 1$  % erfordert die "t-score"-Methode mehr als einen einzelnen grossen t-Wert für die Unterscheidbarkeit, jedoch braucht der bestätigende Beweis nicht so deutlich zu sein. Daher sind drei Ergebnisse bei einem Signifikanzniveau von 5 %, unter der Voraussetzung, dass sie mit gleichem Vorzeichen erscheinen, ausreichend, um Unterscheidbarkeit sicherzustellen.

Dort, wo eine Prüfungsbehörde einen zweiten Versuchsort besitzt, ist zusätzlich zu fordern, dass für den Fall des Auftretens eines "t-score" über 5,2 an dem einen Ort an dem zweiten Ort ein ausreichend bestätigender Beweis erforderlich ist, um für die Unterscheidbarkeit den gesamten "t-score"-Wert 8,5 für beide Orte zu erhalten.

(c) Kriterium der kombinierten Analyse über mehrere Jahre

Beide der oben erwähnten Kriterien basieren auf dem Mittelquadrat des Parzellenfehlers und berücksichtigen daher nicht die Variation der Sorte über mehrere Jahre. Da Messungen an Sorten bei einigen Merkmalen weniger gleichgerichtet sind als bei anderen Merkmalen, bewirkt die Verwendung dieser Kriterien die Annahme von Standards, die zwischen den Merkmalen variieren. Ein Fehlen der Gleichgerichtetheit wird durch grosse Werte von  $\lambda$  angegeben, wobei

$$\lambda^2 = \frac{\text{Sorten} \times \text{Mittelquadrat der Jahre}}{\text{Mittelquadrat des Parzellenfehlers}}$$

ist. Gleichgerichtetheit wird durch Werte von  $\lambda$  nahe bei 1 angegeben. Werte von  $\lambda$  eines Beispieldatensatzes sind in Tabelle 1 wiedergegeben und zeigen einen Bereich von 1,21 bis 2,41. Das Kriterium der Analyse über mehrere Jahre berücksichtigt die oben erwähnte Kritik durch die Verwendung des Wertes Sorte  $\times$  Mittelquadrat der Jahre, zur Berechnung des Fehlers anstelle des Mittelquadrats des Parzellenfehlers. Zwei Sorten sind unterscheidbar, wenn der absolute Wert aus

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{2} \text{ SE}(\bar{x})}$$

grösser ist als der kritische Punkt bei 1 % in der t-Verteilung, wobei  $\bar{x}_1$  und  $\bar{x}_2$  die Mittelwerte der Sorten über drei Jahre darstellen und  $\text{SE}(\bar{x})$  den Standardfehler eines Sortenmittels, der aus dem Wert Sorte  $\times$  Mittelquadrat der Jahre errechnet wurde.

Unter Verwendung dieses Kriteriums stellt der F-Quotient, der wie folgt definiert ist

$$F = \frac{\text{Mittelquadrat der Sorten}}{\text{Sorten} \times \text{Mittelquadrat der Jahre}},$$

einen Masstab für die Unterscheidungskraft eines Merkmals dar. So zeigt Tabelle 1, dass die Merkmale 11 und 20 für die Unterscheidung am besten geeignet, und dass die Merkmale 4 und 19 für die Unterscheidung am wenigsten geeignet sind.

Vergleich der Kriterien für die Annahmewahrscheinlichkeit

Ein geeignetes Mass für die Effizienz der drei oben erwogenen Methoden ist die Wahrscheinlichkeit, zwei hypothetische Sorten als unterscheidbar zu deklarieren, deren Unterschied anhand eines bestimmten Merkmals bekannt ist. Die Wahrscheinlichkeit wird Annahmewahrscheinlichkeit genannt. Der einfachste Fall ist derjenige, bei dem der tatsächliche Unterschied in einem Merkmal mit Null angenommen wird, d.h. die Sorten sind bei diesem Masstab identisch. Aufgrund von

Stichprobentehlern werden die tatsächlichen Unterschiede zwischen den Mittelwerten dieser Sorten, die in den Prüfungen erzielt wurden, nicht Null, sondern um Null verteilt sein, mit einer Varianz, die aus den Prüfungsdaten über drei Jahre errechnet werden kann. Die Annahmewahrscheinlichkeit einer Kandidatensorten kann unter Verwendung dieser Verteilung errechnet werden. In diesem Fall wird sie das Prüfferrisiko genannt, da sie die Wahrscheinlichkeit darstellt, mit der eine nicht unterscheidbare Sorte angenommen werden kann. Werte für diese Risiken in den drei Kriterien bei den gemessenen Merkmalen der ausgewählten Prüfung sind ebenfalls in Tabelle 1 wiedergegeben. Sie zeigen, dass für beide, sowohl für die 2 x 1 %-Kriterien als auch für die Kriterien des "t-score" je nach ihrer Gleichgerichtetheit Unterschiede im Prüfferrisiko zwischen den Merkmalen angegeben wurden. So beträgt für das Merkmal 5, das über die Jahre eine mangelnde Gleichgerichtetheit (d.h. ein grosses  $\lambda$ ) aufgewiesen hat, die Wahrscheinlichkeit für die Annahme einer Sorte, deren tatsächlicher Mittelwert mit demjenigen einer anderen Sorte identisch ist, 9,2 % für die 2 x 1 %-Kriterien bzw. 14,1% für die Kriterien der "t-scores". Ebenfalls sind für das Merkmal 11, das über die Jahre gleichgerichtet war (d.h. kleines  $\lambda$ ) die entsprechenden Annahmewahrscheinlichkeiten nur 0,2 % und 0,5 %. Im Gegensatz dazu sind die Annahmewahrscheinlichkeiten des Kriteriums der kombinierten Analyse über mehrere Jahre für alle Merkmale konstant 1 %, sofern der tatsächliche Unterschied Null beträgt.

Die gleiche Methode kann zur Bestimmung der Annahmewahrscheinlichkeiten bei anderen angenommenen wahren Unterschieden verwendet werden, um bei gegebenen Werten von  $\lambda$  eine Kurve für die Annahmewahrscheinlichkeit gegenüber tatsächlichen Unterschieden zu zeichnen. Solch eine Kurve ist bekannt als Operations-characteristic (OC) des Merkmals. Die Kurven sind in Abbildung 2 für die drei oben behandelten Merkmale aufgezeigt und geben die Annahmewahrscheinlichkeiten gegenüber standardisierten Unterschieden für  $\lambda$ -Werte von 1, 2 und  $\infty$  wieder. Die Neigung eines OC im Mittelabschnitt stellt ein Mass für die Wirksamkeit des Prüfungsschemas dar. Je steiler die Neigung, desto effizienter ist das Schema. So kann daraus ersehen werden, dass die "t-score"-Methode effizienter ist als das 2 x 1 %-Kriterium, auf der anderen Seite aber weniger streng, da sie bei gegebenem  $\lambda$  höhere Annahmewahrscheinlichkeiten liefert. Bei einer Erhöhung von  $\lambda$  gleichen sich die OC-Kurven der 2 x 1 %-Methode und der Methode der abgewandelten t-Werte aneinander an und werden weniger steil. Die Effizienz ist dann gering, und das Prüfferrisiko sehr hoch. Daher ist, obwohl beide Methoden Gleichgerichtetheit über die Jahre voraussetzen, der festgesetzte Standard klein und hat keine Beziehung zu dem für die Variation zwischen den Parzellen angenommenen 1 %-Standard. Im Gegensatz dazu bleibt die OC-Kurve der Kriterien der kombinierten Analyse über mehrere Jahre für alle Werte von  $\lambda$  konstant und garantiert damit für jeden definierten tatsächlichen Unterschied die gleiche Annahmewahrscheinlichkeit, unabhängig von der Gleichgerichtetheit des Merkmals.

#### Homogenität der Sorten x Jahresvarianz

Als Ergebnis des vorangegangenen Absatzes wird empfohlen, die Kriterien der kombinierten Analyse über mehrere Jahre den Methoden 2 x 1 % oder des "t-score" vorzuziehen, allerdings mit einer Einschränkung. Das Sorten x Jahres-Mittelquadrat, das bei dem Kriterium der kombinierten Analyse über mehrere Jahre verwendet wird, ist ein zusammengesetzter Wert, der aus einer grossen Zahl von Sortenvergleichen errechnet wird und daher für einen bestimmten Vergleich nicht zulässig sein könnte. Es muss daher sichergestellt werden, dass der spezifische Wert innerhalb eines Paares aus Sorte x Mittelquadrat der Jahre bei zwei Freiheitsgraden nicht grösser ist als der zusammengesetzte Wert aus Sorten x Mittelquadrat der Jahre, in dem z.B. der  $F_2$ -Quotient errechnet und auf seine Signifikanz geprüft wird. Das Kriterium der kombinierten Analyse über mehrere Jahre sollte in den Fällen mit Vorsicht verwendet werden, in denen ein signifikanter  $F_2$ -Quotient auftritt.

#### Beispiele

In vielen Fällen stimmen die Ergebnisse der 2 x 1 %-Methode und der Methode der Prüfungen über mehrere Jahre überein; es gibt jedoch Fälle, in denen sich die Schlussfolgerungen unterscheiden. Beispiele aus den ausgewählten Daten sind in den Tabellen 2 bis 5 wiedergegeben. Im folgenden werden hierzu einige Bemerkungen gemacht.

(a) Tabelle 2. Die t-Werte des Merkmals 20 bei Sorte A gegen Sorte B sind bei dem Signifikanzniveau von 1 % nur in einem von drei Jahren signifikant unterschiedlich, und die Sorten werden daher bei Verwendung der 2 x 1 %-Kriterien in diesem Merkmal als nicht unterscheidbar angesehen. Der t-Wert für Merkmal 20 ist klein, und es ist daher schwierig, mit Hilfe der 2 x 1 %-Kriterien eine Unterscheidbarkeit festzustellen. Einer der t-Werte ist signifikant mit einem Signifikanzniveau von 0,1 %; die anderen beiden liefern, obgleich sie nicht den hohen Standard, der für das 2 x 1 %-Kriterium erforderlich ist, erfüllen, starke bestätigende Beweise. Das Kriterium 1 % über die Jahre wird einfach erzielt, und der Wert von  $F_2$  ist niedrig. Dem Kriterium des "t-score" ist ebenfalls Genüge getan. Es kann daraus geschlossen werden, dass die Sorten unterscheidbar sind, wobei die Sorte A nach dem Schnitt kleinere Pflanzen als die Sorte B liefert.

(b) Tabelle 3. Weder das 2 x 1 %-Kriterium noch das Kriterium des "t-score" liefert in Merkmal 11 genügend Beweiskraft zur Unterscheidung zwischen den Sorten C und D, jedoch weist das Kriterium der Analyse über mehrere Jahre auf Unterscheidbarkeit hin;  $F_2$  ist kleiner als 1 und zeigt damit an, dass es keine extreme Variation Sortenpaar x Jahr gibt. Nur einer der individuellen t-Werte ist bei einem Signifikanzniveau von 1 % signifikant, die anderen beiden liefern jedoch bestätigende Beweise. Das 2 x 1 %-Kriterium berücksichtigt nicht den signifikanten t-Wert für das Jahr 1981. Hier ist  $\lambda$  wieder klein, und es erscheint vertretbar festzustellen, dass die zwei Sorten unterscheidbar sind.

(c) Tabelle 4. Die Unterschiede im Merkmal 5 bei Sorte E gegen Sorte F sind in zwei Jahren signifikant bei einem Signifikanzniveau von 1 %, und daher sind die Sorten bei Anwendung der 2 x 1 %-Kriterien sowie der t-Wert-Kriterien unterscheidbar; diese Schlussfolgerung wird jedoch nicht durch die Analyse über mehrere Jahre bestätigt. Der  $\lambda$ -Wert für Merkmal 5 ist gross und folglich sind die Prüfungen auf 2 x 1 % und auf den "t-score" ungewöhnlich milde.

(d) Tabelle 5. Alle drei Kriterien führen zu der Schlussfolgerung, dass die Sorte G von der Sorte H unterscheidbar ist, da ihre Pflanzen 30 Tage nach dem Ährenschieben (Merkmal 17) längere Halme aufweisen. Die Prüfung auf Gleichgerichtetheit von Jahr zu Jahr weist jedoch auf die Möglichkeit hin, dass die Schlussfolgerung falsch sein könnte. Die Ergebnisse im Jahre 1979 widersprechen den Ergebnissen der Jahre 1980 und 1981 und der  $F_2$ -Quotient ist signifikant bei einem Signifikanzniveau von 1 %.

#### Multivariates Kriterium

Keines der vorangehenden Kriterien berücksichtigt die substantiellen Korrelationen, die zwischen den Merkmalen bestehen. Dieser Mangel kann durch die Verwendung von Mahalanobis verallgemeinertem Abstand  $D^2$  als einem Masstab für die Unterscheidbarkeit behoben werden; dieser Abstand ist für ein Sortenpaar definiert als  $D^2 = d^T W^{-1} d$ , wobei  $d$  den Vektor der Unterschiede zwischen den Mittelwerten des Sortenpaares über das Jahr für alle Merkmale darstellt,  $d^T$  ist die transponierte Form von  $d$  und  $W$  die Kovarianz Matrix, berechnet aus dem Wert Sorten x Mittelquadrat der Jahre und den Kreuzprodukten für alle Merkmale. So bildet die Matrix  $W$  den analogen multivariaten Ausdruck des Sorten x Jahres-Mittelquadrates, der in dem Kriterium der Analyse über mehrere Jahre verwendet wird.

Der kritische Wert für  $D^2$  wird angegeben durch

$$\frac{2p(m-1)(n-1)}{m(mn-m-n-p+2)} F$$

wobei  $m$  die Anzahl der Jahre darstellt,  $n$  die Anzahl Sorten,  $p$  die Anzahl Merkmale und  $F$  den F-Quotient bei einer Signifikanzzahl von 1 % mit  $p$  und  $mn - m - n - p + 2$  Freiheitsgraden.

Die Tabelle 7 gibt t-Werte aus dem Vergleich zwischen dem Sortenpaar X und Y anhand eines vollständigen Satzes von Merkmalen wieder. Daraus kann gefolgert werden, dass das Kriterium 2 x 1 % an keinem einzelnen Merkmal erfüllt wurde. Obwohl das "t-score"-Kriterium beim Merkmal 4 erfüllt wurde, war dies in diesem Fall nicht ausreichend für die Unterscheidbarkeit, da die Sorten an zwei Orten verglichen wurden und für dies Merkmal am zweiten Ort kein unterstützender Beweis gegeben war. Der  $D^2$ -Wert betrug jedoch 36,43 und

erfüllte somit für diesen Satz von Prüfungen den kritischen  $D^2$ -Wert von 21,5; die Sorten konnten daher aufgrund einer multivariaten Basis als unterscheidbar angesehen werden. Eine Prüfung der Zeichen der einzelnen t-Werte zeigt Gleichgerichtetheit für viele der Merkmale, z.B. 4, 5, 14, 17, 20 und 24. Die beiden ersten Ergebnisse weisen auf eine Tendenz, wonach die Sorte X im Frühjahr aufrechter und höher als die Sorte Y ist, während die Ergebnisse von den Merkmalen 14, 17 und 24 darauf hinweisen, dass X kürzere Blätter als Y hat.

#### Schlussfolgerung

Es wird empfohlen, das 2 x 1  $\chi^2$ -Kriterium und das "t-score"-Kriterium umgehend durch das Kriterium der kombinierten Analyse über mehrere Jahre zu ersetzen, um zu Unterscheidbarkeitsentscheidungen zu gelangen, die sich auf Variationen über die Jahre gründen und so eine grössere Wahrscheinlichkeit aufweisen, langfristig zutreffend zu bleiben. Weiterhin wird empfohlen, die Mahalanobis  $D^2$ -Abstände als Unterscheidbarkeitskriterium zur Unterscheidung zwischen anderen Sortenpaaren zu verwenden, für die es kein einzelnes trennendes Merkmal gibt, die jedoch in ähnlichen Merkmalen mehrere gleichgerichtete Unterschiede aufweisen.

#### Referenzen

Patterson, H.D. und Weatherup, S.T.C. (1984). Statistical criteria for distinctness between varieties of herbage crops. J. Agric. Sci., Camb., 102, 59-68.

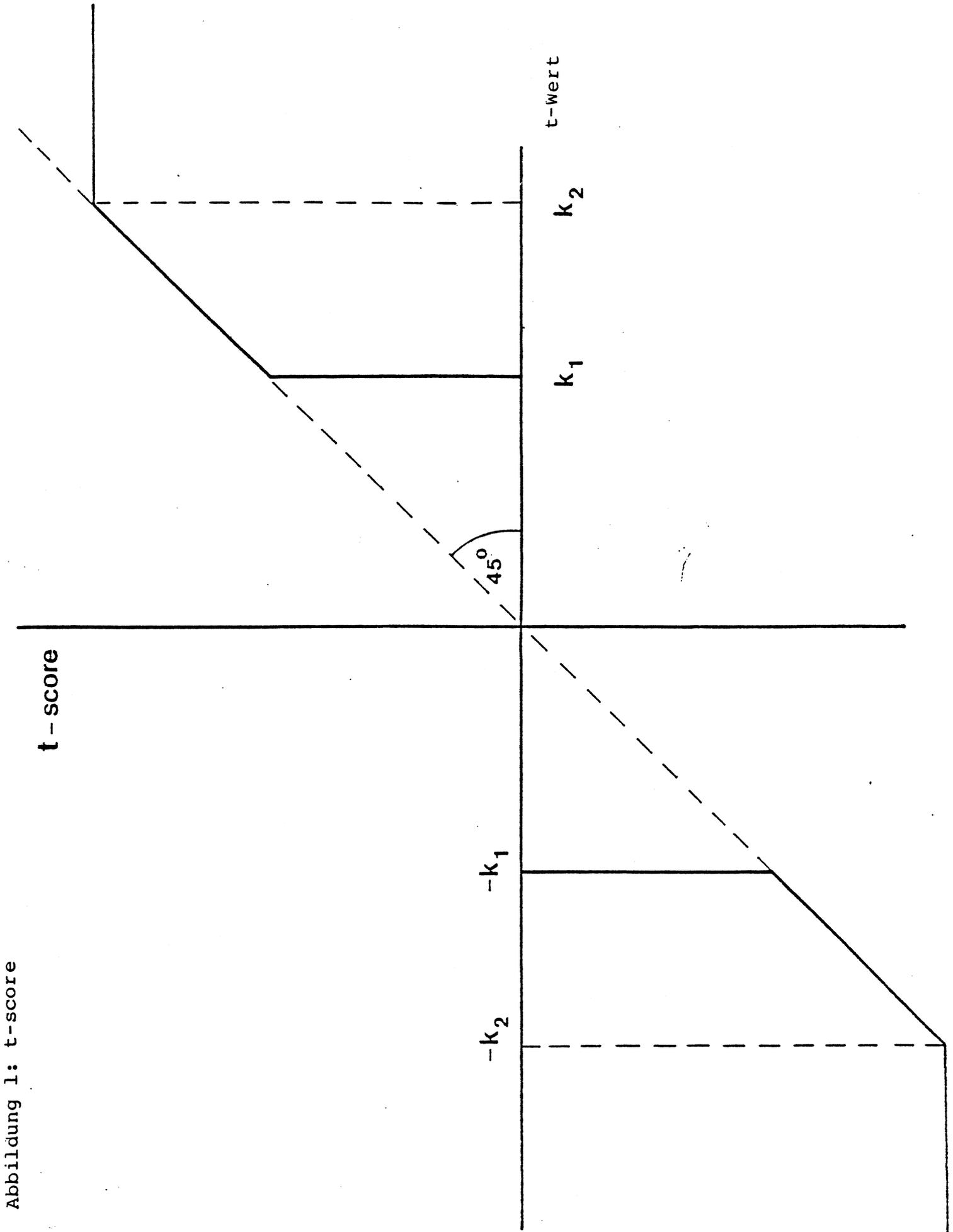


Abbildung 1: t-score

Abbildung 2: Verwandte Merkmale für Unterscheidungskriterien

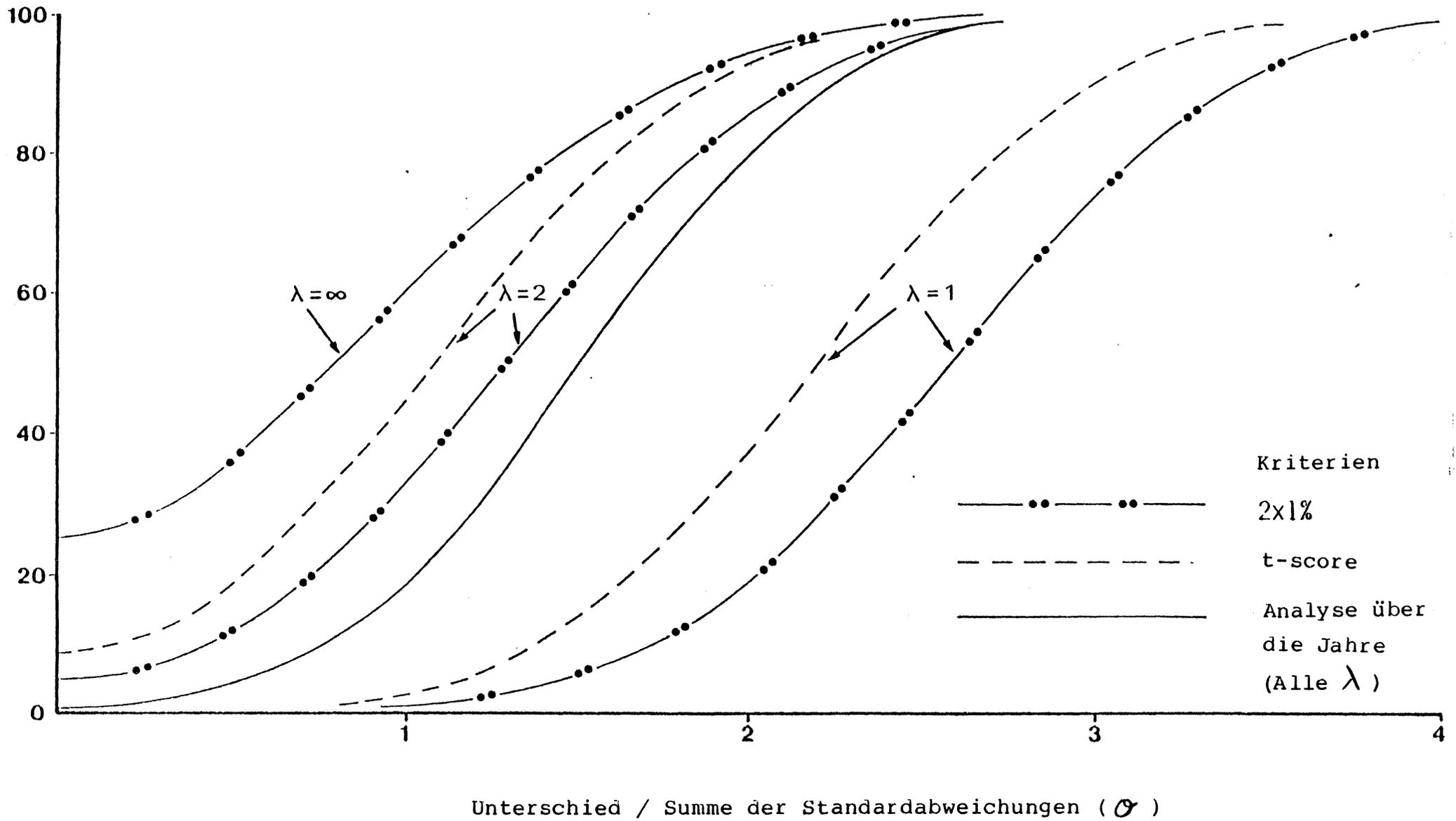


TABELLE 1

Definition der gemessenen Merkmale

Merkmalsnummer	Definition
4	Wachstumswinkel im Aussaatjahr
5	Höhe der aufgerichteten Blätter im Frühjahr (cm)
8	Datum des Ährenschiebens (Tage ab 1. März)
10	Natürliche Pflanzenhöhe zum Zeitpunkt des Ährenschiebens (cm)
11	Pflanzenbreite zum Zeitpunkt des Ährenschiebens (cm)
14	Länge des Spitzenblatts zum Zeitpunkt des Ährenschiebens (cm)
15	Breite des Spitzenblatts zum Zeitpunkt des Ährenschiebens (cm)
17	Länge des Halmes 30 Tage nach dem Zeitpunkt des Ährenschiebens (cm)
*19	Anzahl Blütenstände je Pflanze (in der 0 bis 9-Skala)
*20	Höhe der Pflanze nach dem Schnitt (cm)
24	Länge der Ähre (cm)

\* Alle Pflanzen jeder Sorte werden zu einem bestimmten Zeitpunkt entsprechend ihrem erfassten Datum des Ährenschiebens geschnitten. Die Merkmale 19 und 20 werden an den Pflanzen des Nachwuchses 8 Wochen nach dem Schnitt gemessen.

TABELLE 2

Statistische Masszahlen aus dem Beispieldatensatz

(Prüfungen an frühem Deutschen Weidelgras (diploid) in Crossnacreevy,  
Vereinigtes Königreich, 1979-81)

Merkmal	Mittelquadrat pro Parzelle						Prüferrisiko (%)		
	Jahre (Y)	Sorten (V)	Jahre x Sorten (VxY)	Parzellenmittel (E)	$F_1$ = $V/(VxY)$	$\lambda$ = $\sqrt{(VxY)/E}$	2x1%	t-score	Analyse über mehrere Jahre
4	570.34	214.03	28.41	15.29	7.5	1.36	0.5	1.3	1.0
5	2678.61	351.99	25.50	4.44	13.8	2.41	9.2	14.1	1.0
8	32317.11	351.85	18.84	5.21	18.7	1.90	3.9	7.2	1.0
10	940.04	777.98	44.72	12.79	17.4	1.87	3.6	6.8	1.0
11	2461.98	430.37	19.18	13.16	22.4	1.21	0.2	0.5	1.0
14	30.79	61.95	3.83	1.35	16.2	1.68	2.1	4.4	1.0
15	0.02	2.30	0.18	0.11	13.0	1.25	0.2	0.7	1.0
17	1364.30	533.30	32.91	16.06	16.2	1.41	0.6	1.6	1.0
19	10.96	3.91	0.41	0.17	9.6	1.53	1.2	2.7	1.0
20	3327.77	396.55	19.55	12.96	20.3	1.23	0.2	0.6	1.0
24	412.41	38.84	3.13	1.60	12.4	1.40	0.6	1.6	1.0

TABELLE 3

Sortenpaar AvB

Merkmal 20 ( $\lambda = 1.23$ )

Sorte	Jahr			Mittelwert
	1979	1980	1981	
A	33.9	41.3	36.1	37.1
B	26.6	36.1	31.0	31.2
SE	1.25	1.69	1.43	1.04
t-Wert	4.13**	2.18*	2.52*	3.98 <sup>‡</sup>
2x1%		2.59		
t-score		8.01 <sup>‡</sup>		
Analyse über mehrere Jahre		3.98 <sup>‡</sup>		
F <sub>2</sub>		0.23		

\* Signifikant bei einer Signifikanzzahl von 5 %

\*\* Signifikant bei einer Signifikanzzahl von 1 %

‡ Unterscheidbarkeitskriterium erfüllt

TABELLE 4

## Sortenpaar CvD

Merkmal 11 ( $\lambda = 1.21$ )

Sorte	Jahr			Mittelwert
	1979	1980	1981	
C	58.3	63.1	56.5	59.3
D	52.3	58.7	52.3	54.4
SE	1.40	1.50	1.53	1.04
t-Wert	3.03**	2.07*	1.94	3.33‡
2x1%		2.59		
t-score		8.10‡		
Analyse über mehrere Jahre		3.33‡		
F <sub>2</sub>		0.15		

\* Signifikant bei einer Signifikanzzahl von 5 %

\*\* Signifikant bei einer Signifikanzzahl von 1 %

‡ Unterscheidbarkeitskriterium erfüllt

TABELLE 5

Sortenpaar EvF  
Merkmal 5 ( $\lambda = 2.41$ )

Sorte	Jahr			Mittelwert
	1979	1980	1981	
E	27.1	35.4	37.4	33.3
F	26.6	31.9	32.4	30.3
SE	0.70	0.86	0.99	1.19
t-Wert	0.51	2.88**	3.57**	1.78
2x1%		5.18 <sup>‡</sup>		
t-score		6.19 <sup>‡</sup>		
Analyse über mehrere Jahre		1.78		
F <sub>2</sub>		0.61		

\* Signifikant bei einer Signifikanzzahl von 5 %

\*\* Signifikant bei einer Signifikanzzahl von 1 %

= Unterscheidbarkeitskriterium erfüllt

TABELLE 6

Sortenpaar EvF  
Merkmal 17 ( $\lambda = 1.41$ )

Sorte	Jahr			Mittelwert
	1979	1980	1981	
E	85.2	87.0	88.9	87.0
F	87.6	76.8	76.7	80.4
SE	1.37	1.72	1.78	1.34
t-Wert	-1.24	4.19**	4.85**	3.49 <sup>‡</sup>
2x1%		5.18 <sup>‡</sup>		
t-score		6.63 <sup>‡</sup>		
Analyse über mehrere Jahre		3.49 <sup>‡</sup>		
F <sub>2</sub>		6.54**		

\* Signifikant bei einer Signifikanzzahl von 5 %

\*\* Signifikant bei einer Signifikanzzahl von 1 %

‡ Unterscheidbarkeitskriterium erfüllt

TABELLE 7

t-Werte der Vergleiche zwischen den Sorten X und Y für  
alle Merkmale für 3 Jahre

(t-Werte sind positiv, wenn X grösser als Y ist)

Merkmal	Jahr			t-score	Analyse über mehrere Jahre
	1981	1982	1983		
4	2.40*	4.12***	0.11	5.77	*
5	0.09	2.04*	0.31	2.04	NS
8	-0.25	-3.78***	-1.26	-3.37	NS
10	-0.72	0.59	-1.95	0.00	NS
11	-0.54	0.46	1.01	0.00	NS
14	-1.96	-3.20**	-1.26	-3.20	*
15	1.40	0.31	2.28*	2.28	NS
17	-0.67	-0.29	-0.68	0.00	NS
19	0.34	-0.49	0.80	0.00	NS
20	1.93	0.56	2.00*	2.00	NS
24	-1.83	-0.73	-0.77	0.00	NS

Mahalanobis verallgemeinerter Abstand,  $D^2 = 36.43^{**}$

\*  $P < 0.05$

\*\*  $P < 0.01$

\*\*\*  $P < 0.001$

[Ende der Anlage und des Dokuments]