

# 1 **Reiseführer der gemischten Modelle für randomisierte**

## 2 **Experimente**

3 [engl. Fassung erschienen in Journal of Agronomy and Crop Science 189 (2003), 310-322]

4

5 H. P. Piepho, A. Büchse, K. Emrich

6

7 Adresse der Autoren: Institut für Pflanzenbau und Grünland, Universität Hohenheim, 70599

8 Stuttgart, (e-mail des ersten Autor: piepho@uni-hohenheim.de)

9

## 10 **Zusammenfassung**

11

12 Geplante Versuche in den Pflanzenwissenschaften umfassen in der Regel mehrere zufällige

13 Streuungsursachen. Einschlägige Beispiele sind Spaltanlagen, Versuchsserien und

14 Messwiederholungen auf der selben Parzelle. Die in solchen Versuchen erhobenen Daten

15 können sehr einfach mit gemischten Modellen ausgewertet werden. Obwohl das

16 Instrumentarium für solche Modelle mittlerweile sehr weit entwickelt ist und gute Software

17 zur Verfügung steht, werden diese Verfahren in der Praxis noch zögerlich eingesetzt. Der

18 Zweck der vorliegenden Arbeit ist es daher, zur häufigeren Anwendung gemischter Modelle

19 zu ermuntern. Wir beschreiben einige Grundregeln, welche bei der Formulierung gemischter

20 Modelle helfen, denn dies ist die Haupthürde für den Nutzer statistischer Software. Die

21 wesentlichen Aspekte werden in mehreren Beispielen erläutert. Auf theoretische Grundlagen

22 wird nur in dem Maße eingegangen, wie es für eine kompetente Nutzung von Programmen

23 für gemischte Modelle erforderlich erscheint. Abschließend wird auf einige neuere

24 methodische Entwicklungen hingewiesen, die für die Pflanzenwissenschaften von

25 besonderem Interesse sein dürften.

1 **Schlüsselwörter:** Varianzanalyse - Blockbildung - Fehlerstrata - Geostatistik - longitudinale  
2 Daten - gemischte Modelle - zufälliger Effekt - Randomisation - Messwiederholungen -  
3 Versuchsserien

4

## 5 **1. Einleitung**

6

7 In der agrarwissenschaftlichen Forschung werden geplante Versuche in der Regel unter der  
8 Annahme eines linearen Modells ausgewertet. Häufig ist ein gemischtes Modell erforderlich,  
9 weil mehrere zufällige Streuungsursachen zu berücksichtigen sind. Einige Beispiele sind wie  
10 folgt: (1) Eine Spaltanlage erfordert zwei Fehlerterme für Grossteilstück- und  
11 Kleinteilstückfehler. (2) Messwiederholungen zu verschiedenen Zeitpunkten und/oder an  
12 verschiedenen Punkten innerhalb einer Randomisationseinheit sind korreliert, was durch ein  
13 gemischtes Modell mit entsprechender Varianz-Kovarianz-Struktur berücksichtigt werden  
14 kann. (3) Die räumliche Analyse geostatistischer Daten kann im Rahmen eines gemischten  
15 Modells implementiert werden. (4) Versuche, die an mehreren Orten und/oder in mehreren  
16 Jahren wiederholt werden, erfordern ein lineares Modell mit zufälligen Umwelteffekten, falls  
17 die Umwelten als zufälliger Faktor betrachtet werden können und falls das Ziel der Analyse in  
18 der Berechnung von Mittelwerten über die Umwelten besteht. (5) Die Nutzung der Inter-  
19 Block Information in einer unvollständigen Blockanlage kann durch ein gemischtes Modell  
20 mit zufälligen Blockeffekten realisiert werden.

21

22 Vor der Verbreitung von Computern war die Verrechnung nach einem gemischten Modell  
23 eine große Herausforderung. Eine Analyse war nur praktikabel in einfachen Fällen und bei  
24 balancierten Daten. Eine umfassende Analyse komplexerer Daten, z.B. einer unbalancierten  
25 Versuchsserie mit Varianzheterogenität auf verschiedenen Ebenen (Behandlung-Umwelt  
26 Interaktionen, Fehler) und räumlichen Korrelationen innerhalb eines jeden Versuchs (Smith et

1 al. 2001) war praktisch unmöglich. Während in den vergangenen 20 Jahren die theoretischen  
2 Grundlagen weit entwickelt wurden (McLean et al. 1991, Searle et al. 1992, Verbeke and  
3 Molenberghs 2000) und einfach zu nutzende Software zur Verfügung steht, hat sich die  
4 Nutzung gemischter Modelle in der agrarwissenschaftlichen Forschung noch nicht in dem  
5 Maße verbreitet wie es angemessen wäre. Diese Diskrepanz ist teilweise zurückzuführen auf  
6 die Tatsache, dass die Darstellung in Lehrbüchern oft sehr theorielastig ist, wodurch die  
7 praktische Anwendung schwieriger erscheint als sie tatsächlich ist. Hinzu kommt, das  
8 Ausgabe von entsprechenden Computerprogrammen für den in der Varianzanalyse (analysis  
9 of variance - ANOVA) und der Berechnung von Grenzdifferenzen (LSD) bewanderten Nutzer  
10 ein zunächst etwas ungewohntes Erscheinungsbild aufweist, obwohl es weit mehr  
11 Übereinstimmungen als Unterschiede zwischen beiden Verfahren gibt.

12

13 Wir sind überzeugt, dass Agrarwissenschaftler valide und sinnvolle Auswertungen mit einem  
14 gemischten Modell erhalten können, wenn sie über die entsprechende Software und die  
15 Kenntnis einiger wichtiger Grundprinzipien verfügen. Das Ziel des vorliegenden Artikels ist  
16 die Beschreibung dieser Grundprinzipien. Hierdurch hoffen wir, zur Verbreitung dieses  
17 flexiblen und leistungsstarken Instrumentariums beizutragen. Das Hauptaugenmerk wird  
18 dabei auf der Formulierung eines gemischten Modells gelegt, die Hauptaufgabe in einer  
19 computer-basierten Analyse. Wir geben außerdem einen groben Überblick über einige neuere  
20 Entwicklungen, die für die Pflanzenwissenschaften relevant sind. Leser, die sich bereits mit  
21 gemischten Modellen gut auskennen, werden hier kaum etwas neues finden. Unsere  
22 Zielgruppe sind Wissenschaftler mit soliden Kenntnissen der klassischen Varianzanalyse, die  
23 sich selbstständig in das Gebiet der gemischten Modelle einarbeiten wollen.

24

25 Die Analyse nach einem gemischten Modell umfasst zwei wesentliche Schritte: (1)  
26 Formulierung eines Modells. (2) Ableiten statistischer Schlüsse (Parameterschätzung, Tests,

1 Vertrauensintervalle). Unsere Arbeit widmet jeweils einen Abschnitt für jeden dieser beiden  
2 Schritte. In Abschnitt 3 geben wir eine Reihe von Regeln an, die wir als hilfreich für die  
3 Entwicklung von gemischten Modellen erachten. Die Regeln konzentrieren sich auf  
4 gemischte Modelle vom ANOVA-Typ für qualitative Faktoren (Hocking 1985). Die  
5 Prinzipien statistischer Schlussweisen werden in Abschnitt 4 diskutiert. Unsere  
6 Modellnotation wird in Abschnitt 2 beschrieben. Erweiterungen gemischter Modelle vom  
7 ANOVA-Typ werden kurz am Ende der Arbeit besprochen (Abschnitt 5). Die Darstellung  
8 richtet sich vor allem an Nutzer von Statistikpaketen. Es gibt eine Reihe von guten Paketen  
9 für gemischte Modelle (ASREML, BMDP, GENSTAT, SAS, S-PLUS, SPSS). Während die  
10 hier beschriebenen Grundlagen für jedes dieser Pakete zutreffen, beziehen wir uns zur  
11 Illustration gelegentlich auf das SAS System (Littell et al. 1996), ein Paket, mit dem wir uns  
12 auskennen und welches unter Agronomen sehr verbreitet ist. Es sei aber betont, dass die  
13 Bezugnahme auf ein bestimmtes Paket keineswegs eine spezifische Empfehlung beinhaltet.

14  
15

## 16 **2. Modellsyntax**

### 17 **2.1 Eine praktische Modellnotation**

18

19 **Beispiel 1:** Eine Experiment wird in einer Spaltanlage durchgeführt, wobei der  
20 Grossteilstückfaktor A (Dünger) in vollständigen Blöcken (R) randomisiert ist, während der  
21 Kleinteilstückfaktor B (Genotyp) innerhalb der Grossteilstücke vollständig randomisiert wird.  
22 Ein lineares Modell für die Daten kann geschrieben werden als

23

$$24 \quad y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + r_k + b_{ik} + e_{ijk} \quad , \text{ wobei} \quad (1)$$

25

26  $y_{ijk}$  = Ertrag der  $i$ -ten Düngerstufe und des  $j$ -ten Genotyp im  $k$ -ten Block

- 1  $\mu$  = Gesamtmittel
- 2  $\alpha_i$  = Haupteffekt der  $i$ -ten Düngersstufe
- 3  $\beta_j$  = Haupteffekt des  $j$ -ten Genotyp
- 4  $(\alpha\beta)_{ij}$  = Dünger-Genotyp Wechselwirkung
- 5  $r_k$  = Effekt des  $k$ -ten Block
- 6  $b_{ik}$  = Fehler des  $i$ -ten Grossteilstück im  $k$ -ten Block
- 7  $e_{ijk}$  = Kleinteilstückfehler (Rest).

8

9 Alle Effekte sind fest, mit Ausnahme der zufälligen Fehlereffekte  $b_{ik}$  und  $e_{ijk}$ , welche als

10 normalverteilt mit Mittelwert Null und Varianzen  $\sigma_b^2$  und  $\sigma_e^2$  betrachtet werden. Wir

11 betrachten (1) als gemischtes Modell vom ANOVA-Typ, da es die einfachen Annahmen der

12 Unabhängigkeit und der Varianzhomogenität für alle zufälligen Effekte umfasst. Es gibt

13 zahlreiche Erweiterungen gemischter Modelle vom ANOVA-Typ, welche diese Annahmen

14 abschwächen. Eine kurze Übersicht wird am Ende dieser Arbeit gegeben.

15

16 Es sei angenommen, dass in einem zu analysierenden Datensatz die Düngersstufen mit A

17 kodiert sind, während die Genotypen mit B kodiert sind, und Blöcke mit R. Ein alternativer

18 Ausdruck für das Modell in Gl. (1), welche sich stärker an die in Statistikpaketen verwendete

19 Syntax anlehnt, ist

20

21 
$$Y = A + B + A \bullet B + R + R \bullet A + R \bullet A \bullet B, \tag{2}$$

22

23 mit den folgenden Entsprechungen: Y entspricht  $y_{ijk}$ , A entspricht  $\alpha_i$ , B entspricht  $\beta_j$ , A•B

24 entspricht  $(\alpha\beta)_{ij}$ , R•A entspricht  $b_{ik}$  und R•A•B entspricht  $e_{ijk}$ . Modell (2) umfasst kein

25 Symbol für das allgemeine Mittel bzw. den Achsenabschnitt,  $\mu$ . Diese Schreibweise basiert

1 auf der Annahme dass (fast) jedes lineare Modell einen solchen Effekt umfasst. Außerdem  
2 wird üblicherweise der Resteffekt ( $R \bullet A \bullet B$  im vorliegenden Fall) nicht explizit angegeben. In  
3 dieser Arbeit geben wir diesen Effekt explizit an und unterstreichen ihn der Klarheit halber.  
4 Um die Notation weiter zu vereinfachen, kann man zusätzlich die Zielvariable  $Y$  weglassen  
5 (Wilkinson and Rogers 1973), was zu folgendem Ausdruck führt:

$$6 \quad A + B + A \bullet B + R + R \bullet A + \underline{R \bullet A \bullet B} . \quad (3)$$

8  
9 Die Effekte  $A$ ,  $B$ ,  $A \bullet B$ , und  $R$  sind fest, während  $R \bullet A$  und  $\underline{R \bullet A \bullet B}$  zufällig sind. Dies kann  
10 kenntlich gemacht werden, indem feste und zufällige Effekte durch einen Doppelpunkt  
11 getrennt werden, wobei die festen Effekte vorne stehen (Patterson 1997):

$$12 \quad A + B + A \bullet B + R : R \bullet A + \underline{R \bullet A \bullet B} . \quad (4)$$

14  
15 Wir benutzen aus zwei Gründen durchweg die in (4) angegebene Notation: (i) Sie ist  
16 einfacher als die in (1) verwendete Notation, beinhaltet aber fast dieselbe Information; (ii) sie  
17 entspricht im wesentlichen der in statistischer Software verwendeten Syntax. Die wichtigsten  
18 Anweisungen für Beispiel 1 bei Verwendung der SAS Prozedur MIXED sind

19  
20 `model Y = A B A*B R;`  
21 `random R*A;`

22  
23 Alle festen Effekte werden in der *model* Anweisung aufgeführt, während alle zufälligen  
24 Effekte mit Ausnahme des Restes in der *random* Anweisung festgelegt werden. Der Rest  
25 muss nicht explizit angegeben werden, weil dieser Effekt automatisch angepasst wird.

26

## 1 **2.2 Vier Modelloperatoren**

2

3 Zur Herleitung eines für ein gegebenes Design adäquaten Modells verwenden wir eine  
4 Kurzschreibweise, die von Nelder (1965), Wilkinson und Rogers (1973) sowie McCullagh  
5 und Nelder (1989) vorgeschlagen wurde. Diese Schreibweise umfasst vier Operatoren, die im  
6 folgenden beschrieben werden. Die Operatoren erlauben es, ein Modell in kompakter Form  
7 darzustellen. Die Notation ist intuitiv leicht zugänglich und zeigt schnell die wesentlichen  
8 Aspekte eines Modells.

9

10 **Punktoperator (•):** Der Punktoperator wird verwendet, um gekreuzte Effekte zu definieren.

11 Die Bildung eines gekreuzten Effektes kann als eine mathematische Operation begriffen  
12 werden, ebenso wie die Addition von Effekten in einem Modell mittels des "+"-Operators.

13 Der Punktoperator hat eine höhere Priorität als der "+"-Operator, d.h.,  $A \bullet (B+C) = A \bullet C +$

14  $A \bullet B$ . Der Punktoperator gehorcht außerdem dem Kommutativ- und dem Assoziativgesetz, so

15 dass  $A \bullet B = B \bullet A$  und  $(A \bullet B) \bullet C = A \bullet (B \bullet C)$  gelten. Wenn zwei Effekte gekreuzt werden,

16 welche denselben Faktor enthalten, so werden alle bis auf einen Repräsentanten gelöscht. So

17 gilt zum Beispiel  $(R \bullet A) \bullet (R \bullet B) = R \bullet A \bullet B$ .

18

19 **Produkttermoperator [pt(.)]:** Wenn M ein Modell bezeichnet, so ist pt(M) der mittels des

20 Punktoperators gebildete Produktterm aller in M enthaltenen Effekte. So ist z.B.  $pt(A+B) =$

21  $A \bullet B$  und  $pt(A + B \bullet C) = A \bullet B \bullet C$ .

22

23 **Schachtelungsoperator (/):** Wenn ein Faktor B in einem anderen Faktor A geschachtelt ist,

24 so muss das Modell die Terme A und  $A \bullet B$  enthalten. Eine Kurzschreibweise hierfür ist  $A/B$ ,

25 wobei "/" der Schachtelungsoperator ist. So können wir z.B. schreiben:  $A/B = A + A \bullet B$ .

1 Wenn ein weiterer Faktor C in B geschachtelt ist, so kann man schreiben:  $A/B/C$ . Der  
2 Schachtelungsoperator ist assoziativ, d.h.  $A/(B/C) = (A/B)/C$ . Desweiteren gilt  $A/(B+C) =$   
3  $A/B + A/C$ . Wenn A und B selbst Modelle sind, so ist  $A/B$  definiert als  $A/B = A + pt(A) \bullet B$ .

4

5 **Kreuzungsoperator ( $\times$ ):** Das Modell für zwei gekreuzte Faktoren A und B kann geschrieben  
6 werden als  $A \times B = A + B + A \bullet B$ . Wir benutzen das Symbol " $\times$ " anstelle des Symbols "\*",  
7 welches von Nelder (1965) vorgeschlagen, da wir eine Verwechslung mit der SAS Syntax  
8 vermeiden wollen, in der "\*" den Punktoperator bezeichnet. Zwei wichtige algebraische  
9 Regeln für drei Faktoren A, B und C sind:

10

$$11 \quad A \times (B+C) = A + (B+C) + A \bullet (B+C) = A + B + C + A \bullet B + A \bullet C \quad \text{und} \quad (5)$$

$$12 \quad (A \times B)/C = A \times B + pt(A \times B) \bullet C = A \times B + A \bullet B \bullet C \quad . \quad (6)$$

13

14

### 15 **3. Regeln für die Formulierung eines gemischten Modells**

16

17 In diesem Abschnitt werden eine Reihe von Regeln angegeben und erläutert, die bei der  
18 Formulierung eines gemischten Modelle helfen können. Zur besseren Übersichtlichkeit  
19 werden die Regeln zusammenfassend in Tabelle 1 aufgeführt.

20

#### 21 **3.1 Zufällig versus. fest**

22

23 **Regel 1 (wann ist ein Faktor zufällig?):** Ein Faktor ist zufällig wenn die beobachteten  
24 Stufen des Faktors als Zufallsstichprobe einer Grundgesamtheit aufgefasst werden können  
25 (z.B. Umwelten, Stichprobeneinheiten). Weiterhin ist ein Faktor zufällig, wenn er eine

1 Randomisationseinheit repräsentiert (z.B. Parzellen). Andernfalls wird der Faktor  
2 üblicherweise als fest betrachtet (z.B. nicht randomisierte Blocks, Behandlungen). Wenn die  
3 Stufen eines Faktors verglichen werden sollen (z.B. Behandlungen), so wird der Faktor als  
4 fest betrachtet, und zwar unabhängig davon, ob er dem Design nach zufällig ist. Wenn ein  
5 Faktor zufällig ist, so sind alle Effekte, die diesen Faktor enthalten, ebenfalls zufällig.

6

7 **Beispiel 2:** Es wird eine Serie von Versuchen mit ausgewählten Sorten durchgeführt. Die  
8 Versuche werden an einer Zufallsstichprobe von Orten aus einer Zielregion durchgeführt. Ziel  
9 der Auswertung ist die Berechnung von Sortenmittelwerten über die für die Zielregion  
10 repräsentativen Orte. Das Design umfasst zwei Faktoren: Sorte und Ort. Sorten sind ein fester  
11 Faktor, da Vergleiche zwischen den Stufen dieses Faktors von Interesse sind und weil die  
12 Sorten ausgewählt wurden, so dass sie keine Zufallsstichprobe darstellen. Die Orte sind  
13 dagegen ein zufälliger Faktor, weil diese zufällig ausgewählt wurden und weil die Stufen  
14 dieses Faktors nicht von besonderem Interesse sind; die Prüforte fungieren vielmehr als  
15 Wiederholungen. Die Wechselwirkung zwischen Ort und Sorte ist zufällig, weil sie den  
16 Faktor Ort enthält.

17

18 Es sollte betont werden, dass Orte oft nicht zufällig gewählt werden, beispielsweise wenn  
19 bestehende Versuchsstationen verwendet werden, die so ausgewählt wurden, dass sie  
20 bestimmte agrarökologische Zonen gut repräsentieren. In diesen Fällen ist es angemessen,  
21 Orte als festen Faktor zu betrachten.

22

### 23 **3.2 Zwei Faktorarten**

24

25 **Regel 2 (Zwei Faktorarten):** Man kann Behandlungs- und Blockfaktoren unterscheiden.  
26 *Blockfaktoren* beziehen sich auf zufällig ausgewählte Stichprobeneinheiten (Pflanzen,

1 Bodenproben, etc.), Randomisationseinheiten (Zeilen, Spalten, unvollständige Blöcke,  
2 Grossteilstücke, Kleinteilstücke, etc.) und Blockeinheiten, welche nicht selbst Teil des  
3 Randomisationsprozesses sind (vollständige Blöcke, Umwelten, etc.). Blockfaktoren sind  
4 erforderlich, um jede *Beobachtungseinheit* (Parzelle, Pflanze, etc.) eindeutig zu identifizieren.  
5 Mit anderen Worten: Blockfaktoren sind den Beobachtungseinheiten „angeboren“ (innate;  
6 Brien 1983). Ein *Behandlungsfaktor* und seine Stufen werden vom Versuchsansteller  
7 ausgewählt, um eine Versuchsfrage zu beantworten. Die Stufen des Behandlungsfaktors  
8 werden den Beobachtungseinheiten durch einen Randomisationsprozess zufällig zugeordnet;  
9 somit sind Behandlungen den Beobachtungseinheiten nicht *angeboren*.

10

11 **Beispiel 3:** Eine Serie von randomisierten vollständigen Blockexperimenten mit  
12 verschiedenen Sorten wird an verschiedenen Standorten durchgeführt. Sorten sind ein  
13 randomisierter Behandlungsfaktor, während Orte, vollständige Blöcke und Parzellen  
14 Blockfaktoren sind. Insbesondere können Orte als Superblöcke betrachtet werden, die jeweils  
15 aus vollständigen Blöcken aufgebaut sind. Die Blockfaktoren identifizieren die  
16 Beobachtungseinheiten (Parzellen) eindeutig und sind diesen *angeboren*. Im Gegensatz dazu  
17 werden die Sorten den Beobachtungseinheiten durch einen Randomisationsprozess  
18 zugeordnet und sind diesen somit nicht *angeboren*.

19

### 20 ***3.3 Behandlungs- und Blockmodell für ein randomisiertes Experiment***

21

22 Die hier angegebenen Regeln sind im wesentlichen angelehnt an Nelder (1965) sowie  
23 Wilkinson und Rogers (1973). Sie sind außerdem ähnlich den Anweisungen zur Formulierung  
24 eines Blockmodells wie sie in GENSTAT (Payne and Wilkinson 1977) und GLIM  
25 (McCullagh und Nelder 1989) implementiert sind. Die Regeln basieren auf R. A. Fisher's  
26 Annahme, dass keine Interaktion zwischen Behandlung und Block bestehen (Nelder 1965,

1 Brien 1983, Bailey, 1991).

2

3 **Regel 3 (Behandlungs- und Blockmodell getrennt halten, jedenfalls zu Beginn):** Bei der  
4 Modellierung geplanter Versuche ist es hilfreich, das Modell für die Behandlungen und das  
5 Modell für die Blöcke getrennt zu halten (Nelder 1965). Die Behandlungsstruktur kann allein  
6 mit Hilfe der Behandlungsfaktoren modelliert werden, während das Blockmodell  
7 ausschließlich mit Hilfe der Blockfaktoren formuliert werden kann. Das Blockmodell  
8 beschreibt die Daten wenn es keine Behandlungseffekte gibt, und es repräsentiert die den  
9 Beobachtungseinheiten angeborene Struktur.

10

11 **Regel 4 (Effekte des Blockmodells):** Bei einem geplanten Versuch erhält jede  
12 Randomisationseinheit (jedes Fehlerstratum) einen eigenen zufälligen Effekt. Eine Versuchs-  
13 oder Blockeinheit wird zu einer Randomisationseinheit, wenn die Stufen eines Faktors  
14 (Behandlungen oder andere) den verschiedenen Einheiten zufällig zugeordnet werden. Wenn  
15 zwei Randomisationseinheiten (Fehlerstrata) *gekreuzt* werden, d.h. eine Einheit der einen  
16 Randomisation reicht über mehrere Einheiten der anderen Randomisation, so erzeugt die  
17 Kreuzung beider Randomisationsschritte eine weitere Versuchseinheit (z.B. eine Parzelle).  
18 Das Modell enthält dann einen zufälligen Effekt, der durch die Kreuzung derjenigen beiden  
19 Variablen erhalten wird, welche die gekreuzten Randomisationseinheiten repräsentieren.  
20 Wenn mehr als zwei Randomisationsschritte gegenseitig gekreuzt werden, so enthält das  
21 Modell zufällige Effekte für alle möglichen Kreuzungen der Variablen, welche die  
22 Randomisationseinheiten repräsentieren (2-Weg-, 3-Weg-, etc.). Desweiteren erhält jede  
23 Stichprobeneinheit einen eigenen zufälligen Effekt. Schließlich enthält das Blockmodell feste  
24 Effekte für Blockfaktoren, welche nicht selbst Teil der Randomisation oder der  
25 Stichprobennahme sind, z.B. Blöcke in einer randomisierten vollständigen Blockanlage. Der  
26 zufällige Effekt, welcher der Beobachtungseinheit zugeordnet ist, entspricht dem Restfehler.

1

2 **Regel 5 (Kodierung der Blockfaktoren):** Jeder Blockfaktor kann eindeutig durch eine  
3 eigene Variable kodiert werden. Diese Kodierung erlaubt es auf einfache Weise, ein korrektes  
4 Blockmodell zu formulieren. Nach der Formulierung des Blockmodells können Variablen,  
5 welche Blockfaktoren kodieren, oft durch eine Variable ersetzt werden, welche einen  
6 Behandlungsfaktor oder gekreuzte Effekte bezüglich mehrere Behandlungsfaktoren  
7 repräsentieren. In diesem Fall ist es möglich, dass die den Blockfaktor kodierende Variable  
8 ganz aus dem Modell gelassen wird. Es sollte jedoch betont werden, dass diese Art der  
9 Ersetzung nicht notwendig ist für eine korrekte Analyse. Es kann im Einzelfall sogar besser  
10 sein, ganz auf eine solche Ersetzung zu verzichten, vor allem in komplexen Versuchen, wo  
11 hierdurch leicht die wesentlichen Eigenschaften des Designs verschleiert werden können  
12 (Brien 1983).

13

14 **Beispiel 4:** Ein Behandlungsfaktor A wird in randomisierten vollständigen Blöcken (R)  
15 geprüft. Das Blockmodell umfasst einen festen Effekt für die vollständigen Blöcke und einen  
16 zufälligen Effekt für die Parzellen (Randomisationseinheiten), der innerhalb der vollständigen  
17 Blöcke *geschachtelt* sind. Es sei angenommen, dass die Parzellen eindeutig durch die  
18 Variable PARZELLE kodiert sind, dass also jede Parzelle eines Blocks einen anderen Wert  
19 für diese Variable hat. Das Blockmodell lautet dann  $R : \underline{R \bullet PARZELLE}$ , während das  
20 Behandlungsmodell durch A gegeben ist. Eine Kurzschreibweise für die *geschachtelte*  
21 Randomisation von Parzellen innerhalb der Blöcke lautet R/PARZELLE. Das vollständige  
22 Modell heisst  $A + R : \underline{R \bullet PARZELLE}$ . Nun wird eine Parzelle innerhalb eines Blocks  
23 eindeutig durch die Stufe des Faktors A identifiziert, welche auf der betreffenden Parzelle  
24 geprüft wird. Daher kann die Variable PARZELLE durch A ersetzt werden, was zu der  
25 gleichwertigen Schreibweise  $A + R : \underline{R \bullet A}$  führt.

26

1 **Beispiel 5:** Ein Behandlungsfaktor A wird in einem Lateinischen Quadrat geprüft. Hierbei  
 2 gibt es zwei *gekreuzte* Randomisationseinheiten: Zeilen (ZEILE) und Spalten (SPALTE). Das  
 3 Blockmodell lautet daher  $ZEILE \times SPALTE = ZEILE + SPALTE + \underline{ZEILE \bullet SPALTE}$ . Der  
 4 gekreuzte Effekt  $ZEILE \bullet SPALTE$  wird durch die Kreuzung der Faktoren Zeile und Spalte  
 5 erzeugt und repräsentiert die Beobachtungseinheiten (Parzellen). Nach Hinzunahme des  
 6 Behandlungsmodells lautet das volle Modell:  $A : ZEILE + SPALTE + \underline{ZEILE \bullet SPALTE}$ .

7

8 **Beispiel 6:** Eine Spaltanlage mit zwei Faktoren A und B wird so angelegt, dass die  
 9 Großteilstücke in vollständigen Blocks (R) randomisiert sind. Faktor A ist der  
 10 Großteilstückfaktor, d.h., die Stufen von A werden zufällig den Großteilstücken innerhalb  
 11 eines Blocks zugeordnet. Die Stufen von B werden zufällig den Kleinteilstücken innerhalb  
 12 der Großteilstücke zugeordnet. Es gibt daher zwei *geschachtelte* Randomisationseinheiten:  
 13 Großteilstücke (GROSS) innerhalb der vollständigen Blöcke und Kleinteilstücke (KLEIN)  
 14 innerhalb der Großteilstücke. Eine Kurzschreibweise für das Blockmodell lautet  
 15  $R/GROSS/KLEIN$ , was zu

16

17  $R/(GROSS/KLEIN)$

18 
$$= R/(GROSS + GROSS \bullet KLEIN) = R + R \bullet GROSS + \underline{R \bullet GROSS \bullet KLEIN} \quad (7)$$

19

20 entwickelt wird. Ein Großteilstück innerhalb eines vollständigen Blocks kann eindeutig  
 21 identifiziert werden, indem die zugehörige Stufe des Faktors A benannt wird. Ebenso wird ein  
 22 Kleinteilstück innerhalb eines Großteilstückes durch die zugehörige Stufe des Faktors B  
 23 eindeutig angesprochen. Daher kann das Blockmodell auch als  $R/A/B$  ausgedrückt werden,  
 24 was nach Auflösung  $R : R \bullet A + \underline{R \bullet A \bullet B}$  ergibt. Das Behandlungsmodell ist  $A \times B = A + B +$   
 25  $A \bullet B$ , wobei  $A \bullet B$  die Interaktion von A und B ist, so dass das volle Modell mit dem in Gl. (4)

1 entspricht.

2

3 **Beispiel 7:** Wenn viele Behandlungen zu prüfen sind, ist es oft sinnvoll, die vollständigen  
4 Wiederholungen in unvollständige Blöcke zu unterteilen. Jede Behandlung wird einmal je  
5 Wiederholung geprüft. Diese Art von Versuchsanlage wird als Gitteranlage bezeichnet, wobei  
6 verschiedene Varianten möglich sind (Mead 1997). Die Anlage eines Gitters für einen  
7 einzigen Behandlungsfaktor beinhaltet zwei *geschachtelte* Randomisationseinheiten:  
8 Parzellen (PARZELLE) innerhalb unvollständiger Blöcke und unvollständige Blöcke  
9 (UBLOCK) innerhalb vollständiger Wiederholungen (R). Das Blockmodell lautet daher  
10 R/UBLOCK/PARZELLE, was zu

11

$$12 \quad R : R \bullet \text{IBLOCK} + \underline{R \bullet \text{IBLOCK} \bullet \text{PARZELLE}} \quad (8)$$

13

14 entwickelt wird.

15

16 **Beispiel 8:** Zwei Faktoren A und B werden in einer Streifenanlage mit Faktor A in den  
17 Großzeilen (GROSSZEILE) und B in den Großspalten (GROSSSPALTE) geprüft. Die  
18 Randomisation der Großzeilen und Großspalten ist *gekreuzt*. Jede Parzelle dieser Anlage wird  
19 als Großteilstück betrachtet, welches als Block zur Prüfung zweier weiterer Faktoren C und D  
20 dient. Diese werden in einer Streifenanlage mit C in den Kleinzeilen (KLEINZEILE) und D  
21 in den Kleinspalten (KLEINSPALTE) geprüft. Das Behandlungsmodell lautet  $A \times B \times C \times D$ ,  
22 während das Blockmodell gegeben ist durch  $R / (\text{GROSSZEILE} \times \text{GROSSSPALTE}) /$   
23  $(\text{KLEINZEILE} \times \text{KLEINSPALTE})$ . Innerhalb eines Blocks können Großzeilen und Spalten  
24 eindeutig durch die Stufen der Faktoren A und B identifiziert werden. Ebenso werden  
25 Kleinzeilen und –spalten eines Großteilstückes eindeutig durch die jeweiligen Stufen der  
26 Faktoren C und D angesprochen. Daher kann das Blockmodell auch als  $R / (A \times B) / (C \times D)$

1 geschrieben werden. Dies berücksichtigt die Tatsache, dass die *gekreuzte* Randomisation für  
2 C und D *geschachtelt* ist in derjenigen für A und B. Das Blockmodell kann wie folgt  
3 ausgeschrieben werden:

4

$$5 \quad R : R \bullet A + R \bullet B + R \bullet A \bullet B + R \bullet A \bullet B \bullet C + R \bullet A \bullet B \bullet D + \underline{R \bullet A \bullet B \bullet C \bullet D} . \quad (9)$$

6

7 Die ersten drei zufälligen Effekte repräsentieren die Randomisationsstruktur einer  
8 Streifenanlage für A und B. Zufällige Effekte, welche die Faktoren C und D umfassen, sind  
9 alle geschachtelt in den Großteilstücken, die durch  $R \bullet A \bullet B$  repräsentiert sind.

10

### 11 ***3.4 Interaktionen zwischen Block- und Behandlungsfaktoren***

12

13 **Regel 6 (Interaktion zwischen Block und Behandlungsfaktoren):** Nach der Formulierung  
14 von Block- und Behandlungsmodell, welche zunächst die Abwesenheit von Block-  
15 Behandlung-Interaktionen voraussetzt, kann man prüfen, ob Grund zu der Annahme von  
16 Interaktionen zwischen einzelnen Block- und Behandlungsfaktoren besteht. Die  
17 Notwendigkeit zur Annahme einer Interaktion besteht z.B. dann, wenn die Heterogenität  
18 zwischen den verschiedenen Blockeinheiten groß ist. Ein Interaktionseffekt kann dem Modell  
19 hinzugefügt werden, sofern die Versuchsanlage eine Schätzung des betreffenden Effektes  
20 erlaubt. Eine notwendige Voraussetzung für die Schätzbarkeit einer Block-Behandlung-  
21 Interaktion ist, dass eine valide Auswertung auch getrennt für jede Stufe des Blockfaktors  
22 möglich ist. Eine Analyse von Block-Behandlung-Interaktionen kann sogar von vorrangigem  
23 Interesse sein, beispielsweise dann, wenn Blöcke verschiedenen Umwelten (Standorten  
24 und/oder Jahren) entsprechen.

25

26 **Beispiel 9:** Eine Serie von randomisierten vollständigen Blockanlagen zur Prüfung

1 verschiedener Genotypen (G) wird an einer zufällig ausgewählten Stichprobe von Standorten  
2 durchgeführt (L). Blöcke seien mit dem Buchstaben R kodiert, während Parzellen mit der  
3 Bezeichnung PARZELLE kodiert werden. Standorte sind ein zusätzlicher Blockfaktor,  
4 welcher den vollständigen Blöcken übergeordnet werden. Das Blockmodell ist gegeben durch  
5  $L/R/PARZELLE$ , während das Behandlungsmodell durch den Buchstaben G repräsentiert ist.  
6 Unter der Annahme, dass keine Block-Behandlung-Interaktion besteht, lautet das volle  
7 Modell:  $G : L + L \bullet R + \underline{L \bullet R \bullet PARZELLE}$ .

8  
9 Unterschiede zwischen den Standorten sind oft so groß, dass von einer Genotyp-Umwelt-  
10 Interaktion ( $L \bullet G$ ) auszugehen ist. Eine Schätzung der Interaktion  $L \bullet G$  ist möglich, weil eine  
11 valide Auswertung je Ort ebenfalls möglich ist. Unter Hinzunahme der Interaktion ist das  
12 Modell gegeben durch

$$14 \quad G : L + L \bullet G + L \bullet R + \underline{L \bullet R \bullet PARZELLE} \quad . \quad (10)$$

### 16 *3.5 Mehrphasige Experimente*

17  
18 Einige Experimente können in verschiedene Phasen unterteilt werden, wobei in jeder Phase  
19 eine andere Versuchsanlage verwendet wird. In einer zweiphasigen Versuchsanlage werden  
20 beispielsweise Beobachtungseinheiten von der ersten in die zweite Phase übernommen, wo  
21 sie entsprechend einem neuen Versuchsplan randomisiert werden (Brien 1983).

22  
23 **Regel 7 (mehrphasige Experimente):** Wenn ein Experiment mehrere Phasen umfasst, so  
24 stelle man das Blockmodell für jede Phase auf und addiere die Blockmodelle der  
25 verschiedenen Phasen zum Gesamtmodell. Jede Wiederholung eines Effektes in den  
26 Blockmodellen verschiedener Phasen wird gelöscht. Falls Effekte verschiedener Phasen

1 vermengt sind, so wird nur ein Repräsentant im Modell behalten. Das Blockmodell einer  
2 Phase bezieht sich auf die Beobachtungseinheiten dieser Phase. In anderen Worten: Alle  
3 Blockfaktoren einer Phase identifizieren die Beobachtungseinheiten dieser Phase und sind  
4 diesen angeboren. Stichproben von den Beobachtungseinheiten einer Phase werden den  
5 Beobachtungseinheiten der folgenden Phase durch einen Randomisationsschritt zugeordnet.  
6 Alle Schritte im Verlauf eines Experiments, die zeitlich vor der Randomisation der folgenden  
7 Phase liegen, werden der unmittelbar vorhergehenden Phase zugerechnet.

8

9 Regel 7 entspricht im wesentlichen den von Brien (1983) sowie Brien und Payne (1999)  
10 vorgeschlagenen Regeln (siehe auch [www.maths.unisa.edu.au/~matcjb/multitier/](http://www.maths.unisa.edu.au/~matcjb/multitier/)). Unter  
11 Berücksichtigung von Regel 6 können nach Bedarf Interaktionen verschiedener Phasen  
12 hinzugefügt werden, sofern die Versuchsanlage eine Schätzung dieser Effekte erlaubt.

13

14 **Beispiel 10:** Das vorliegende Beispiel bezieht sich auf eine Variante der Spaltanlage, die von  
15 Cochran und Cox (1957, Abschnitt 7.33) sowie in leicht abgewandelter Form von  
16 Hinkelmann und Kempthorne (1994, Abschnitt 13.4.3) diskutiert wird. Angenommen in der  
17 ersten Phase wird der Behandlungsfaktor „Erntezeitpunkt“ (A) in randomisierten  
18 vollständigen Blöcken (R) geprüft. Auf jeder Parzelle werden die geernteten Pflanzen zu einer  
19 Mischprobe zusammengefasst und dann in drei Teilproben aufgeteilt, von denen jede mit  
20 jeweils einer von drei verschiedenen chemischen Analysemethoden auf ihren Gehalt an einem  
21 bestimmten Nährelement untersucht wird. Die drei Analysemethoden stellen einen weiteren  
22 Behandlungsfaktor B mit drei Stufen dar. Für eine Parzelle werden die drei Methoden zufällig  
23 den drei Teilproben zugeordnet. In Phase I liegt eine Spaltanlage mit Blockmodell  $R/A/B = R$   
24  $: R \bullet A + \underline{R \bullet A \bullet B}$  zugrunde. Die chemische Analyse im Labor entspricht Phase II. Die Proben  
25 werden vor der Analyse getrocknet. Es ist eine Besonderheit und eine Schwäche der  
26 Versuchsanlage in Phase II, dass der Behandlungsfaktor A (Erntezeitpunkt) zwangsläufig als

1 zeitlicher Blockfaktor verwendet werden, weil die geernteten Pflanzen umgehend  
 2 weiterverarbeitet werden müssen; denn Proben von verschiedenen Erntezeitpunkten können  
 3 nicht zeitgleich getrocknet werden. Die Haupteffekte für den Erntezeitpunkt sind daher mit  
 4 den Blockeffekten in Phase II vermengt, so dass streng genommen keine statistischen  
 5 Schlüsse bezüglich des Erntezeitpunktes möglich sind.  
 6  
 7 Es stehen drei Öfen (O) für die Trocknung in Phase II zur Verfügung. Sie müssen simultan  
 8 genutzt werden, um die anschließenden chemischen Analysen vergleichbar zu machen. Die  
 9 Kapazität der Öfen ist gerade groß genug zur Trocknung aller zu einem Zeitpunkt geernteten  
 10 Pflanzen. Sie unterscheiden sich in der Geschwindigkeit, mit der die Proben getrocknet  
 11 werden. Die unterschiedliche Trockengeschwindigkeit kann eine Auswirkung auf das  
 12 Ergebnis der chemischen Analyse haben. Aus diesem Grund wird der Ofen als Blockfaktor  
 13 verwendet. Insbesondere wird das Grossteilstück der ersten Phase als Spalte eines  
 14 Lateinischen Quadrats zur Prüfung des Faktors B verwendet, während der Ofentyp (O) als  
 15 Zeile des Lateinischen Quadrates fungiert (Abb. 1). Für jeden Erntezeitpunkt (A) wird ein  
 16 eigenes Lateinisches Quadrat randomisiert, d.h., die Lateinischen Quadrate A sind innerhalb  
 17 von A geschachtelt. Für eine gegebene Stufe des Faktors A ist das Grossteilstück eindeutig  
 18 identifiziert durch den zugehörigen Block aus Phase I (R). Daher spezifizieren O und R die  
 19 Zeilen und Spalten eines jeden Lateinischen Quadrates. Das Blockmodell in der zweiten  
 20 Phase ist somit gegeben durch  $A/(R \times O) = A : A \bullet R + A \bullet O + \underline{A \bullet R \bullet O}$ . Man beachte dass der  
 21 Effekt für das Großteilstück  $A \bullet R$  im Blockmodell beider Phasen erscheint, weil beide Phasen  
 22 das Großteilstück als Randomisationseinheit genutzt wird. Daher enthält das Blockmodell der  
 23 Phase II den Haupteffekt für den Faktor A, welcher als Blockvariable fungiert. Nur ein  
 24 Repräsentant der Effekte A und  $A \bullet R$  wird im endgültigen Modell beibehalten. Desweiteren  
 25 sind sowohl  $\underline{R \bullet A \bullet B}$  als auch  $\underline{A \bullet R \bullet O}$  Effekte mit einer separaten Stufe für jede  
 26 Beobachtungseinheit (Stichprobe), so dass beide dem Resteffekt entsprechen. Beide Effekte

1 sind daher vermengt und nur einer der beiden Effekte muss im Modell verbleiben. Man beachte  
2 insbesondere, dass es nicht möglich ist, für vermengte zufällige Effekte jeweils eine eigene  
3 Varianz zu schätzen. Es sollte allerdings betont werden, dass die beiden Residualeffekte nicht  
4 identisch sind. Der Term  $\underline{R \bullet A \bullet B}$  repräsentiert den Stichprobenfehler von Phase I, während  
5  $\underline{A \bullet R \bullet O}$  die zusätzliche Variation aufgrund des Trocknungsprozesses in Phase II umfasst.

6 Unter Einschluss des Behandlungsmodells  $A \times B$  lautet das volle Modell

7

$$8 \quad A + B + A \bullet B + R : A \bullet O + R \bullet A + \underline{R \bullet A \bullet B} . \quad (11)$$

9

10 Dies ist im wesentlichen das von Cochran und Cox (1957) sowie von Hinkelmann und  
11 Kempthorne (1994) formulierte Modell, mit der einzigen Ausnahme, dass wir den Effekt  $A \bullet O$   
12 als zufällig betrachtet, um die Randomisationsstruktur abzubilden.

13

### 14 **3.6 Zufällige Effekte als fest betrachten**

15

16 **Regel 8 (Zufällige Effekte als fest betrachten):** Ein zufälliger Effekt, der nicht den mittels  
17 F-Test oder t-Test zu prüfenden festen Effekt *enthält*, kann für die Analyse als fest betrachtet  
18 werden. Dies ist vor allem dann von Vorteil, wenn der zufällige Effekt nur fünf oder zehn  
19 Stufen aufweist, so dass eine verlässliche Schätzung der Varianz nicht möglich ist. Um zu  
20 entscheiden, ob ein zufälliger Effekt den zu testenden festen Effekt enthält, muss man den  
21 zufälligen Effekt nach Möglichkeit mittels der im festen Effekt enthaltenen Faktoren  
22 ausdrücken (siehe Regel 5). Zusätzlich kann folgende einfache Prüfung erfolgen: Ersetze den  
23 zufälligen Effekt durch seine Kreuzung mit dem zu testenden Behandlungseffekt. Wenn die  
24 resultierende Analyse unverändert bleibt, so ist der feste Behandlungseffekt im zufälligen  
25 Effekt enthalten, aber sonst nicht.

26

1 **Beispiel 7 (Fortsetzung):** Das Modell für eine Gitteranlage in Gl. (8) umfasst einen  
2 zufälligen Term für die unvollständigen Blöcke:  $R \bullet IBLOCK$ . Dieser Effekt enthält nicht den  
3 Behandlungseffekt A, wie man leicht bestätigen kann, indem man  $R \bullet IBLOCK$  durch  
4  $R \bullet IBLOCK \bullet A$  ersetzt, was das Resultat der Analyse verändert. Daher kann man formal den  
5 Effekt  $R \bullet IBLOCK$  als fest betrachten, wenn statistische Schlüsse bezüglich A gezogen  
6 werden sollen. Die resultierende Analyse nutzt nur die Intra-Block-Information und wird  
7 daher als Intra-Block-Analyse bezeichnet (Cochran and Cox 1957, Mead 1997, Federer and  
8 Wolfinger 1998, Federer et al. 2001). Wird dagegen der Blockeffekt als zufällig betrachtet, so  
9 nutzt die entsprechende Analyse die sog. Inter-Block-Information. Eine Nutzung der Inter-  
10 Block-Information ist nur dann lohnend, wenn die Blockvarianz nicht groß ist im Verhältnis  
11 zur Restvarianz und wenn die Zahl der Blöcke ausreichend ist, um die Blockvarianz genau zu  
12 schätzen. Oftmals liefern beide Analysen recht ähnliche Ergebnisse.

13  
14 **Beispiel 9 (Fortsetzung):** In Modell (10) können alle Effekte als fest betrachtet werden, die  
15 nicht den Faktor G enthalten. Dies führt zu dem Modell

$$17 \quad G + L + L \bullet R : L \bullet G + \underline{L \bullet R \bullet PARZELLE} \quad . \quad (12)$$

18  
19 Bei der resultierenden Analyse basieren Genotypvergleiche ausschließlich auf Vergleichen  
20 innerhalb eines Experiments, während die Inter-Experiment-Information ungenutzt bleibt. Die  
21 Auswertung entspricht der einer Intra-Block-Analyse für eine unvollständige Blockanlage  
22 (Patterson 1997). Eine Analyse basierend auf Modell (10) nutzt dagegen sowohl die Inter-  
23 Experiment- als auch die Inter-Block-Information, sofern die Daten unbalanciert sind.  
24 Oftmals sind bei Serien von Sortenversuchen die Varianzen für Effekte, welche den  
25 genotypischen Faktor G enthalten, relativ groß, so dass die Inter-Experiment-Information  
26 klein ist und beide Analysen ähnliche Ergebnisse liefern. Eine Analyse nach Modell (12) ist

1 vor allem dann nützlich, wenn die Zahl der Versuche gering ist.

2

3

#### 4 **4. Schließende Statistik**

5

6 Hier soll keine umfassende Beschreibung der theoretischen Grundlagen der schließenden  
7 Statistik für gemischte Modelle gegeben werden, diese kann an anderer Stelle nachgelesen  
8 werden (z.B., Hocking 1985; Searle et al. 1992, Verbeke und Molenberghs 2000). Unser Ziel  
9 ist vielmehr eine kurze Diskussion der für die Nutzer von Statistikpaketen für gemischte  
10 Modelle relevanten Fakten. Insbesondere werden wir die Unterschiede zwischen der  
11 klassischen Varianzanalyse und der Likelihood-basierten Analyse von gemischten Modellen.

12

#### 13 **4.1 Fehlende Werte**

14

15 Bei geplanten Versuchen treten relativ häufig fehlende Werte durch unvorhersehbare  
16 Umstände auf, so dass der für die Auswertung zur Verfügung stehende Datensatz kleiner ist  
17 als geplant. Glücklicherweise ist eine statistische Auswertung mittels gemischter Modelle  
18 ohne weiteres möglich, sofern die Daten die Voraussetzung des völlig zufälligen Fehlens  
19 (*missing completely at random* - MCAR) erfüllen (Verbeke und Molengerghs 1997).  
20 Vereinfacht gesagt fordert die MCAR Annahme, dass das Fehlmuster der Daten unabhängig  
21 von der Versuchsanlage ist, insbesondere der Behandlungsstruktur. Eine genauere Definition  
22 der MCAR-Annahme wird z.B. bei Verbeke und Molenberghs (1997) gegeben. Tatsächlich  
23 sind die statistischen Schlussfolgerungen auch unter etwas schwächeren Annahmen als  
24 MCAR gültig, wie bei Verbeke und Molenberghs (1997) diskutiert wird. Wie anhand der  
25 folgenden beiden Beispiele gezeigt wird, erfordert die Entscheidung, ob die MCAR Annahme  
26 gilt, eine genaue Betrachtung.

1

2 **Beispiel 11:** Drei Stichproben je Pflanze werden im Labor auf einen Pflanzennährstoff  
3 untersucht. Ein Reagenzglas wird versehentlich fallengelassen, so dass fehlende Werte  
4 vorliegen. Welches Reagenzglas verloren geht, wird in der Regel nicht von der Behandlung  
5 oder den zugehörigen Randomisationseinheiten beeinflusst, so dass die Daten die MCAR  
6 Annahme erfüllen.

7

8 **Beispiel 12:** Drei Düngemittel werden in einem Gefäßversuch untersucht. Jedes Gefäß enthält  
9 zu Beginn des Versuches zehn Pflanzen. Einer der Düngemittel hat einen negativen Einfluss  
10 auf einige Pflanzen, so dass diese im Laufe des Versuches absterben. Das resultierende  
11 Fehlmuster erfüllt nicht die MCAR Annahme, denn es besteht eine Abhängigkeit des  
12 Fehlmusters von der Behandlung.

13

#### 14 *4.2 Einfache Kleinste Quadrate vs. Generalisierte Kleinste Quadrate*

15

16 Feste Effekte eines linearen Modells können mit der Methode der Kleinsten Quadrate  
17 (ordinary least squares - OLSE) geschätzt werden. Diese Methode ist optimal für lineare  
18 Modelle mit einem Fehlerterm und homogenen Varianzen (Searle et al. 1992). In einem  
19 gemischten linearen Modell ist diese Methode dagegen nicht generell optimal. Eine bessere  
20 Methode ist bekannt unter der Bezeichnung Generalisierte Kleinste Quadrate (generalized  
21 least squares - GLSE). Diese Methode verwendet Gewichte, welche aus den  
22 Varianzkomponenten der zufälligen Effekte berechnet werden. GLSE hat optimale  
23 Eigenschaften, falls die Varianzkomponenten bekannt sind. Die Parameterschätzer sind dann  
24 sog. Beste Lineare Unverzerrte Schätzer (best linear unbiased estimators - BLUE), d.h. sie  
25 haben minimale Varianz unter allen unverzerrten linearen Schätzern. In der Praxis müssen die  
26 Varianzen zwar geschätzt werden, aber GLSE ist meist auch dann besser als OLSE, wenn

1 Schätzwerte der Varianzen verwendet werden. OLSE ist die von der SAS Prozedur GLM  
2 verwendete Methode, während die MIXED Prozedur auf GLSE basiert. Die mit beiden  
3 Methoden erhaltenen Parameterschätzwerte stimmen in der Regel nicht überein, mit  
4 Ausnahme von Spezialfällen (Searle et al. 1992), z.B. bei balancierten Daten (Piepho und  
5 Spilke 1999).

6

### 7 ***4.3 Schätzen der Varianzkomponenten***

8

9 Es gibt verschiedene Methoden zur Schätzung der Varianzkomponenten (Searle et al. 1992).  
10 Restricted maximum likelihood (REML) (Patterson und Thompson 1971) hat sich  
11 mittlerweile als Methode der Wahl durchgesetzt, und es ist die Voreinstellung bei den meisten  
12 Statistikpaketen für gemischte Modelle. Die Ausgabe von REML-basierten Prozeduren für  
13 gemischte Modelle umfasst meist eine Reihe deskriptiver Maßzahlen, die sich aus der  
14 maximierten Likelihood (oder Log-Likelihood) berechnen, welche weniger geläufig ist, wenn  
15 man vornehmlich die klassische Varianzanalyse gewohnt ist. Diese statistischen Maßzahlen  
16 können verwendet werden, um die Güte der Modellanpassung zu beurteilen und um unter  
17 einer Menge verschiedener Modellkandidaten eine Auswahl zu treffen (siehe Wolfinger  
18 1996). Sie sind vor allem dann relevant, wenn das gemischte Modell nicht vom ANOVA-Typ  
19 ist.

20

21 Wenn ein zufälliger Effekt wenige Stufen hat, so wird die Varianzschätzung ungenau. In  
22 diesem Fall sollte man prüfen, ob der betreffende Effekt entsprechend Regel 8 auch als fest  
23 betrachtet werden kann. Für balancierte Daten ist es außerdem sinnvoll, die Nicht-  
24 Negativitätsbedingung für die Varianzkomponentenschätzungen aufzuheben (NOBOUND  
25 Option in SAS PROC MIXED). Dies führt zu einer generellen Übereinstimmung von REML-  
26 und sog. ANOVA-Schätzern (Searle et al. 1992), welche bei balancierten Daten optimal sind

1 und zu exakten Wald-Tests für die festen Effekte führen (Piepho und Spilke 1999).

2

#### 3 **4.4 ANOVA vs. REML-basierte Wald-Tests**

4

5 Tests der festen Effekte in einem gemischten Modell können auf zwei verschiedenen Wegen  
6 durchgeführt werden, wobei die Ergebnisse nicht notwendigerweise übereinstimmen. Die eine  
7 Möglichkeit beruht auf einer klassischen Quadratsummenzerlegung in einzelne  
8 Streuungsursachen. F-Tests werden basierend auf den Erwartungswerten der Mittelquadrate  
9 der Varianzanalyse konstruiert. Diese Erwartungswerte können nach einem allgemeinen  
10 Algorithmus berechnet werden, der z.B. bei Milliken und Johnson (1984) beschrieben ist.  
11 Hierbei werden F-Statistiken als Quotient von linearen Kombinationen der Mittelquadrate der  
12 Varianzanalyse so konstruiert, dass Zähler und Nenner bei Gültigkeit der Nullhypothese  
13 denselben Erwartungswert haben. Unter der Alternative kommt im Nenner ein Term hinzu,  
14 der vom zu prüfenden Effekt abhängt. Dieses Verfahren ist beispielsweise in der SAS  
15 Prozedur GLM implementiert. Es ist zu bevorzugen gegenüber *ad hoc* Verfahren wie dem  
16 von Heyland and Kochs (1978) vorgeschlagenen. Falls Zähler oder Nenner mehr als ein  
17 Mittelquadrat umfassen, kann die Satterthwaite-Methode (Milliken and Johnson 1984)  
18 verwendet werden, um approximative Freiheitsgrade zu berechnen.

19

20 Die zweite Methode berechnet F-Statistiken vom Wald-Typ basierend auf gewichteten  
21 Kleinstquadratschätzungen (GLSE) der festen Effekte, wobei REML-Schätzungen der  
22 Varianzkomponenten verwendet werden (Littell et al. 1996). Dies Verfahren ist in der SAS  
23 Prozedur MIXED implementiert. Die Ausgabe von MIXED stimmt mit der klassischen  
24 Varianzanalyse insofern überein, als für jeden festen Effekt ein F-Test berechnet wird. Ein  
25 Unterschied ist das Fehlen einer Quadratsummenzerlegung. Die Verteilung der F-Statistik  
26 unter der Nullhypothese muss in der Regel approximiert werden. Die derzeit beste

1 Approximation wurde von Kenward und Roger (1997) vorgeschlagen und ist über die DDFM  
2 Option der MIXED Prozedur verfügbar (Spilke et al. 2003).

3

4 Für balancierte Daten liefern der varianzanalytische F-Test und der REML-basierte Wald-  
5 Test identische Ergebnisse, sofern die REML-Schätzung ohne Nicht-Negativitäts-Bedingung  
6 für die Varianzkomponenten durchgeführt wird (Piepho und Spilke 1999). Bei unbalancierten  
7 Daten differieren die Resultate dagegen, und es ist nicht ohne weiteres klar, welche der beiden  
8 Verfahren in diesem Fall zu bevorzugen ist. Wald-Test nutzen die Methode der gewichteten  
9 Kleinsten Quadrate (GLSE), welche nur im Fall bekannter Varianzkomponenten optimal ist,  
10 ein in der Praxis eher seltener Fall.

11

12 In der Analyse linearer Modelle mit nur einem Fehlerterm können die Quadratsummen mit  
13 verschiedenen Methoden berechnet werden. Die gängigsten Verfahren sind die sequentiellen  
14 Quadratsummen (Type I SS in SAS) und Typ III Quadratsummen (Type III SS in SAS)  
15 (Searle 1987). Diese Verfahren werden ebenfalls zur Auswertung gemischter Modelle  
16 verwendet. Ebenso können Wald-F-Tests zur Prüfung von Typ I oder Typ III Hypothesen  
17 definiert werden. Nelder (1994) setzt sich intensiv mit Typ I und Typ III Hypothesentests  
18 auseinander, und er liefert starke Argumente zu Gunsten von Type I Hypothesentests, z.B.  
19 bessere Teststärke. Typ III Hypothesen sind vor allem dann relevant, wenn Haupteffekte  
20 geprüft werden sollen, obwohl signifikante Wechselwirkungen vorliegen. Der Gegensatz  
21 zwischen Typ III und Typ I Hypothesen löst sich weitgehend auf, wenn man Haupteffekte  
22 und Randmittelwerte nur dann prüft, wenn Wechselwirkungen nicht signifikant sind. Details  
23 hierzu findet man bei Nelder (1994).

24

25

26

## 1 **4.5 Mittelwertvergleiche**

2

3 Behandlungsmittelwerte können aus den Kleinst-Quadrat-Schätzungen der festen Effekte  
4 berechnet werden (OLSE oder GLSE). Diese Mittelwerte werden auch als Kleinst-Quadrat-  
5 Mittelwerte bezeichnet (least square means - LS-means). Bei balancierten Daten ist der  
6 Standardfehler einer Differenz (s.e.d.) unter den Kleinst-Quadrat-Mittelwerten identisch für  
7 jeden Paarvergleich. In diesem Fall kann eine gemeinsame kritische Differenz berechnet  
8 werden. Fisher's Grenzdifferenz (least significant difference) wird dann beispielsweise  
9 berechnet nach  $LSD = t \times s.e.d.$ , wobei  $t$  der kritische Wert einer t-Verteilung mit den jeweils  
10 gültigen Freiheitsgraden ist. Außerdem kann in diesen Fällen eine Linien- oder  
11 Buchstabendarstellung mit Standardmethoden erhalten werden (Steel and Torrie 1980).

12

13 Bei unbalancierten Daten ist der Standardfehler einer Differenz dagegen nicht konstant für  
14 verschiedene Vergleiche, so dass keine gemeinsame kritische Differenz existiert. Ebenso kann  
15 eine Liniendarstellung nicht mit Standardverfahren berechnet werden. Dies ist enttäuschend  
16 für diejenigen, die Liniendarstellungen und Grenzdifferenzen gewohnt sind. Einige Statistik-  
17 Pakete, z.B. GENSTAT und ASREML, geben einen durchschnittlichen s.e.d. über alle  
18 Paarvergleiche aus, der als Genauigkeitsmaßzahl anstelle einer Grenzdifferenz verwendet  
19 werden kann. Die MIXED Prozedur von SAS gibt alle paarweisen Vergleiche zeilenweise  
20 untereinander aus, zusammen mit einem s.e.d. für jeden Vergleich, aber keinen  
21 durchschnittlichen s.e.d. (dieser kann in einem anschließenden Datenschnitt berechnet  
22 werden). Piepho (2003) hat einen Algorithmus entwickelt, der eine Buchstabendarstellung für  
23 unbalancierte wie für balancierte Daten erzeugt. Ein SAS Makro ist auf Nachfrage vom Autor  
24 erhältlich.

25

26 Die GLM Prozedur von SAS berechnet Kleinst-Quadrat-Mittelwerte (LS-means) basierend

1 auf einfachen Kleinst-Quadrat-Schätzungen (OLSE). Im allgemeinen werden keine adäquaten  
2 s.e.d. erhalten. Daher sollte die GLM Prozedur nicht für Mittelwertvergleiche verwendet  
3 werden. Die MIXED Prozedur verwendet dagegen gewichtete Kleinste Quadrate um  
4 Mittelwerte zu berechnen. Standardfehler werden basierend auf REML Schätzungen der  
5 Varianzkomponenten berechnet, und Freiheitsgrade können mit der Methode von Kenward  
6 und Roger (1997) berechnet werden.

7

8 Multiples Testen beinhaltet die Gefahr eines erhöhten Fehlers 1. Art. Für balancierte Daten  
9 gibt es zahlreiche Verfahren zur Einhaltung des versuchsbezogenen Fehlers 1. Art (family-  
10 wise error rate - FWE), z.B. den Tukey-Test (Hsu 1996). Diese Verfahren sind nicht  
11 anwendbar bei unbalancierten gemischten linearen Modellen. Eine allgemein anwendbare  
12 Methode zur Einhaltung des versuchsbezogenen Fehlers 1. Art ist das Simulationsverfahren  
13 von Edwards und Berry (1987), welches in der MIXED Prozedur implementiert ist. Eine  
14 Reihe alternativer Verfahren wird bei Westfall et al. (1999) diskutiert. Die Autoren stellen  
15 verschiedene SAS Makros bereit.

16

#### 17 ***4.6 Vorhersage zufälliger Effekte (BLUP)***

18

19 Es gibt eine wichtige Ausnahme von Regel 1. Ein Behandlungsfaktor kann aufgrund der  
20 Stichprobennahme auch dann als zufällig betrachtet werden, wenn besonderes Interesse an  
21 den im Versuch geprüften Stufen des Faktors besteht. Wenn die Zahl der Stufen groß ist, ist  
22 es von Vorteil, den Faktor als zufällig zu betrachten und Schätzwerte der zufälligen Effekte  
23 unter Annahme dieses Modells zu berechnen (Searle et al. 1992, p.18). Ein prominentes  
24 Beispiel sind pflanzenzüchterische Versuche, in denen eine große Zahl von Linien evaluiert  
25 wird, die auf eine einzige Kreuzung zurückgehen. Eine in der Tier- und Pflanzenzüchtung  
26 populäre Methode zur Schätzung zufälliger Effekte ist bekannt unter der Bezeichnung "best

1 linear unbiased prediction" (BLUP) (Searle et al. 1992, Mrode 1998). BLUPs können  
2 effizienter sein als Schätzer, welche feste Behandlungseffekte annehmen. In den meisten  
3 pflanzenbaulichen Versuchen sind die Stufen des Behandlungsfaktors gezielt ausgewählt und  
4 die Zahl der Stufen ist klein, so dass die Annahme zufälliger Effekte nicht sinnvoll ist.

5

## 6 **5. Ein kurzer Überblick über Erweiterungen des varianzanalytischen** 7 **gemischten Modells mit Relevanz für Pflanzenbauwissenschaften**

8

9 Dieser Beitrag hat sich auf varianzanalytische gemischte Modelle konzentriert. Es gibt eine  
10 Reihe von Erweiterungen dieser Art von gemischtem Modell, die hier kurz angesprochen  
11 werden sollen. Eine umfassende Darstellung würde hier den Rahmen sprengen, und der  
12 interessierte Leser wird auf die einschlägige Literatur verwiesen.

13

14 Oftmals werden wiederholte Messungen auf derselben Untersuchungseinheit vorgenommen,  
15 wobei die Wiederholung entweder räumlicher oder zeitlicher Art oder beides ist. Wiederholte  
16 Messungen erfordern besondere Aufmerksamkeit wegen der Korrelation, die dadurch  
17 induziert wird, dass dieselbe Versuchseinheit Gegenstand mehrerer nicht-randomisierter  
18 Messungen ist.

19

20 **Beispiel 13:** Vier Düngerarten werden in einer randomisierten vollständigen Blockanlage mit  
21 einer perennierenden Grassart geprüft. Auf jeder Parzelle wird der Ertrag in mehreren  
22 aufeinander folgenden Jahren erfasst, wodurch ein zweiter ("wiederholter")  
23 Behandlungsfaktor entsteht. Beobachtungen, die auf derselben Parzelle in verschiedenen  
24 Jahren gemacht werden, sind miteinander korreliert. Es ist zu erwarten, dass die Korrelation  
25 zwischen benachbarten Jahren enger ist als die zwischen weiter entfernten Jahren.

26

1 Die Analyse von wiederholten Messungen muss der serialen Korrelation von Beobachtungen  
2 auf derselben Versuchseinheit Rechnung tragen. Es gibt verschiedene Modelle, die für die  
3 Korrelationsstruktur angenommen werden können, beispielsweise autoregressive und  
4 räumliche Modelle. Diese Korrelationsstrukturen können im Rahmen eines gemischten  
5 Modells angepasst werden, aber die resultierenden Modelle sind nicht mehr vom ANOVA-  
6 Typ (Diggle et al. 1996, Davidian und Giltinan 1995, Hand und Crowder 1996, Verbeke und  
7 Molenberghs 1997, 2000, Schabenberger und Pierce 2000). Gute Schlüsselwörter für eine  
8 Literaturrecherche zu diesen Modellen sind *spatial statistics*, *geostatistics*, *time series* und  
9 *longitudinal data*. Langfristig angelegte Fruchtfolgeversuche beinhalten eine Reihe  
10 spezifischer Probleme hinsichtlich wiederholter Messungen, welche spezielle Modellansätze  
11 erfordern (Singh et al. 1997).

12

13 Eine einfache Methode für die Analyse von Messwiederholungen besteht in der Berechnung  
14 von zusammenfassenden Maßzahlen für die jeweils kleinste randomisierte  
15 Beobachtungseinheit (Parzelle, Gefäß, etc.). Dies führt zu jeweils einem Wert pro Einheit, so  
16 dass die aggregierten Daten mittels gemischter Modelle vom ANOVA-Typ ausgewertet  
17 werden können (Diggle et al. 1996). Phytopathologen verwenden beispielsweise die Fläche  
18 unter der Befallskurve' als integrale Maßzahl für die Befallsintensität über einen definierten  
19 Zeitraum (Campbell and Madden 1990). Räumliche oder zeitliche Trends über eine kurze  
20 Serie von Messwiederholungen können oft durch einfache lineare Regression erfasst werden,  
21 indem je Versuchseinheit die Steigung der Regression berechnet wird (Bürkert et al. 2002).  
22 Alternativ können Parameterschätzwerte einer nichtlinearen Regression zur Erfassung des  
23 Trends verwendet werden. Oftmals ist auch der einfache Mittelwert über eine Serie von  
24 Messungen eine sinnvolle Maßzahl, z.B. in einem mehrjährigen Beweidungsversuch.

25

26 Viele Autoren haben vorgeschlagen, dass wiederholte Messungen so ausgewertet werden

1 können, als wäre der "wiederholte" Behandlungsfaktor der Kleinteilstückfaktor einer  
2 Spaltanlage ('split-plot in time', Peterson 1994, Steel und Torrie 1980). Dieser Ansatz ist  
3 kritisch einzuschätzen, vor allem deswegen, weil das implizierte Modell nicht mittels  
4 Randomisationstheorie begründet werden kann (Hinkelmann und Kempthorne 1994). Seine  
5 Popularität begründet sich vor allem in der relativ einfachen Analyse. Mit der Verfügbarkeit  
6 rechenstarker Computer und leistungsfähiger Statistik-Pakete für gemischte Modelle ist dieser  
7 Einsatz weitgehend überflüssig geworden. Das Modell für eine Spaltanlage impliziert eine  
8 sehr einfache Korrelationsstruktur für die wiederholten Messungen, und strenggenommen ist  
9 diese Struktur nur gerechtfertigt, wenn sie zu einer besseren Modellanpassung führt als  
10 andere, komplexere Modelle (Wolfinger 1996).

11

12 Die Auswertung wiederholter Messungen mittels gemischter Modelle steht im Gegensatz zur  
13 multivariaten Varianzanalyse (multivariate analysis of variance - MANOVA) (Cole und  
14 Grizzle 1966). Der MANOVA-Ansatz hat zwei wichtige Nachteile: (i) Der für die  
15 Berechenbarkeit der multivariaten Teststatistiken notwendige Stichprobenumfang ist im  
16 allgemeinen größer als für Teststatistiken, die basierend auf einem gemischten Modell  
17 berechnet werden. Der notwendige Stichprobenumfang kann sogar unpraktikabel für  
18 landwirtschaftliche Versuche sein. (ii) Der MANOVA Ansatz kann nur vollständige Serien  
19 von wiederholten Messungen analysieren. Wenn etwa in Beispiel 13 eine der Messungen auf  
20 derselben Parzelle fehlt, müssen alle Messungen von dieser Parzelle von der Analyse  
21 ausgeschlossen werden. Mit einem gemischten Modell können dagegen alle Daten in die  
22 Auswertung einfließen. Der wesentliche Vorteil des MANOVA-Ansatzes besteht in den im  
23 Vergleich zum gemischten Modell weniger restriktiven Annahmen. Details findet man z.B.  
24 bei Hand und Crowder (1996) sowie Diggle et al. (1996).

25

26 Große Feldversuche führen zu großen Problemen hinsichtlich der Fehlerkontrolle, und

1 klassische Blockanlagen sind oft wenig effektiv. Eine alternative Methode besteht in der  
2 Verwendung räumlich-statistischer Modelle für die Korrelation benachbarter Parzellen. Viele  
3 der geo-statistischen Methoden können im Rahmen eines gemischten Modells implementiert  
4 werden (Cullis und Gleeson 1991, Gleeson 1997, Gilmour et al. 1997, Wu et al. 1998).  
5 Räumlich-statistische Modelle können auch mit Glättungs- und Interpolationsverfahren zur  
6 Modellierung großräumiger Trends kombiniert werden (Verbyla et al. 1999).  
7  
8 In der vorliegenden Arbeit wurde die Behandlungsstruktur weitgehend ignoriert, während das  
9 Hauptaugenmerk auf dem Blockmodell lag, da dieses mehrere zufällige Effekte umfassen  
10 kann, so dass das lineare Modell vom gemischten Typ ist. Der Einfachheit halber wurden alle  
11 Behandlungsfaktoren als qualitativ betrachtet. Wenn ein Behandlungsfaktor quantitativ ist, wie  
12 beispielsweise ein in verschiedenen Aufwandmengen geprüftes Düngemittel oder Pestizid, so  
13 kann zur Auswertung eine Regression verwendet werden. Die Regression im Rahmen eines  
14 gemischten Modells hat eine varianzanalytische Form (ANOVA-Typ), sofern alle  
15 Regressionskoeffizienten fest sind. Wenn einige der Koeffizienten zufällig sind, so sind  
16 weitergehende Modelle notwendig (Wolfinger 1996, Verbeke und Molenberghs 1997).  
17  
18 Die Auswertung von Versuchsserien wird erschwert durch das regelmäßige Auftreten von  
19 Behandlung-Umwelt-Interaktionen. Gemischte Modelle vom ANOVA-Typ sind nicht immer  
20 befriedigend zur Modellierung von Interaktionen, und verschiedene Erweiterungen sind  
21 vorgeschlagen worden, z.B. heteroskedastische (Denis et al. 1997, Frensham et al. 1997,  
22 Piepho 1999a) und multiplikative (Piepho 1997, Smith et al. 2001) Modelle. In diesen  
23 Zusammenhang fallen auch Verfahren zur Erfassung der phänotypischen Stabilität.  
24 Maßzahlen für die phänotypische Stabilität, welche die Variabilität des Ertrages unter  
25 wechselnden Umweltbedingungen erfassen, sind sehr populär in der Pflanzenzüchtung und im  
26 Pflanzenbau. Viele dieser Maßzahlen können als Funktionen von Parametern eines

1 gemischten Modells aufgefasst werden (Piepho 1998 and 1999a), so dass gemischte Modelle  
2 eine Schlüsselrolle bei der Analyse der phänotypischen Stabilität spielen. Desweiteren können  
3 räumliche (geostatistische) Modelle mit weiterführenden Modellen für Behandlung-Umwelt-  
4 Interaktion im Rahmen eines gemischten Modells kombiniert werden (Smith et al. 2001).

5

6 Gemischte Modelle sind auch für den Fall nicht-normalverteilter Daten erweitert worden. Die  
7 Erweiterung ist unter der Bezeichnung "generalisiertes gemischtes lineares Modell" geläufig  
8 (Schabenberger und Pierce 2000, McCulloch und Searle 2001), und es ist hoch relevant für  
9 Problemdaten wie Zählwerte (z.B. bei Unkräutern und Insekten) und Prozentzahlen (z.B.  
10 Befallsintensität, Unkrautdeckungsgrad oder Auflaufraten) (Piepho 1999b). Darüber hinaus  
11 können nicht-lineare gemischte Modelle verwendet werden, um intrinsisch nicht-lineare  
12 Zusammenhänge zu modellieren (Davidian and Giltinan 1995), z.B. Modelle für das  
13 Wachstum von Pflanzen (Gregoire und Schabenberger 1996, Schabenberger und Pierce  
14 2000).

15

## 16 **6. Danksagung**

17

18 Wir sind Hanspeter Thöni (Universität Hohenheim) und Christel Richter (Humboldt-  
19 Universität Berlin) dankbar für hilfreiche Kommentare zu früheren Manuskriptversionen.  
20 Unser Dank gilt außerdem Brian Cullis, der uns auf die wichtigen Arbeiten von Chris Brien  
21 aufmerksam gemacht hat.

22

23

24

25

26

## 1 **7. Literatur**

2

3 Bailey, R. A., 1991: Strata for randomized experiments (with discussion). *Journal of the*  
4 *Royal Statistical Society B* **53**, 27-78.

5 Brien, C. J., 1983: Analysis of variance tables based on experimental structure. *Biometrics* **39**,  
6 53-59.

7 Brien, C. J., and R. W. Payne, 1999: Tiers, structure formulae and the analysis of complicated  
8 experiments. *The Statistician* **48**, 41-52.

9 Bürkert, A., H. P. Piepho, and A. Batino, 2002: Multi-site time trend analysis of crop residue,  
10 phosphorus, nitrogen and legume rotation effects on cereal yields in sub-Saharan West Africa.  
11 *Experimental Agriculture* **38**, 163-183.

12 Campbell, and L. V. Madden, 1990: *Introduction to plant disease epidemiology*. Wiley, New  
13 York.

14 Cochran, W. G., and G. M. Cox, 1957: *Experimental designs*. Wiley, New York.

15 Cole, J. W. L., and J. E. Grizzle, 1966: Applications of multivariate analysis of variance to  
16 repeated measures experiments. *Biometrics* **22**, 810-828.

17 Cullis, B. R. , and A. C. Gleeson, 1991: Spatial analysis of field experiments -- An extension  
18 to two dimensions. *Biometrics* **47**, 1449-1460.

19 Davidian, M., and D. M. Giltinan, 1995: *Nonlinear models for repeated measurement data*.  
20 Chapman and Hall, London.

21 Denis, J. B., H. P. Piepho, and F. A. van Eeuwijk, 1997: Modelling expectation and variance  
22 for genotype by environment data. *Heredity* **79**, 162-171.

23 Diggle, P. J., K. Y. Liang, and S.L. Zeger, 1996: *Analysis of longitudinal data*. Clarendon  
24 Press, London.

25 Edwards, D., and J. J. Berry, 1987: The efficiency of simulation-based multiple comparisons.  
26 *Biometrics* **43** , 913-928.

1 Federer, W. T., and R. D. Wolfinger, 1998: SAS code for recovering intereffect information  
2 in experiments with incomplete block and lattice rectangle designs. *Agronomy Journal* **90**,  
3 545-551.

4 Federer, W. T., M. Reynolds, and J. Crossa, 2001: Combining results from augmented  
5 designs over sites. *Agronomy Journal* **93**, 389-395.

6 Frensham, A. , B. R. Cullis, and A. P. Verbyla, 1997: Genotype by environment variance  
7 heterogeneity in a two-stage analysis. *Biometrics* **53**, 1373-1383.

8 Gilmour, A. R., B. R. Cullis, and A. P. Verbyla, 1997: Accounting for natural and extraneous  
9 variation in the analysis of field experiments. *Journal of Agricultural, Biological, and*  
10 *Environmental Statistics* **2**, 269-293.

11 Gleeson, A. C. 1997: Spatial analysis. In: R. A. Kempton, and P. N. Fox (eds.), *Statistical*  
12 *methods for plant variety evaluation*, pp. 68-85. Chapman and Hall, London.

13 Gregoire, T. G. and O. Schabenberger, 1996: Nonlinear mixed-effects modeling of  
14 cumulative bole volume with spatially correlated within-tree data. *Journal of Agricultural,*  
15 *Biological and Environmental Statistics* **1**, 107-119.

16 Hand, D. J., and M. Crowder, 1996: *Practical longitudinal data analysis*. Chapman and Hall,  
17 London.

18 Heyland, K.-U., and H.-J. Kochs, 1978: Varianzanalytische Testung mehrfaktorieller  
19 pflanzenbaulicher Versuche. *Zeitschrift für Acker und Pflanzenbau* **146**, 109-119.

20 Hinkelmann, K., and O. Kempthorne, 1994: *Design and analysis of experiments*. Vol. 1:  
21 *Introduction to experimental design*. Wiley, New York.

22 Hocking, R. R., 1985: *The analysis of linear models*. Brooks and Cole, Monterey.

23 Hsu, J. C., 1996: *Multiple comparisons*. Chapman and Hall, London.

24 Kenward, M. G., and J. H. Roger, 1997: Small sample inference for fixed effects from  
25 restricted maximum likelihood. *Biometrics* **53**, 983-997.

26 Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger, 1996: *SAS system for*

1 mixed models. SAS Institute, Cary.

2 McCullagh, P., and J. A. Nelder, 1989: Generalized linear models. 2nd edition. Chapman and  
3 Hall, London.

4 McCulloch, C. E., and S. R. Searle, 2001: Generalized, linear, and mixed models. Wiley, New  
5 York.

6 McLean, R. A. , W. L. Sanders, and W. W. Stroup, 1991: A unified approach to mixed linear  
7 models. *The American Statistician* **45**, 54-64.

8 Mead, R. 1997: Design of plant breeding trials. In: R. A. Kempton, and P. N. Fox (eds.),  
9 *Statistical methods for plant variety evaluation*, pp. 40-67. Chapman and Hall, London.

10 Milliken, G. A., and D. E. Johnson, 1984: Analysis of messy data. Volume 1: Designed  
11 experiments. Chapman and Hall, London.

12 Mrode, R. A., 1998: Linear models for the prediction of animal breeding values. CAB  
13 International, Wallingford.

14 Nelder, J. A., 1965: The analysis of randomized experiments with orthogonal block structure.  
15 I. Block structure and the null analysis of variance. II. Treatment structure and the general  
16 analysis of variance. *Proceedings of the Royal Society of London A* **283**, 147-178.

17 Nelder, J. A., 1994: The statistics of linear models: back to basics. *Statistics and Computing*  
18 **4**, 221-234.

19 Patterson, H. D., 1997: Analysis of series of variety trials. In: R. A. Kempton, and P. N. Fox  
20 (eds.), *Statistical methods for plant variety evaluation*, pp. 139-161. Chapman and Hall,  
21 London.

22 Patterson, H. D., and R. Thompson, 1971: Recovery of inter-block information when block  
23 sizes are unequal. *Biometrika* **58**, 545-554.

24 Payne, R. W., and G. N. Wilkinson, 1977: A general algorithm for analysis of variance.  
25 *Applied Statistics* **26**, 251-260.

26 Peterson, R. G., 1994: Agricultural field experiments. Design and analysis. Marcel Dekker,

1 New York.

2 Piepho, H.P., 1997: Analyzing genotype-environment data by mixed models with  
3 multiplicative effects. *Biometrics* **53**, 761-766.

4 Piepho, H. P., 1998: Methods for comparing the yield stability of cropping systems - A  
5 review. *Journal of Agronomy and Crop Science* **180**, 193-213.

6 Piepho, H. P., 1999a: Stability analysis using the SAS system. *Agronomy Journal* **91**, 154-  
7 160.

8 Piepho, H. P., 1999b: Analysing disease incidence data from designed experiments by  
9 generalized linear mixed models. *Plant Pathology* **48**, 668-674.

10 Piepho, H. P., 2003: An algorithm for representing all pairwise comparisons. submitted

11 Piepho, H. P., and J. Spilke, 1999: Anmerkungen zur Analyse balancierter gemischter  
12 Modelle mit der SAS Prozedur MIXED. *Zeitschrift für Agrarinformatik* **7**, 39-46.

13 Schabenberger, O., and F. J. Pierce, 2000: Contemporary statistical models. CRC Press, Boca  
14 Raton.

15 Searle, S. R., 1987: Linear models for unbalanced data. Wiley, New York.

16 Searle, S. R., G. Casella, and C. E. McCulloch, 1992. Variance components. Wiley, New  
17 York.

18 Singh, M., S. Christiansen, and B. K. Chakraborty, 1997: An assessment of the effect of  
19 covariances of plot errors over time on the precision of means of rotation experiments.  
20 *Experimental Agriculture* **33**, 469-475.

21 Smith, A., B. R. Cullis, and R. Thompson, 2001: Analyzing variety by environment data  
22 using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* **57**,  
23 1138-1147.

24 Spilke, J., X. Hu, and H. P. Piepho, 2003: A simulation study on tests of hypotheses for fixed  
25 effects in mixed models for blocked experiments with missing data. submitted

26 Steel, R. G. D., and J. H. Torrie, 1980: Principles and procedures of statistics: a biometrical

- 1 approach. Second edition. McGraw-Hill, New York.
- 2 Verbeke, G., and G. Molenberghs (eds), 1997: Linear mixed models in practice. A SAS-  
3 oriented approach. Springer, Berlin.
- 4 Verbeke, G., and G. Molenberghs, 2000: Linear mixed models for longitudinal data. Springer,  
5 Berlin.
- 6 Verbyla, A. P., B. R. Cullis, M. G. Kenward, and S. J. Welham, 1999: The analysis of  
7 designed experiments and longitudinal data by using smoothing splines. *Applied Statistics* **48**,  
8 269-300.
- 9 Westfall, P. H., R. D. Tobias, D. Rom, R. D. Wolfinger, and Y. Hochberg, 1999: Multiple  
10 comparisons and multiple tests. SAS Institute, Cary.
- 11 Wilkinson, G. N., and C. E. Rogers, 1973: Symbolic description of factorial models for  
12 analysis of variance. *Applied Statistics* **22**, 392-399.
- 13 Wolfinger, R.D., 1996: Heterogeneous variance-covariance structures for repeated measures.  
14 *Journal of Agricultural, Biological and Environmental Statistics* **1**, 205–230.
- 15 Wu, T. X., D. E. Mather, and P. Dutilleul, 1998: Application of geostatistical and neighbor  
16 analyses to data from plant breeding trials. *Crop Science* **38**, 1545-1553.

**Tabelle 1:** Zusammenfassung der Regeln für die Formulierung eines gemischten Modells vom ANOVA-Typ.

Regel	Schlüsselsatz	Seite
1	Wann ist ein Faktor zufällig?	8
2	Zwei Faktorarten	9
3	Behandlungs- und Blockmodell getrennt halten, jedenfalls zu Beginn	11
4	Effekte des Blockmodells	11
5	Kodierung der Blockfaktoren	12
6	Interaktion zwischen Block und Behandlungsfaktoren	15
7	Mehrphasige Experimente	16
8	Zufällige Effekte als fest betrachten	19

$a_1$				$a_2$				$a_3$				$a_4$			
$b_3$	$b_2$	$b_1$	$o_1$	$b_1$	$b_3$	$b_2$	$o_3$	$b_3$	$b_1$	$b_2$	$o_2$	$b_2$	$b_3$	$b_1$	$o_2$
$b_1$	$b_3$	$b_2$	$o_2$	$b_2$	$b_1$	$b_3$	$o_2$	$b_1$	$b_2$	$b_3$	$o_3$	$b_1$	$b_2$	$b_3$	$o_1$
$b_2$	$b_1$	$b_3$	$o_3$	$b_3$	$b_2$	$b_1$	$o_1$	$b_2$	$b_3$	$b_1$	$o_1$	$b_3$	$b_1$	$b_2$	$o_2$
I	II	III		III	I	II		II	I	III		I	III	II	

**Abb. 1:** Versuchsplan für Phase II (Labor) mit vier Lateinischen Quadraten für Faktor B (chemische Methode) mit Stufen  $b_1$ ,  $b_2$  und  $b_3$  (Beispiel 10). Für jede Stufe des Faktors A (Erntezeitpunkt;  $a_1, a_2, a_3, a_4$ ) gibt es ein Lateinisches Quadrat. Für jedes Lateinische Quadrat entsprechen die Zeilen den Öfen ( $o_1, o_2, o_3$ ), während die Spalten den Blöcken (I, II, III) aus Phase I entsprechen. Die Öfen in den Lateinischen Quadraten sind jeweils dieselben.