

# Spaltanlage - Messwiederholung - Dauerversuch: Hinweise zur Auswertung komplexer pflanzenbaulicher Versuche mit gemischten Modellen

Hans-Peter Piepho und Andreas Büchse

Skript zu einem Workshop gehalten bei der Jahrestagung der Gesellschaft für Pflanzenbauwissenschaften am 25.9.2003 in Giessen

## Einleitung

Viele pflanzenbauliche Versuche werden nach komplexen Designs angelegt, die mehrere Randomisationsschritte beinhalten, z.B. Spalt- und Streifenanlagen. Das für die Auswertung adäquate Modell umfasst für jeden Randomisationsschritt einen oder mehrere zufällige Effekte, so dass ein sog. gemischtes Modell vorliegt (mixed model). Hinzu kommt oft eine über mehrere Zeitpunkte wiederholte Erfassung von Zielvariablen, so z.B. bei mehrschnittigen und/oder mehrjährigen Grünlandversuchen sowie bei Dauerversuchen. Wegen der wiederholten Messung müssen bei der Analyse Korrelationen zwischen den auf der gleichen Parzelle oder an den gleichen Pflanzen erhobenen Daten berücksichtigt werden. Eine ähnliche Problematik ergibt sich bei Versuchen mit Bodenproben in mehreren Tiefen. Für solche Fälle stehen verschiedene Erweiterungen des gemischten Modells zur Verfügung. Darüber hinaus gibt es eine ganze Reihe weiterer Versuchsanlagen und -strukturen, die zur Auswertung ein gemischtes Modell erfordern, z.B. Versuchsserien und Nächstnachbaranalysen großer Sortenversuche.

Was ist ein gemischtes Modell? Diese Frage beantwortet sich am einfachsten ausgehend von einem einfachen Beispiel, das kein gemischtes Modell ist, nämlich dem Modell für eine Blockanlage:

$$y_{ij} = \mu + \beta_j + \tau_i + e_{ij}$$

$y_{ij}$  = Ertrag der  $i$ -ten Behandlung im  $j$ -ten Block

$\mu$ = Gesamtmittelwert	}	fest
$\beta_j$ = Effekt des $j$ -ten Blocks		
$\tau_i$ = Effekt der $i$ -ten Behandlung		
$e_{ij}$ = Fehler	}	zufällig

Das Modell hat drei Arten von Effekten: Blockeffekte, Behandlungseffekte und Fehler. Block- und Behandlungseffekte sind in der Regel sog. feste Effekte, während Fehlereffekte wegen der Randomisation (!) zufällig sind.

Das Modell für die Blockanlage hat nur einen zufälligen Effekt. Es gibt kompliziertere Modelle, die mehr als einen zufälligen Effekt haben. Solche Modelle sind gemischte Modelle. Modelle mit nur einem Fehlerterm werden dagegen in der Regel nicht als gemischte Modelle bezeichnet.

Anwender sind meist vertraut mit varianzanalytischen Auswertungen auf Basis linearer Modelle, z.B. mit PROC GLM in SAS. Dagegen ist die Verwendung von Prozeduren für gemischte Modelle wie z.B. PROC MIXED in SAS viel weniger geläufig, als es das regelmäßige Auftreten solcher Versuchsanlagen erfordern würde. Ziel des Workshops ist daher eine anwendungsorientierte Einführung in gemischte Modelle, die deutlich macht, dass es im Vergleich zwischen gemischten Modellen und altbekannter Varianzanalyse viel mehr Gemeinsamkeiten als Neuerungen gibt und dass gemischte Modelle ein sehr flexibles Instrumentarium zur Auswertung pflanzenbaulicher Versuche bieten. Anhand einiger typischer Beispiele sollen die wesentlichen Analyseschritte erläutert werden, wobei zu Beginn eine Streifenanlage aus Standardfall relativ ausführlich behandelt wird, um die grundlegenden Aspekte darzulegen. Danach wird auf einige weiterführende Beispiele eingegangen, wobei jeweils einzelne besondere Aspekte beleuchtet werden.

## Beispiel 1: Streifenanlage

Die langfristige Auswirkung von 2-, 3-, 4-, und 5 Schnittnutzung auf Pflanzenbestandszusammensetzung, TM-Ertrag und Futterinhaltsstoffe eines aus einer einheitlichen Saatmischung hervorgegangenen Grünlandfläche wurde in einem Feldversuch geprüft (Dr. U. Thumm, Institut für Pflanzenbau und Grünland, Hohenheim). Hierbei wurden zwei Faktoren untersucht: Nutzungshäufigkeit mit 4 Stufen (2 bis 5 Schnitte pro Jahr) und N-Düngung mit 2 Stufen (30 und 60 kg N/ha). Aus technischen Gründen wurde der Versuch als Streifenanlage randomisiert. Die Ertragserfassung erfolgt mit einem HALDRUP-Grünfuttermesser. Daher wurden die Varianten mit gleicher Nutzungshäufigkeit, die also stets gleichzeitig geschnitten werden, hintereinander angeordnet (in Streifen). Ansonsten wären zwischen allen Parzellen breite Zwischenwege erforderlich, soviel Versuchsfläche stand jedoch nicht zur Verfügung. Die Düngung erfolgte mit einem Kastendüngerstreuer. Diese Geräte brauchen immer eine Anlaufstrecke (ca. 1-2m) bis gleichmäßig Dünger ausgebracht wird. Durch die Anordnung in Streifen kann auf mehreren Parzellen hintereinander eine einheitliche Düngermenge ohne Unterbrechung ausgestreut werden.

### Kodierung der Faktoren:

N- Düngung: N1 = 30 kg N/ha pro Aufwuchs  
(n\_stufe) N2 = 60 kg N/ha pro Aufwuchs

Nutzungshäufigkeit: S2 = 2 Schnitte/Jahr  
(nutz\_h) S3 = 3 Schnitte/Jahr  
S4 = 4 Schnitte/Jahr  
S5 = 5 Schnitte/Jahr

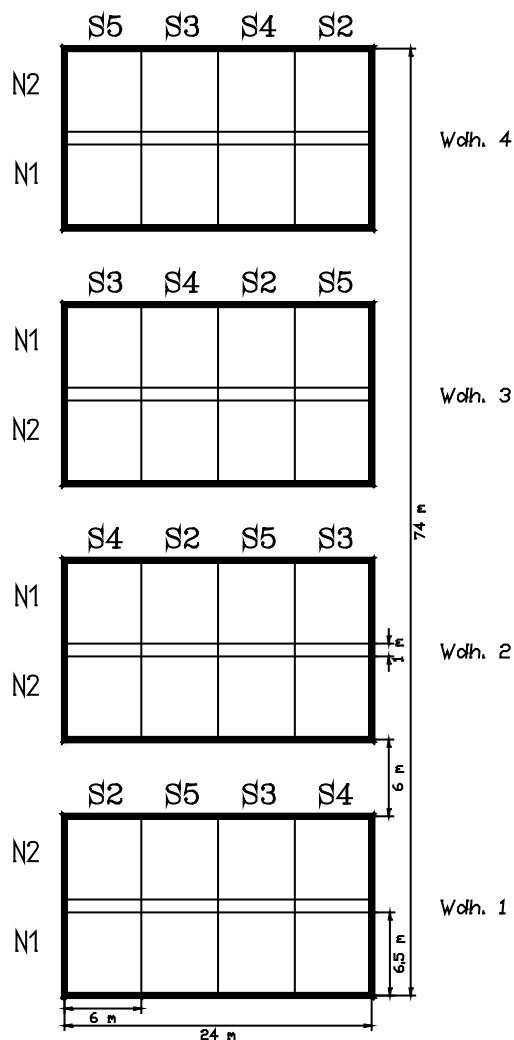
Wiederholungen: 4 Blöcke  
(block)

Der Versuch wurde als Streifenanlage angelegt (Abb. 1.1). Hierbei wurde wie folgt randomisiert:

- Je Block 2 Zeilen für Faktor N-Düngung und 4 Spalten für Nutzungshäufigkeit
- Je Block N1 und N2 zufällig auf Zeilen verteilen
- Je Block S2 bis S5 zufällig auf Spalten verteilen

Eigenschaften der Streifenanlage:

- Wechselwirkungen werden genauer geprüft als Hauptwirkungen



**Abb. 1.1:** Feldplan für Streifenanlage.

Die Daten sind nach den einzelnen Schnitten (**aufwuchs**) aufgeschlüsselt. Zunächst müssen die Aufwüchse auf derselben Parzelle summiert werden, um den Gesamtertrag der Vegetationsperiode zu ermitteln.

Ein Modell für den Gesamterträge je Parzelle kann man getrennt für die Behandlungen und die Versuchseinheiten entwickeln unter der üblichen Annahme, dass keine Wechselwirkung zwischen Behandlung und Blockeffekten besteht (Piepho et al., 2003a).

Behandlungen:

$$\alpha_i + \beta_j + (\alpha\beta)_{ij} \quad , \text{ wobei}$$

$\alpha_i$  = Haupteffekt der  $i$ -ten N-Menge

$\beta_j$  = Haupteffekt der  $j$ -ten Nutzungshäufigkeit

$(\alpha\beta)_{ij}$  = Wechselwirkung Nutzungshäufigkeit  $\times$  N-Menge

Versuchseinheiten (Blockeffekte):

$r_k + z_{ik} + s_{jk} + e_{ijk}$  , wobei

$r_k$  = Effekt des  $k$ -ten Blocks

$z_{ik}$  = Effekt der  $i$ -ten Zeile (N-Menge) im  $k$ -ten Block;  $z_{ik} \sim N(0, \sigma_z^2)$

$s_{jk}$  = Effekt der  $j$ -ten Spalte (Nutzungshäufigkeit) im  $k$ -ten Block;  $s_{jk} \sim N(0, \sigma_s^2)$

$e_{ijk}$  = Restfehler;  $e_{ijk} \sim N(0, \sigma_e^2)$

Der Auswertungsgang nach diesem gemischten Modell ist ähnlich wie bei einer Spaltanlage, d.h., es gibt in der Varianzanalyse verschiedene Fehlerterme im Nenner der F-Statistik, und bei Mittelwertvergleichen sind die Grenzdifferenzen aus verschiedenen Fehler-Mittelquadraten ( $MQ$ s) zu berechnen. Man erhält die "richtige" Auswertung automatisch, wenn man nach einem gemischten Modell mit MIXED auswertet. Die Varianzanalyse-Tabelle hat die folgende Struktur:

**Tabelle 1.1:** Varianzanalyse-Tabelle einer zweifaktoriellen Streifenanlage.

Ursache	Freiheitsgrade ( $FG$ )	$MQ$	$E(MQ)^\S$
Blöcke	$(r - 1)$		
A (N-Menge)	$(a - 1)$		$\sigma_e^2 + b\sigma_s^2 + br \frac{\sum_{i=1}^a [\alpha_i]^2}{(a - 1)}$
Fehler(A)	$(a - 1)(r - 1)$	$E_a$	$\sigma_e^2 + b\sigma_s^2$
B (Nutzungshäufigkeit)	$(b - 1)$		$\sigma_e^2 + a\sigma_z^2 + ar \frac{\sum_{j=1}^b [\beta_j]^2}{(b - 1)}$
Fehler(B)	$(b - 1)(r - 1)$	$E_b$	$\sigma_e^2 + a\sigma_z^2$
A x B	$(a - 1)(b - 1)$		$\sigma_e^2 + r \frac{\sum_{i=1}^a \sum_{j=1}^b [(\alpha\beta)_{ij}]^2}{(a - 1)(b - 1)}$
Fehler(C)	$(a - 1)(b - 1)(r - 1)$	$E_c$	$\sigma_e^2$

$r$  = Zahl der Blöcke,  $a$  = Zahl der Stufen des Faktors A (N-Menge),  $b$  = Zahl der Stufen des Faktors B (Nutzungshäufigkeit)

§ Für Haupteffekte gelten  $E(MQ)$  für den Fall das keine Wechselwirkung vorliegt.

Außerdem wird mit der folgenden Restriktion gearbeitet:  $\sum_{i=1}^a \alpha_i = \sum_{j=1}^b \beta_j = \sum_{i=1}^a \sum_{j=1}^b (\alpha\beta)_{ij}$

Aus den Erwartungswerten der  $MQ$  ergibt sich, dass N-Menge gegen Fehler(A), Nutzungshäufigkeit gegen Fehler (B) und Nutzungshäufigkeit  $\times$  N-Menge gegen Fehler(C) zu testen ist.

Wir verwenden folgende SAS Anweisungen, um die Varianzanalyse mit GLM zu erhalten.

```

data a;
input
aufwuchs      N_Stufe      Nutz_h      Block      dt_ha;
datalines;
  1           1           2           1           58.4616
  1           1           2           2           72.5483
  1           1           2           3           75.2509
  1           1           2           4           70.4225
  1           1           3           1           38.8985
  1           1           3           2           48.0374
  1           1           3           3           32.4051
  1           1           3           4           33.8352
  1           1           4           1           13.5999
  1           1           4           2           12.6620
  1           1           4           3           15.5651
  1           1           4           4           21.1731
  1           1           5           1           9.1519
  1           1           5           2           15.0174
  1           1           5           3           17.3829
  1           1           5           4           8.4984
  1           2           2           1           56.9507
  1           2           2           2           81.3566
  1           2           2           3           76.6414
  1           2           2           4           64.4613
  1           2           3           1           35.8087
  1           2           3           2           35.9437
  1           2           3           3           38.3631
  1           2           3           4           40.5439
  1           2           4           1           14.9449
  1           2           4           2           12.2055
  1           2           4           3           14.7063
  1           2           4           4           19.6358
  1           2           5           1           10.0524
  1           2           5           2           14.5975
  1           2           5           3           20.4165
  1           2           5           4           11.3968
  2           1           2           1           25.4434
  2           1           2           2           25.0982
  2           1           2           3           21.9189
  2           1           2           4           21.8732
  2           1           3           1           25.3443
  2           1           3           2           22.8939
  2           1           3           3           17.8980
  2           1           3           4           17.6992
  2           1           4           1           35.2326
  2           1           4           2           37.2596
  2           1           4           3           34.5837
  2           1           4           4           38.5288
  2           1           5           1           13.6766
  2           1           5           2           15.0926
  2           1           5           3           18.0105
  2           1           5           4           17.0302

```

2	2	2	1	28.0825
2	2	2	2	31.8989
2	2	2	3	29.1086
2	2	2	4	29.3683
2	2	3	1	24.2449
2	2	3	2	23.7486
2	2	3	3	21.0978
2	2	3	4	22.7213
2	2	4	1	44.0350
2	2	4	2	40.2271
2	2	4	3	45.1668
2	2	4	4	40.3142
2	2	5	1	22.1082
2	2	5	2	19.0585
2	2	5	3	23.0442
2	2	5	4	19.9732
3	1	3	1	22.1728
3	1	3	2	19.8071
3	1	3	3	19.7152
3	1	3	4	18.0864
3	1	4	1	13.3374
3	1	4	2	13.0653
3	1	4	3	14.8682
3	1	4	4	15.3952
3	1	5	1	12.5986
3	1	5	2	15.1414
3	1	5	3	13.3397
3	1	5	4	10.1243
3	2	3	1	26.6436
3	2	3	2	21.2608
3	2	3	3	27.0843
3	2	3	4	24.4858
3	2	4	1	19.8501
3	2	4	2	20.8244
3	2	4	3	20.7007
3	2	4	4	19.7127
3	2	5	1	19.2076
3	2	5	2	23.1589
3	2	5	3	23.4496
3	2	5	4	18.9526
4	1	4	1	8.7680
4	1	4	2	8.4214
4	1	4	3	6.8829
4	1	4	4	8.8399
4	1	5	1	8.8810
4	1	5	2	8.6831
4	1	5	3	10.2463
4	1	5	4	8.7041
4	2	4	1	12.9199
4	2	4	2	12.8284
4	2	4	3	13.4860
4	2	4	4	12.7340
4	2	5	1	13.3933
4	2	5	2	10.0650
4	2	5	3	8.3225
4	2	5	4	12.9515
5	1	5	1	5.3005
5	1	5	2	8.6417
5	1	5	3	5.7641
5	1	5	4	7.1087
5	2	5	1	13.4777

```

5          2          5          2          13.5009
5          2          5          3          6.7861
5          2          5          4          11.2373
;
proc sort data=a out=a;
by N_Stufe Nutz_h Block;

proc means data=a;
var dt_ha;
by N_Stufe Nutz_h Block;
output out=b sum=;

proc glm;
class n_stufe nutz_h block;
model dt_ha=n_stufe nutz_h n_stufe*nutz_h block n_stufe*block nutz_h*block;
random n_stufe*block nutz_h*block/test;
run;

```

Ergebnis:

The GLM Procedure

Source	Type III Expected Mean Square
Block	Var(Error) + 4 Var(Block*N_Stufe) + 2 Var(Block*Nutz_h) + Q(Block)
Nutz_h	Var(Error) + 2 Var(Block*Nutz_h) + Q(Nutz_h,N_Stufe*Nutz_h)
N_Stufe	Var(Error) + 4 Var(Block*N_Stufe) + Q(N_Stufe,N_Stufe*Nutz_h)
N_Stufe*Nutz_h	Var(Error) + Q(N_Stufe*Nutz_h)
Block*Nutz_h	Var(Error) + 2 Var(Block*Nutz_h)
Block*N_Stufe	Var(Error) + 4 Var(Block*N_Stufe)

Die Erwartungswerte der  $MQ$  zeigen, dass "NUTZ\_H" gegen den Fehler "BLOCK\*NUTZ\_H" [Fehler(B)] zu prüfen ist, "N\_STUFE" gegen BLOCK\*N\_STUFE" [Fehler(A)] und "NUTZ\_H\*N\_STUFE" gegen den Rest ["Error"; Fehler(C)]. Die F-Tests sind wie folgt:

The GLM Procedure  
Tests of Hypotheses for Mixed Model Analysis of Variance

Dependent Variable: dt\_ha

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Block	3	202.852222	67.617407	1.42	0.4173
Error	2.378	113.384843	47.680929		
Error: MS(Block*Nutz_h) + MS(Block*N_Stufe) - MS(Error)					

Source	DF	Type III SS	Mean Square	F Value	Pr > F
* Nutz_h	3	3176.944124	1058.981375	13.64	0.0011
Error	9	698.586933	77.620770		
Error: MS(Block*Nutz_h)					
* This test assumes one or more other fixed effects are zero.					

Source	DF	Type III SS	Mean Square	F Value	Pr > F
* N_Stufe	1	1311.612312	1311.612312	91.40	0.0024
Error	3	43.051361	14.350454		
Error: MS(Block*N_Stufe)					



\* This test assumes one or more other fixed effects are zero.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
N_Stufe*Nutz_h	3	344.725394	114.908465	2.59	0.1171
Block*Nutz_h	9	698.586933	77.620770	1.75	0.2080
Block*N_Stufe	3	43.051361	14.350454	0.32	0.8081
Error: MS(Error)	9	398.612653	44.290295		

Die Wechselwirkungen sind nicht signifikant, also können Randmittelwerte für NUTZ\_H und N\_STUFE verglichen werden. Hier werden aus illustrativen Gründen die NUTZ\_H\*N\_STUFE Mittelwerte verglichen. Dieselbe Varianzanalyse erhalten wir mit MIXED:

```
proc mixed data=b nobound;
class block n_stufe nutz_h;
model dt_ha=block nutz_h n_stufe nutz_h*n_stufe;
random nutz_h*block n_stufe*block;
run;
```

Ergebnis:

Cov Parm	Estimate
Block*Nutz_h	16.6652
Block*N_Stufe	-7.4850
Residual	44.2903

#### Fit Statistics

-2 Res Log Likelihood	156.8
AIC (smaller is better)	162.8
AICC (smaller is better)	164.2
BIC (smaller is better)	165.1

#### Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
2	2.61	0.2713

#### Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Block	3	3	1.42	0.3905
Nutz_h	3	9	13.64	0.0011
N_Stufe	1	3	91.40	0.0024
N_Stufe*Nutz_h	3	9	2.59	0.1171

Die F-Tests sind dieselben wie mit GLM, aber sie werden anders berechnet:

- (1) Mit Hilfe der Restricted Maximum Likelihood (REML) Methode werden die Varianzen für Zeilen-, Spalten, und Restfehler geschätzt.
- (2) Die Varianzen werden dann verwendet, um die festen Effekte zu schätzen. Die Schätzungen sind hier identisch mit denen von GLM, aber nicht bei unbalancierten Daten.

(3) Mittels der geschätzten Effekte und Varianzen wird eine sog. Wald-F-Statistik berechnet. Diese ist im vorliegenden Fall identisch mit der F-Statistik in GLM, aber nicht in anderen Fällen, z.B. bei unbalancierten Daten.

Man beachte, dass die Varianzkomponente für den Spaltenfehler negativ geschätzt wird (-7.49). Diese Fehlerursache scheint also sehr klein oder nicht existent zu sein.

**Tabelle 1.2:** Varianzen von Mittelwertdifferenzen in einer zweifaktoriellen Streifenanlage (Gomez und Gomez, 1984, S. 203).

Vergleich	Varianz der Differenz ( $V$ ) <sup>*</sup>
A Mittelwerte (N-Stufe)	$2E_a/(br)$
B Mittelwerte (Nutzungshäufigkeit)	$2E_b/(ar)$
A x B Mittelwerte auf gleicher Stufe von B	$2[(b - 1)E_c + E_a]/(br)$
A x B Mittelwerte auf gleicher Stufe von A	$2[(a - 1)E_c + E_b]/(ar)$

\*: Standardfehler der Differenz ergibt sich aus der Quadratwurzel der Varianz  $V$ .  $E_a$ ,  $E_b$  und  $E_c$ : Fehler-MQ aus Varianzanalyse (Tab. 1.1).

Die Mittelwertvergleiche führt man mit MIXED durch, da nur hier die richtigen Standardfehler/Varianzen einer Differenz gefunden werden. Die Varianz einer Differenz berechnet sich wie in Tab. 1.2 angegeben. Die Besonderheit einer Streifenanlage besteht darin, dass die Varianz für den Vergleich von AB-Mittelwerten von zwei  $MQ$  abhängt. Die Grenzdifferenz (least significant difference - LSD) berechnet sich dann generell nach

$$LSD = t \times \sqrt{V}$$

wobei  $t$  der  $t$ -Wert bei den entsprechenden Fehler-Freiheitsgraden ( $FG$ ) ist. In die Berechnung der ersten beiden Vergleiche in Tab. 1.2 geht nur ein  $MQ$  ein, so dass exakte Tests resultieren. Zur Berechnung der Freiheitsgrade für die beiden letzten Vergleiche kann die Methode von Satterthwaite (1941, 1946) angewendet werden (Milliken und Johnson, 1984, S. 281). Hier soll exemplarisch der Vergleich der A-Mittelwerte (N-Menge) auf gleicher Stufe von B (Nutzungshäufigkeit) betrachtet werden. Die Varianz einer Differenz ( $V$ ) berechnet sich als Linearkombination der MQs  $E_a$  und  $E_c$ :

$$V = h_a E_a + h_c E_c,$$

wobei  $h_c = 2(b - 1)/(br)$  und  $h_a = 2/(br)$  ist. Die approximativen Satterthwaite-Freiheitsgrade ( $FG_{Satterth}$ ) berechnen sich zu

$$FG_{Satterth} = \frac{V^2}{(h_a E_a)^2 / FG_a + (h_c E_c)^2 / FG_c},$$

wobei  $FG_a = (a - 1)(r - 1)$  und  $FG_c = (a - 1)(b - 1)(r - 1)$  ist. Wir führen zunächst die Varianzanalyse mit der Prozedur GLM durch, um  $E_a$  und  $E_c$  zu erhalten und daraus die Freiheitsgrade von Hand zu berechnen. Wir finden  $E_a = 43.05$  und  $E_c = 398.61$ . Hiermit berechnen sich  $V = 30,47$  und  $FG_{Satterth} = 16,7$  (Standardfehler der Differenz =  $\sqrt{V} = 5,52$ ). Exakt dieselben Ergebnisse werden mit der Prozedur MIXED unter Angabe der SATTERTH-Option oder der KR-Option für die Kenward-Roger Methode erhalten. Auch für die anderen Vergleiche werden die adäquaten Standardfehler und Freiheitsgrade berechnet.

SAS Anweisungen:

```
ods output diffs=diffs;
proc mixed data=b nobound;
class block n_stufe nutz_h;
model dt_ha=block nutz_h n_stufe nutz_h*n_stufe/ddfm=kr;
random nutz_h*block n_stufe*block;
lsmeans nutz_h*n_stufe/pdiff;
run;
```

```
data diffs;
set diffs;
lsd=StdErr*tinv(0.975,DF);
```

```
proc print data=diffs;
run;
```

Ergebnis:

#### Least Squares Means

Effect	N_Stufe	Nutz_h	Estimate	Standard Error	DF	t Value	Pr >  t
N_Stufe*Nutz_h	1	2	92.7543	3.6562	15.4	25.37	<.0001
N_Stufe*Nutz_h	1	3	79.1983	3.6562	15.4	21.66	<.0001
N_Stufe*Nutz_h	1	4	74.5458	3.6562	15.4	20.39	<.0001
N_Stufe*Nutz_h	1	5	57.0985	3.6562	15.4	15.62	<.0001
N_Stufe*Nutz_h	2	2	99.4671	3.6562	15.4	27.21	<.0001
N_Stufe*Nutz_h	2	3	85.4866	3.6562	15.4	23.38	<.0001
N_Stufe*Nutz_h	2	4	91.0730	3.6562	15.4	24.91	<.0001
N_Stufe*Nutz_h	2	5	78.7876	3.6562	15.4	21.55	<.0001

Obs	Effect	N_Stufe	Nutz_h	N_Stufe	Nutz_h	Estimate	StdErr	DF	tValue	Probt	lsd
1	N_Stufe*Nutz_h	1	2	1	3	13.5560	5.5207	16.7	2.46	0.0253	11.6609
2	N_Stufe*Nutz_h	1	2	1	4	18.2085	5.5207	16.7	3.30	0.0043	11.6609
3	N_Stufe*Nutz_h	1	2	1	5	35.6558	5.5207	16.7	6.46	<.0001	11.6609
4	N_Stufe*Nutz_h	1	2	2	2	-6.7128	4.2898	10.7	-1.56	0.1468	9.4770
5	N_Stufe*Nutz_h	1	2	2	3	7.2676	5.1706	15.4	1.41	0.1797	10.9941
6	N_Stufe*Nutz_h	1	2	2	4	1.6813	5.1706	15.4	0.33	0.7494	10.9941
7	N_Stufe*Nutz_h	1	2	2	5	13.9667	5.1706	15.4	2.70	0.0161	10.9941
8	N_Stufe*Nutz_h	1	3	1	4	4.6525	5.5207	16.7	0.84	0.4112	11.6609
9	N_Stufe*Nutz_h	1	3	1	5	22.0998	5.5207	16.7	4.00	0.0009	11.6609
10	N_Stufe*Nutz_h	1	3	2	2	-20.2688	5.1706	15.4	-3.92	0.0013	10.9941
11	N_Stufe*Nutz_h	1	3	2	3	-6.2884	4.2898	10.7	-1.47	0.1715	9.4770
12	N_Stufe*Nutz_h	1	3	2	4	-11.8747	5.1706	15.4	-2.30	0.0360	10.9941

13 N_Stufe*Nutz_h	1	3	2	5	0.4107	5.1706	15.4	0.08	0.9377	10.9941
14 N_Stufe*Nutz_h	1	4	1	5	17.4473	5.5207	16.7	3.16	0.0058	11.6609
15 N_Stufe*Nutz_h	1	4	2	2	-24.9213	5.1706	15.4	-4.82	0.0002	10.9941
16 N_Stufe*Nutz_h	1	4	2	3	-10.9409	5.1706	15.4	-2.12	0.0510	10.9941
17 N_Stufe*Nutz_h	1	4	2	4	-16.5272	4.2898	10.7	-3.85	0.0028	9.4770
18 N_Stufe*Nutz_h	1	4	2	5	-4.2418	5.1706	15.4	-0.82	0.4245	10.9941
19 N_Stufe*Nutz_h	1	5	2	2	-42.3686	5.1706	15.4	-8.19	<.0001	10.9941
20 N_Stufe*Nutz_h	1	5	2	3	-28.3881	5.1706	15.4	-5.49	<.0001	10.9941
21 N_Stufe*Nutz_h	1	5	2	4	-33.9745	5.1706	15.4	-6.57	<.0001	10.9941
22 N_Stufe*Nutz_h	1	5	2	5	-21.6891	4.2898	10.7	-5.06	0.0004	9.4770
23 N_Stufe*Nutz_h	2	2	2	3	13.9805	5.5207	16.7	2.53	0.0216	11.6609
24 N_Stufe*Nutz_h	2	2	2	4	8.3941	5.5207	16.7	1.52	0.1470	11.6609
25 N_Stufe*Nutz_h	2	2	2	5	20.6795	5.5207	16.7	3.75	0.0016	11.6609
26 N_Stufe*Nutz_h	2	3	2	4	-5.5863	5.5207	16.7	-1.01	0.3260	11.6609
27 N_Stufe*Nutz_h	2	3	2	5	6.6991	5.5207	16.7	1.21	0.2418	11.6609
28 N_Stufe*Nutz_h	2	4	2	5	12.2854	5.5207	16.7	2.23	0.0401	11.6609

Wir sehen z.B., dass alle Vergleiche, die dieselbe Nutzungshäufigkeit, aber verschiedene N-Mengen umfassen, den Standardfehler  $\sqrt{V} = 4,2898$  sowie die Satterthwaite-Freiheitsgrade  $FG_{Satterth} = 10,7$  haben. Multiplikation des Standardfehler mit den t-Wert für die entsprechenden Satterthwaite-Freiheitsgrade liefert die Grenzdifferenz, die wir hier in einem Datenschnitt berechnet haben. Für Vergleiche verschiedener Nutzungshäufigkeiten bei derselben N-Menge finden wir  $LSD = 9,477$ .

Das ganze fassen wir in einer Mittelwert-Tabelle zusammen.

**Tabelle 1.3:** Vergleich der Nutzungshäufigkeiten bei derselben N-Stufe.

Nutzungshäufigkeit	TM-Ertrag §	
	N-Menge	
	N <sub>1</sub>	N <sub>2</sub>
S <sub>2</sub>	92.8 c	99.5 c
S <sub>3</sub>	79.2 b	85.5 ab
S <sub>4</sub>	74.5 b	91.1 bc
S <sub>5</sub>	57.1 a	78.8 a

LSD = 9,48

§ Mittelwerte in derselben **Spalte**, die einen Buchstaben gemeinsam haben, sind nicht signifikant verschieden (t-Test,  $\alpha = 5\%$ )

**Tabelle 1.4:** Vergleich der N-Stufen bei derselben Nutzungshäufigkeit.

Nutzungshäufigkeit	TM-Ertrag	
	N-Menge	
	N <sub>1</sub>	N <sub>2</sub>
S <sub>2</sub>	92.8 a	99.5 a
S <sub>3</sub>	79.2 a	85.5 a
S <sub>4</sub>	74.5 a	91.1 b
S <sub>5</sub>	57.1 a	78.8 b

LSD = 11,66

§ Mittelwerte in derselben **Zeile**, die einen Buchstaben gemeinsam haben, sind nicht signifikant verschieden (t-Test,  $\alpha = 5\%$ )

Die Buchstabendarstellungen ist leider für Vergleiche von A×B-Mittelwerten mit MIXED nicht direkt zu erhalten. Es gibt jedoch ein SAS Makro (Piepho 2003a), welches in für diesen Zweck verwendet werden kann. Eine Buchstabendarstellung wird mit den folgenden Anweisungen erhalten:

```
ods output diffs=diffs;
ods output lsmeans=lsmeans;

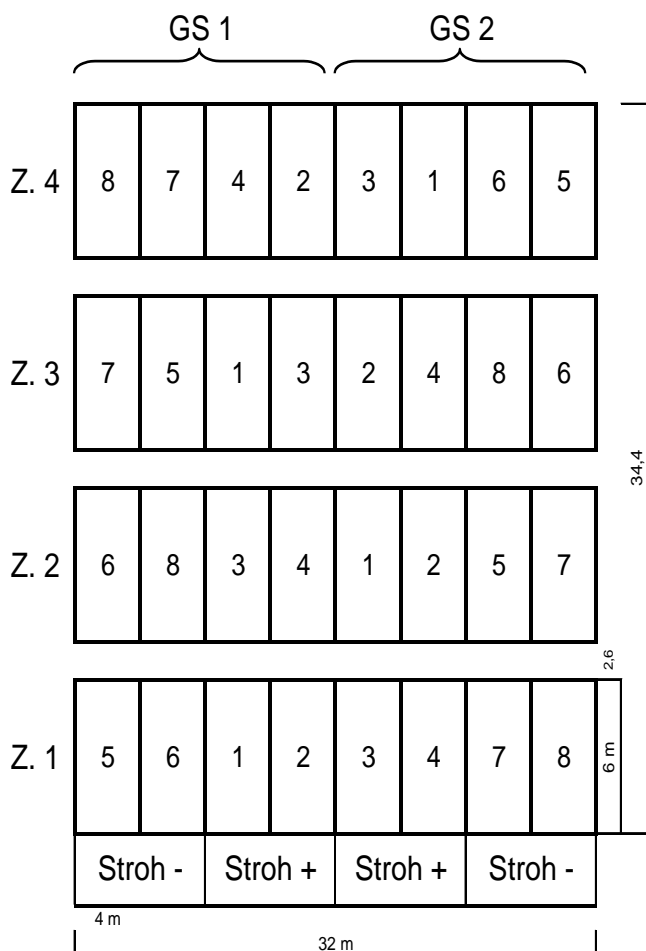
proc mixed;
class block nutz_h n_stufe;
model y=block nutz_h n_stufe nutz_h*n_stufe/ddfm=kr;
random nutz_h*block n_stufe*block;
lsmeans nutz_h*n_stufe/pdiff;
run;

%mult(trt=nutz_h, by=n_stufe, level=1);
%mult(trt=nutz_h, by=n_stufe, level=2);
%mult(trt=n_stufe, by=nutz_h, level=2);
%mult(trt=n_stufe, by=nutz_h, level=3);
%mult(trt=n_stufe, by=nutz_h, level=4);
%mult(trt=n_stufe, by=nutz_h, level=5);
```

Abschließende Bemerkung: Wir haben hier Nutzungshäufigkeiten getrennt für die N-Stufen verglichen und umgekehrt, um die Berechnung approximativer Freiheitsgrade nach Satterthwaite zu illustrieren. Wegen der nicht-signifikanten Wechselwirkungen ist es effizienter, die Mittelwerte der Nutzungshäufigkeiten über die N-Stufen und umgekehrt zu vergleichen.

## Beispiel 2: Eine ungewöhnliche Spaltanlage

Zur Untersuchung der Einflussmöglichkeiten auf die Festbrennstoffqualität von Energietriticale wurde ein zweifaktorieller Versuch mit den Faktoren Stroh und K-Düngung angelegt (Dr. D. Kauter, Institut für Pflanzenbau und Grünland, Hohenheim; Abb. 2.1). Aus technischen Gründen musste die Randomisation des Faktors Stroh stark eingeschränkt werden. Die Versuchsfläche wurde in 4 Zeilen (Z) und 8 Spalten unterteilt. Die 8 Spalten wurden zu 4 Grossteilstücken à 2 Spalten zusammengefasst,



auf welche der Faktor Stroh in einer Blockanlage randomisiert wurde (Block = Großspalte = GS). Je Zeile wurden die 4 K-Stufen getrennt für beiden Strohstufen randomisiert. Für die Auswertung dieser Spaltanlage ist ein adäquates gemischtes Modells erforderlich, welches wie in Piepho et al. (2003) beschriebenen hergeleitet werden können. Man beachte, dass für die betrachtete Anlage keine Formeln in Lehrbüchern existieren und eine vollständige Analyse mit GLM nicht möglich ist.

**Abb. 2.1:** Spaltanlage für Versuch mit Energietriticale. Varianten 1 bis 4 mit Strohdüngung: 1: KCl im Herbst, 2: KCl im Frühjahr; 3:  $K_2SO_4$  im Herbst; 4: Kontrolle. Varianten 5 bis 8: Wie 1 bis 4, aber ohne Strohdüngung.

Zunächst ist darauf hinzuweisen, dass die Versuchsanlage eine geringe Effizienz der Prüfung des Haupteffektes Stroh aufweist, was der Preis für die aus technischen Gründen eingeschränkte Randomisation ist.

Nun zur Formulierung eines für die Versuchsanlage adäquaten Modells. Die Methoden von Piepho et al. (2003) gehen davon aus, dass die Randomisationseinheiten getrennt von den Behandlungsfaktoren kodiert sind. Im vorliegenden Beispiel werden folgende Kodierungen verwendet:

Versuchseinheiten:

- GS = Großspalten (bestehend aus je 4 Spalten)
- S = Grossteilstücke (bestehend aus je 2 Spalten)
- Z = Zeilen (Blöcke)

PLOT = Parzellen-Nummer

Behandlungsfaktoren:

A = Strohvariante

B = K-Variante

Das Modell wird getrennt für Behandlungs- und Blockstruktur aufgestellt (Piepho et al., 2003a).

Modell für Behandlungsfaktoren (Syntax nach Patterson; siehe Piepho et al., 2003a; ähnlich der Modellsyntax bei Prozeduren für lineare Modelle):

$$A \times B = A + B + A \bullet B \quad [\text{entspricht } \alpha_i + \beta_j + (\alpha\beta)_{ij}]$$

Die Versuchsanlage umfasst zwei Randomisationsschritte, einen für die Grossteilstücke und einen für die Kleinteilstücke. Die beiden Randomisationsschritte sind nicht hierarchisch sondern überlagern sich, so dass wir getrennt für beide ein Teilmodell entwickeln und anschließend beide Teilmodelle zusammenfassen.

Modell für Grossteilstücke: Die Großspalten (GS) stellen Blöcke dar. Der Faktor Stroh (A) wird auf Grossteilstücke (S) innerhalb der Großspalten (GS) randomisiert, so dass eine hierarchische Struktur "Grossteilstücke innerhalb Großspalten" vorliegt. Das entsprechende Modell ist

$$GS/S = GS + GS \bullet S$$

Modell für Kleinteilstücke: Kombinationen aus Zeile (Z) und Faktor A stellen Blöcke dar, innerhalb derer der Faktor K-Düngung auf die Parzellen (PLOT) randomisiert wird. Das entsprechende Modell ist

$$(Z \bullet A)/PLOT = Z \bullet A + Z \bullet A \bullet PLOT$$

Es ergibt sich folgendes Gesamt-Modell:

$$A + B + A \bullet B + GS + Z \bullet A : GS \bullet S + Z \bullet A \bullet PLOT.$$

Hierbei werden feste Effekte vor zufälligen Effekten aufgeführt (nach dem Doppelpunkt). Spezifizierung dieses Modells in MIXED in Kombination mit der Kenward-Roger Methode liefert automatisch die richtigen F-Tests und Mittelwertvergleiche:

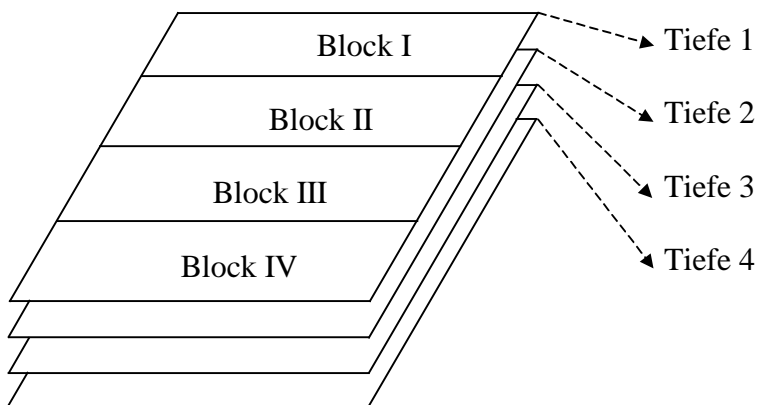
```
proc mixed;
class a b gs s z;
model y= a b a*b gs z*a/ddfm=kr;
random gs*s;
run;
```

Das Beispiel zeigt, wie auch für Nicht-Standard-Anlagen ein adäquates gemischtes Modell aufgestellt werden kann, sofern die Randomisationsstruktur eindeutig beschrieben werden kann.

### Beispiel 3: Feldversuch mit Messwiederholungen im Raum

(F. Wald, Institut für Pflanzenbau und Grünland, Hohenheim): Leguminosen-Grasgemenge sind als N-Quelle in Fruchtfolgen des Organischen Landbaus von zentraler Bedeutung. Nach meist mehrjährigem Anbau wird der von den Leguminosen assimilierte Stickstoff mit dem Umbruch der Narbe durch die Mineralisation der organischen Substanz wieder freigesetzt. In der Praxis steht für eine Steuerung der Mineralisierung ausschließlich das Instrument der Bodenbearbeitung zur Verfügung. Ziel dieser Untersuchungen war es, die Beeinflussbarkeit der N-Mineralisierung durch eine Variation des Narbenumbruchs zu quantifizieren, um so eine Datengrundlage für die Modellierung der N-Dynamik in Abhängigkeit von der Bodenbearbeitung zu schaffen. Nach dem Umbruch eines 3-jährigen Klee-Luzernegrasgemenges wurde im Herbst die Weizensorte Rektor gesät. Die Bodenbearbeitung zum Umbruch der Narbe wurde in 3 Stufen in einer Blockanlage variiert. Die Stickstoffgehalte des Bodens wurden durch wiederholtes Beprobieren in den Tiefen 0-10, 10-20, 20-30 und 30-60 cm ermittelt. Die Messungen erfolgten zu 3 verschiedenen Zeitpunkten.

Bei der Auswertung ist zu berücksichtigen, dass wiederholte Messungen in derselben Parzelle zu verschiedenen Zeitpunkten korreliert sind. Ebenso sind Messungen in verschiedenen Tiefen auf derselben Parzelle korreliert. Diese Korrelationen müssen bei der Auswertung berücksichtigt werden, um valide statistische Tests für Hauptwirkungen und Interaktionen zu erhalten. Dies gilt insbesondere für den Vergleich von Stufen eines "wiederholten" Faktors (Tiefe, Zeit). Ignoriert man dagegen die Korrelationen, so sind die resultierenden statistischen Auswertungen inadäquat bzw. irreführend. Bei Messwiederholungen ist in der Regel davon auszugehen, dass die Korrelation mit dem räumlichen bzw. zeitlichen Abstand sinkt. Es gibt eine Reihe entsprechender Korrelationsstrukturen, die im Rahmen eines gemischten Modells leicht geschätzt werden können. Man passt zunächst verschiedene Modelle für die Korrelationsstruktur an und selektiert dann das am besten an die Daten passende Korrelationsmodell mittels eines statistischen Kriteriums (Likelihood-Quotienten-Test, Akaike Information Criterion - AIC). Anschließend werden die Behandlungseffekte des Modells mittels Wald-F-Tests und t-Tests geprüft.



**Abb. 3.1:** Vier Tiefenschichten (1, 2, 3, 4) im Bodenkörper einer Blockanlage.



**Tab. 3.1:** Bodenbearbeitungsmaßnahmen:

Nummer	Beschreibung
1	zweifache Rotortillerbearbeitung (RT, 10 cm tief) in etwa zweiwöchigem Abstand abschließend Pflugfurche (Pflug, 25 cm tief)
2	RT einmal und abschließend Pflugeinsatz
3	ausschließlicher Pflugumbruch

Wir betrachten hier nur den dritten Termin. Wir können uns den Bodenkörper unter der Versuchsanlage in vier Schichten geteilt vorstellen. In jeder der vier Schichten liegt eine Blockanlage vor. Das Modell einer Blockanlage lautet

$$y_{ij} = \mu + b_j + \tau_i + e_{ij}$$

wobei

$y_{ij}$  = N<sub>min</sub>-Gehalt der  $i$ -ten Behandlung (Bodenbearbeitung) im  $j$ -ten Block

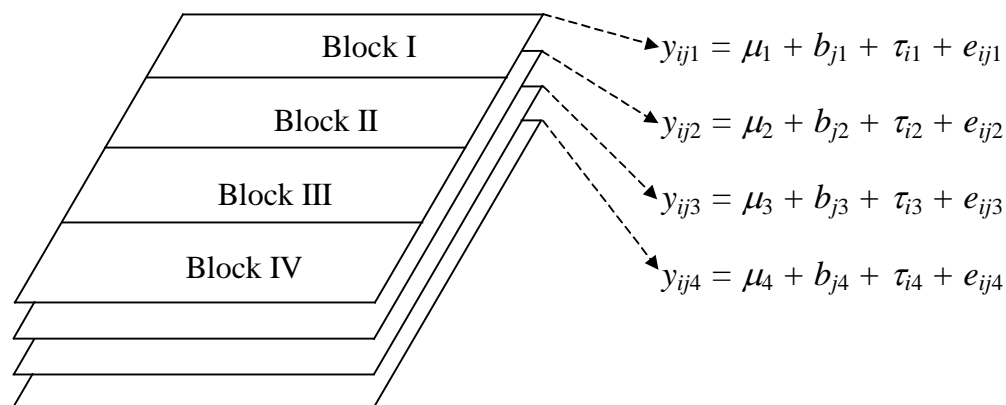
$\mu$  = Gesamteffekt

$b_j$  = Effekt des  $j$ -ten Block

$\tau_i$  = Effekt der  $i$ -ten Behandlung

$e_{ij}$  = Fehler der  $ij$ -ten Parzelle

Nun soll der Faktor Tiefe in das Modell mit aufgenommen werden, um eine zweifaktorielle Auswertung Bodenbearbeitung  $\times$  Tiefe durchführen zu können. Hierzu können wir uns zunächst vorstellen, dass in jeder Tiefe der Blockanlage ein Blockmodell gelten muss. In Abb. 3.2 ist dies veranschaulicht, wobei ein zusätzlicher Index für die Tiefe eingeführt wurde.



**Abb. 3.2:** Vier Tiefenschichten (1, 2, 3, 4) im Bodenkörper einer Blockanlage mit dazugehörigem linearem Modell.

Zusammenfassend, einschließlich des Index für die Tiefe, lautet das zweifaktorielle Modell

$$y_{ijt} = \mu_t + b_{jt} + \tau_{it} + e_{ijt}$$

wobei  $t = 1, 2, 3, 4$  ein Index für die Tiefe ist. Der Behandlungseffekt trägt nun eine doppelte Indizierung für die beiden Faktoren Bodenbearbeitungsmaßnahme und Tiefe. Wir können diesen Effekt auch wie üblich explizit zweifaktoriell aufspalten:

$$\tau_{it} = \alpha_i + \beta_t + (\alpha\beta)_{it}$$

wobei

$\alpha_i$  = Haupteffekt der  $i$ -ten Bodenbearbeitungsmaßnahme

$\beta_t$  = Haupteffekt der  $t$ -ten Tiefe

$(\alpha\beta)_{it}$  =  $it$ -te Wechselwirkung

Wichtig bei der Auswertung ist nun zunächst die Betrachtung des Fehlerterms  $e_{ijt}$ , und hier insbesondere der Fehler in derselben Parzelle bei verschiedenen Tiefen. Die einfachste Annahme wäre, dass die Fehler in einer Parzelle unabhängig sind, also unkorreliert sind. Dann gilt für zwei Tiefen  $s$  und  $t$ :

$$\text{corr}(e_{ijs}, e_{ijt}) = 0 \quad (\text{Unabhängigkeit})$$

Diese Annahme ist allerdings sehr unrealistisch, da die Tiefenschichten unmittelbar benachbart sind und eine Randomisation nicht möglich ist. Man kann daher erwarten, dass die Messwerte aus verschiedenen Tiefen positiv korreliert sind. Eine entsprechende einfache Erweiterung des Modells besteht in der Annahme einer konstanten Korrelation für jedes Paar von Schichten:

$$\text{corr}(e_{ijs}, e_{ijt}) = \rho \quad (\text{Compound Symmetry})$$

Dieses Modell, welches auch als "compound symmetry" (CS) bezeichnet wird, ist sehr eng verwandt mit dem einer Spaltanlage. Wir können nämlich das CS Modell auch erzeugen, indem wir für die Fehler  $e_{ijt}$  folgendes Modell postulieren:

$$e_{ijt} = f_{ij} + g_{ijt}$$

wobei

$f_{ij}$  =  $ij$ -ter Parzellenhaupteffect; normalverteilt mit Varianz  $\sigma_f^2$

$g_{ijt}$  = Tiefenspezifische Abweichung in der  $ij$ -ten Parzelle; normalverteilt mit Varianz

$$\sigma_g^2$$

Man kann zeigen, dass bei diesem Modell gilt

$$\text{corr}(e_{ijs}, e_{ijt}) = \rho = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_g^2} \quad ,$$

was im wesentlichen dem CS-Modell entspricht. Das Modell entspricht formal dem einer Spaltanlage, wenn wir  $f_{ij}$  als Grossteilstückfehler und  $g_{ijt}$  als Kleinteilstückfehler auffassen. In anderen Worten, die Bodensäule unter einer Parzelle ist das Grossteilstück, während die 4 Tiefenschichten die Kleinteilstücke darstellen. In der Tat werden Versuche mit Messwiederholungen gerne wie eine Spaltanlage ausgewertet. Allerdings ist dies aus einem wichtigen Grund problematisch: Man kann die verschiedenen Tiefen innerhalb der Bodensäule nicht randomisieren. Daher kann die Auswertung als Spaltsanlage nicht über die Randomisationstheorie motiviert werden wie in echten Spaltanlagen, bei denen Grossteilstücke und Kleinteilstücke randomisiert werden.

Wegen der fehlenden Randomisation ist davon auszugehen, dass die Höhe der Korrelation nicht für alle Paare von Schichten dieselbe ist. Insbesondere findet man oft, dass unmittelbar benachbarte Schichten höher korreliert sind als weiter entfernte. Ein sehr einfaches Modell für eine solche Situation ist das folgende:

$$\text{corr}(e_{ijs}, e_{ijt}) = \rho^{|s-t|} \quad [\text{AR}(1)]$$

Dies Modell wird als autoregressives Modell bezeichnet. Die Korrelation nimmt mit der Höhe des Abstandes  $|s-t|$  ab, wobei die Abnahme umso stärker ausfällt, je kleiner die Autokorrelation  $\rho$  ist (Abb. 3.3).

Bisher haben wir nur die Korrelation betrachtet, nicht aber die Varianz. Bei allen drei Korrelationsmodellen sind zwei Annahmen möglich: entweder die Varianzen sind bei allen Tiefen gleich oder aber die Varianz unterscheidet sich zwischen den Schichten. Schließlich kann man fordern, dass sowohl die Varianzen als auch die Korrelationen völlig frei wählbar sind. In diesem Fall spricht man von einem unstrukturierten Modell.

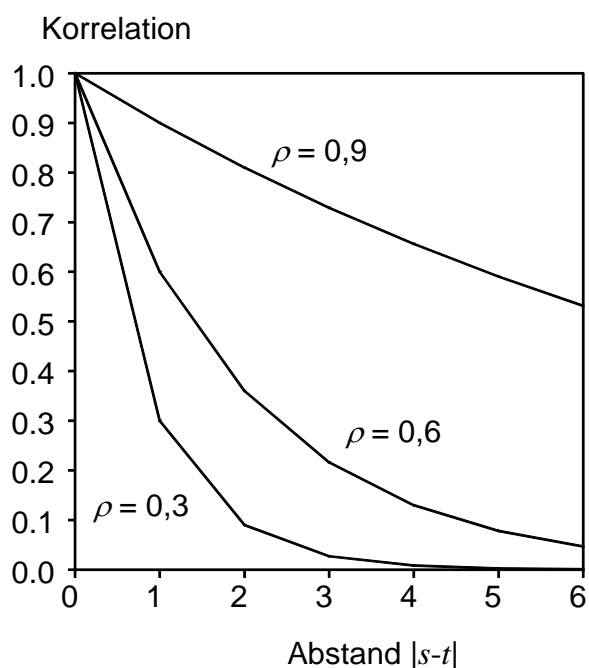
Nun haben wir verschiedene Korrelationsstrukturen zur Auswahl. Welches sollen wir nehmen? Eine sehr einfache Strategie besteht darin, ein Modellselektionskriterium wie das AIC (Akaike Information Criterion) zu verwenden. Für gut passende Modelle nimmt dieses einen kleinen Wert an. Es berechnet sich aus der sog. Log-likelihood Funktion ( $\log L$ ), welche ein Maß dafür ist, wie wahrscheinlich die beobachteten Daten unter der Annahme des jeweils betrachteten Modells sind. *AIC* ist wie folgt definiert:

$$AIC = -2 \log L + 2p$$

wobei  $p$  die Zahl der Korrelationsparameter ist. Der zweite Term auf der rechten Seite ist ein Strafterm, der komplexe Modelle "bestraft". Es stellt sicher, dass man nicht mehr Parameter schätzt als zur Beschreibung der Daten notwendig. Die *AIC*-Werte für verschiedene Modelle sind in Tab. 3.1 angegeben. Offenbar passt das AR(1)-Modell mit heterogenen Varianzen am besten an die Daten.

**Tab. 3.1:** Anpassung verschiedener Varianz-Kovarianz-Modelle.  $N_{min}$ -Werte logarithmisch transformiert zur Erzielung einer Normalverteilung.

Modell	$-2 \log L$	AIC (kleiner ist besser)
Homogene Varianzen:		
Unabhängig	21.9	23.9
CS	15.7	19.7
AR(1)	13.9	17.9
Heterogene Varianzen:		
Unabhängig	15.1	23.1
CS	6.7	16.7
AR(1)	6.4	16.4
UN	-2.5	17.5



**Abb. 3.3:** Abhängigkeit der Korrelation vom Abstand  $|s-t|$  beim AR(1)-Modell.

Im weiteren werden nun die festen Effekte mit dem selektierten Modell für die Korrelationsstruktur geprüft. Die Anweisungen für MIXED unter Verwendung der Kenward-Roger Methode sind wie folgt:

```
proc mixed;
class tiefe block boden;
model log_nmin=         tiefe
                    block block*tiefe
                    boden boden*tiefe/ddfm=kr;
repeated tiefe/sub=block*boden type=arh(1);
run;
```

Für die Wald-F-Tests finden wir:

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
tiefe	3	9.36	271.44	<.0001
block	3	6.73	3.08	0.1024
tiefe*block	9	9.58	0.33	0.9429
boden	2	6.73	78.96	<.0001
tiefe*boden	6	9.91	20.98	<.0001

Die Wechselwirkungen sind signifikant, so dass wir Bodenbearbeitungen getrennt für jede Tiefe und Tiefen getrennt für jede Bodenbearbeitungsmaßnahme vergleichen.

Nebenbemerkung: Wegen der tiefenspezifischen Blockeffekte  $b_{jt}$  können die Hauptwirkungen für die Tiefe nicht von den Blockeffekten getrennt geprüft werden. Daher erlaubt die Auswertung keine Aussage über Haupteffekte für den wiederholten Faktor Tiefe. Dies ist im vorliegenden Fall unproblematisch, da die Wechselwirkungen signifikant sind, so dass Hauptwirkungen irrelevant sind.

Die Mittelwerte werden mit der Methode der gewichteten kleinsten Quadrate geschätzt, die in MIXED implementiert ist. Hierbei werden die Varianzkomponenten zur Berechnung optimaler Gewichte verwendet. Wegen der Art der Korrelationsstruktur sind die Standardfehler einer Differenz nicht konstant und es kann keine gemeinsame Grenzdifferenz berechnet werden. Somit ist auch die Buchstabendarstellung mit Standardverfahren nicht verfügbar. Stattdessen verwenden wir das Makro von Piepho (2003a). Die SAS-Anweisungen für Mittelwertvergleiche sind wie folgt:

```
ods output diffs=diffs;
ods output lsmeans=lsmeans;

proc mixed;
class tiefe block boden;
model log_nmin=          tiefe
                block block*tiefe
                boden boden*tiefe/ddfm=kr;
repeated tiefe/sub=block*boden type=arh(1);
lsmeans boden*tiefe/pdiff;
%mult(trt=boden, by=tiefe, level=1);
%mult(trt=boden, by=tiefe, level=2);
%mult(trt=boden, by=tiefe, level=3);
%mult(trt=boden, by=tiefe, level=4);
%mult(trt=tiefe, by=boden, level=1);
%mult(trt=tiefe, by=boden, level=2);
%mult(trt=tiefe, by=boden, level=3);
run;
```

Die Ergebnisse sind in Tab. 3.2 dargestellt.

**Tab. 3.2:** Mittelwertvergleiche für  $N_{min}$ -Versuch.  $N_{min}$ -Werte logarithmisch transformiert.

Bodenbearbeitung	Tiefe			
	1	2	3	4
2*Rotortiller+Pflug	3,70 <sup>bC</sup>	2,95 <sup>cB</sup>	1,99 <sup>cA</sup>	1,99 <sup>bA</sup>

1*Rotortiller+Pflug	3,63 <sup>b C</sup>	2,26 <sup>b B</sup>	1,42 <sup>b A</sup>	1,52 <sup>b A</sup>
Pflug	1,72 <sup>a B</sup>	1,49 <sup>a B</sup>	0,92 <sup>a A</sup>	0,80 <sup>a A</sup>

Mittelwerte in einer Zeile, die einen großen Buchstaben gemeinsam haben, sind nicht signifikant voneinander verschieden (t-Test,  $\alpha = 5\%$ ).

Mittelwerte in einer Spalte, die einen kleinen Buchstaben gemeinsam haben, sind nicht signifikant voneinander verschieden (t-Test,  $\alpha = 5\%$ ).

Zur besseren Interpretierbarkeit werden die Mittelwerte von der logarithmischen Skala auf die ursprüngliche Skala zurücktransformiert (Tab 3.3). Hierbei ist lediglich zu beachten, dass die zurücktransformierten Mittelwerte Schätzwerte für den Median, nicht aber das arithmetische Mittel (Erwartungswert) sind (siehe Anhang). Die Buchstabendarstellung in Tab. 3.3 ist dieselbe die auch in Tab. 3.2 angegeben ist.

**Tab. 3.3:** Median-Vergleiche für  $N_{min}$ -Versuch. Rücktransformierte  $N_{min}$ -Mittelwerte.

Bodenbearbeitung	Tiefe			
	1	2	3	4
2*Rotortiller+Pflug	40,4 <sup>b C</sup>	19,1 <sup>c B</sup>	7,3 <sup>c A</sup>	7,3 <sup>b A</sup>
1*Rotortiller+Pflug	37,7 <sup>b C</sup>	9,6 <sup>b B</sup>	4,1 <sup>b A</sup>	4,6 <sup>b A</sup>
Pflug	5,6 <sup>a B</sup>	4,4 <sup>a B</sup>	2,5 <sup>a A</sup>	2,2 <sup>a A</sup>

Mediane in einer Zeile, die einen großen Buchstaben gemeinsam haben, sind nicht signifikant voneinander verschieden (t-Test,  $\alpha = 5\%$ ).

Mediane in einer Spalte, die einen kleinen Buchstaben gemeinsam haben, sind nicht signifikant voneinander verschieden (t-Test,  $\alpha = 5\%$ ).

Wir stellen fest, dass der  $N_{min}$ -Gehalt bei allen Bodenbearbeitungsvarianten mit der Tiefe abnimmt. Dabei ist der  $N_{min}$ -Gehalt in der Pflug-Variante jeweils signifikant am geringsten.

Abschließende Bemerkung: Wir haben hier nur einen Termin betrachtet, und verschiedene Tiefen verglichen. Alternativ können wir die Tiefe konstant halten und die Termine vergleichen. Hierbei haben wir im wesentlichen dieselbe Fehlerstruktur, nur dass räumliche durch zeitliche Korrelation ersetzt wird. Insbesondere ist zu erwarten, dass zeitlich benachbarte Termine sich ähnlicher sind als weiter entfernte. Schließlich können wir die beiden wiederholten Faktoren Termin und Tiefe simultan betrachten, was eine komplexe Korrelationsstruktur erfordert.

SAS-Anweisungen:

```
data;
input
termin block tiefe boden Nmin;
if termin=3;
log_nmin=log(nmin);
datalines;
1 1 1 1 8.48
1 1 2 1 5.74
1 1 3 1 6.49
1 1 4 1 7.18
1 1 1 2 20.61
```

1	1	2	2	5.16
1	1	3	2	4.35
1	1	4	2	4.41
1	1	1	3	9.10
1	1	2	3	5.86
1	1	3	3	3.45
1	1	4	3	4.65
1	2	1	2	4.95
1	2	2	2	3.23
1	2	3	2	2.91
1	2	4	2	2.14
1	2	1	1	7.35
1	2	2	1	3.81
1	2	3	1	2.88
1	2	4	1	2.37
1	2	1	3	10.56
1	2	2	3	4.55
1	2	3	3	2.94
1	2	4	3	4.28
1	3	1	1	5.41
1	3	2	1	3.23
1	3	3	1	2.63
1	3	4	1	2.93
1	3	1	3	13.66
1	3	2	3	7.90
1	3	3	3	4.88
1	3	4	3	9.15
1	3	1	2	7.13
1	3	2	2	5.14
1	3	3	2	2.76
1	3	4	2	2.35
1	4	1	2	6.85
1	4	2	2	4.43
1	4	3	2	2.87
1	4	4	2	1.94
1	4	1	3	6.78
1	4	2	3	3.83
1	4	3	3	2.98
1	4	4	3	2.51
1	4	1	1	8.26
1	4	2	1	3.39
1	4	3	1	2.91
1	4	4	1	4.05
2	1	1	1	29.57
2	1	2	1	1.18
2	1	3	1	5.65
2	1	4	1	8.74
2	1	1	2	9.68
2	1	2	2	4.14
2	1	3	2	3.07
2	1	4	2	4.31
2	1	1	3	4.45
2	1	2	3	3.27
2	1	3	3	2.04
2	1	4	3	2.91
2	2	1	2	5.75
2	2	2	2	4.22
2	2	3	2	2.29
2	2	4	2	4.38
2	2	1	1	36.47
2	2	2	1	8.84

2	2	3	1	3.96
2	2	4	1	6.02
2	2	1	3	7.71
2	2	2	3	6.26
2	2	3	3	3.94
2	2	4	3	4.03
2	3	1	1	23.82
2	3	2	1	6.53
2	3	3	1	5.89
2	3	4	1	14.93
2	3	1	3	4.95
2	3	2	3	4.64
2	3	3	3	4.17
2	3	4	3	4.29
2	3	1	2	6.04
2	3	2	2	3.04
2	3	3	2	2.10
2	3	4	2	2.31
2	4	1	2	10.61
2	4	2	2	2.87
2	4	3	2	2.26
2	4	4	2	2.85
2	4	1	3	4.61
2	4	2	3	2.46
2	4	3	3	2.02
2	4	4	3	2.41
2	4	1	1	37.06
2	4	2	1	9.26
2	4	3	1	5.30
2	4	4	1	7.09
3	1	1	1	39.36
3	1	2	1	17.49
3	1	3	1	6.89
3	1	4	1	6.54
3	1	1	2	44.82
3	1	2	2	11.75
3	1	3	2	5.54
3	1	4	2	9.49
3	1	1	3	6.01
3	1	2	3	5.08
3	1	3	3	2.28
3	1	4	3	2.04
3	2	1	2	43.72
3	2	2	2	12.63
3	2	3	2	4.62
3	2	4	2	4.33
3	2	1	1	45.92
3	2	2	1	26.49
3	2	3	1	8.51
3	2	4	1	8.65
3	2	1	3	7.35
3	2	2	3	4.92
3	2	3	3	2.97
3	2	4	3	2.71
3	3	1	1	34.04
3	3	2	1	13.80
3	3	3	1	5.90
3	3	4	1	6.00
3	3	1	3	4.43
3	3	2	3	5.06
3	3	3	3	2.67



```

3      3      4      3      1.86
3      3      1      2      31.28
3      3      2      2      6.43
3      3      3      2      2.81
3      3      4      2      2.93
3      4      1      2      32.87
3      4      2      2      8.96
3      4      3      2      4.05
3      4      4      2      3.69
3      4      1      3      4.90
3      4      2      3      3.06
3      4      3      3      2.23
3      4      4      3      2.36
3      4      1      1      43.66
3      4      2      1      21.06
3      4      3      1      8.53
3      4      4      1      8.69
;
proc print;
run;

proc sort;
by block boden tiefe;

ods output diffs=diffs;
ods output lsmeans=lsmeans;

proc mixed;
class tiefe block boden;
model log_nmin=          tiefe
                block block*tiefe
                boden boden*tiefe/ddfm=kr;
repeated tiefe/sub=block*boden type=arh(1);
lsmeans boden*tiefe/pdiff;
%mult(trt=boden, by=tiefe, level=1);
%mult(trt=boden, by=tiefe, level=2);
%mult(trt=boden, by=tiefe, level=3);
%mult(trt=boden, by=tiefe, level=4);
%mult(trt=tiefe, by=boden, level=1);
%mult(trt=tiefe, by=boden, level=2);
%mult(trt=tiefe, by=boden, level=3);
run;

```

## Beispiel 4: Streifen-Spaltanlage, Messwiederholungen in der Zeit

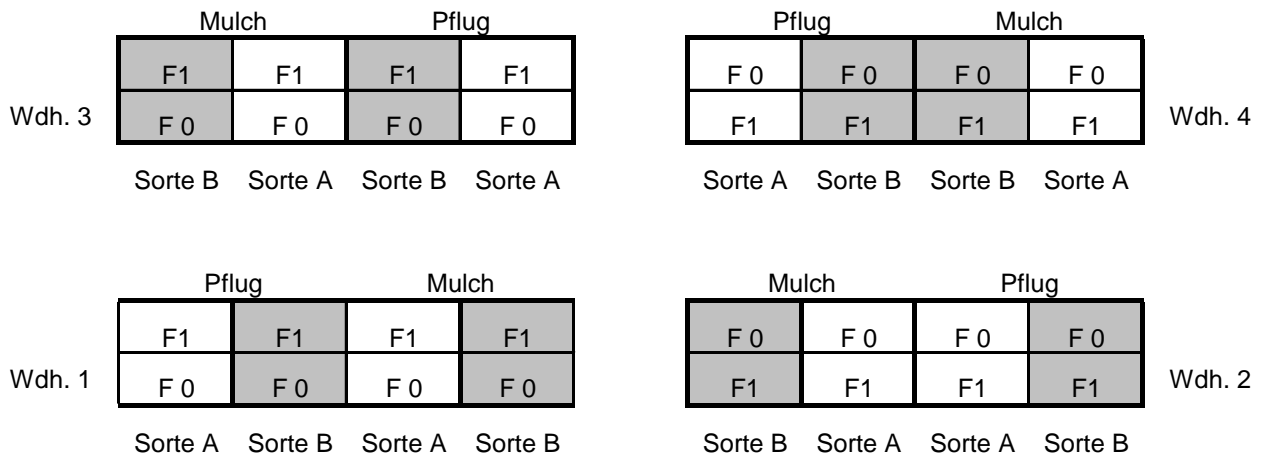
### Beschreibung

Der Befall mit *Fusarium*-Pilzen ist nach wie vor ein Problem im Weizenanbau. Im Rahmen einer mehrjährigen Versuchsserie wurde die Wirkung von Bodenbearbeitung, Sorte und Fungizidbehandlung auf den Fusarienbefall von Winterweizen gemessen (Daten von C. Pringas, Institut für Zuckerrübenforschung Göttingen). Zielvariable ist die Belastung des Erntegutes mit *Deoxynivalenol* [mg/kg].

Der Versuch war in vier vollständigen Blöcken angelegt. Jeder Block war in zwei Spalten (Pflug, Mulch) geteilt. Die Spalten waren jeweils in zwei Unterspalten (Sorte A, Sorte B) geteilt. In rechtem Winkel dazu war jeder Block in zwei Zeilen (mit Fungizid, ohne Fungizid) geteilt. Das Versuchsdesign entspricht damit dem einer dreifaktori-

riellen Streifen-Spaltanlage in vier Wiederholungen. Der Versuch wurde zweijährig auf der gleichen Fläche mit ortsfesten Behandlungen durchgeführt.

Die Daten sind im Anhang wiedergegeben.



**Abb. 4.1:** Versuchsplan

### Statistisches Modell für die einjährige Auswertung

Zunächst sollen die beiden Versuchsjahre separat ausgewertet werden. Für jede Randomisierungseinheit ist ein Fehlerterm ins Modell aufzunehmen. Neben dem Restfehler (Kleinteilstückfehler) treten deshalb zwei Großteilstückfehler (Fungizid\*Wdh. und B\*Wdh.) und zwei Mittelteilstückfehler auf (B\*Fungizid\*Wdh., B\*Sorte\*Wdh.) Fixe Faktoren sind Bodenbearbeitung, Sorte und Fungizidbehandlung.

#### Das Modell für die Behandlungseffekte:

Mit Bodenbearbeitung = B, Fungizid = F, Sorte = S

$$F \times B \times S = F + B + S + F \cdot B + F \cdot S + B \cdot S + F \cdot B \cdot S$$

[entspricht  $[\alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk}]$ ]

#### Das Modell für die Designeffekte:

Zeile (Ze)

Spalte (Sp)

Unterspalte (U)

Block bzw. Wiederholung (R)

$$= R + Ze \cdot R + Sp \cdot R + Ze \cdot Sp \cdot R + Sp \cdot U \cdot R + PLOT$$

[entspricht  $r_l + r\alpha_{il} + r\beta_{jl} + r\alpha\beta_{ijl} + r\beta\gamma_{jkl} + e_{ijkl}$ ]

Zeile, Spalte und Unterspalte lassen sich ebenso gut durch Stufen der korrespondierenden Behandlungsfaktoren darstellen. Nach PIEPHO et al. (2003a) folgt

$$Wdh / ((Bodenbearbeitung/Sorte) \times Fungizid) = Wdh / (B + B \cdot S + F + B \cdot F + B \cdot S \cdot F)$$

$$= Wdh + Wdh \cdot B + Wdh \cdot B \cdot S + Wdh \cdot F + Wdh \cdot B \cdot F + Wdh \cdot B \cdot S \cdot F$$

### Komplettes Modell:

Fixe Effekte: BB + F + S + BB•F + BB•S + F•S + BB•F•S + Wdh

Zufällige Effekte: Z•Wdh + Sp•Wdh + Ze•Sp•Wdh + Sp•U•Wdh + Plot

Oder mit anderer Schreibweise

$$y_{ijkl} = \alpha_i + \beta_j + \gamma_k + \alpha\beta_{ij} + \alpha\gamma_{ik} + \beta\gamma_{jk} + \alpha\beta\gamma_{ijk} + r_l + r\alpha_{il} + r\beta_{jl} + r\alpha\beta_{ijl} + r\beta\gamma_{jkl} + e_{ijkl}$$

wobei:

$r_l$  Effekt der l-ten Wiederholung

$\alpha_i$  Effekt der i-ten Fungizidstufe

$\beta_j$  Effekt der j-ten Bodenbearbeitung

$\gamma_k$  Effekt der k-ten Sorte

$\alpha\beta\gamma_{ijk}$  Interaktionen zwischen Fungizid, Bodenbearbeitung und Sorte

$r\alpha_{il}$  Fehler des il-ten Großteilstücks  $\sim N(0, \sigma^2_{ra})$

$r\beta_{jl}$  Fehler des jl-ten Großteilstücks  $\sim N(0, \sigma^2_{rb})$

$r\alpha\beta_{ijl}$  Fehler des ijl-ten Mittelteilstücks (Kombination Zeile/Spalte)  $\sim N(0, \sigma^2_{rab})$

$r\beta\gamma_{jkl}$  Fehler des ikl-ten Mittelteilstücks (Komb. Spalte/U.spalte)  $\sim N(0, \sigma^2_{rbc})$

$e_{ijkl}$  Fehler der ijk-l-ten Parzelle  $\sim N(0, \sigma^2_e)$

### Umsetzung in SAS-Code:

In Mixed werden im Gegensatz zu der Prozedur GLM zufällige Effekte nicht in die „Model“-Zeile geschrieben, sondern tauchen nur in der Zeile „random“ auf. Für die Approximation der Freiheitsgrade sollte die Methode nach Kenward-Roger (DDFM=KR) gewählt werden, eine Weiterentwicklung der Satterthwaite-Methode.

Kasten 4.1: Auswertung nach Modell (1) über SAS, Prozedur Mixed

```
Proc mixed data=fusarium; by jahr;
class BB Sorte Fungi WDH BD;
model DON = Wdh BB|Sorte|Fungi/ddfm=kr outpred=pred;
random BB*Wdh Fungi*Wdh BB*Fungi*Wdh BB*Sorte*Wdh;
run;
```

Für das Merkmal *Deoxynivalenol*-Gehalt besteht der Verdacht, dass wichtige Voraussetzungen der Varianzanalyse wie Normalverteilung der Residuen und Homogenität der Varianzen nicht eingehalten werden. Wir können die Varianzhomogenität mit einem Plot der Residuen gegen die geschätzten Erwartungswerte und die Normalverteilungshypothese mit einem sogenannten QQ-Plot überprüfen (Plot von „studentisierten“ Residuen gegen die „Normal Scores“, die geordneten Erwartungswerte aus einer Standardnormalverteilung). Die Anweisung „outpred=pred“ sorgt dafür, dass die Residuen und die Erwartungswerte in eine Datei mit dem Namen „pred“ geschrieben werden.

Ein QQ-Plot der Residuen zeigt für das Jahr 2001 Abweichungen von der Normalverteilung. Bei Normalverteilung sollten die Punkte im QQ-Plot auf der Diagonalen liegen (Abb. 4.2). Wir versuchen Normalverteilung durch eine Transformation zu erreichen. Hier bieten sich zunächst einmal die Wurzel und die Log-Transformation an.

$y' = \sqrt{y}$

$y' = \log(y)$

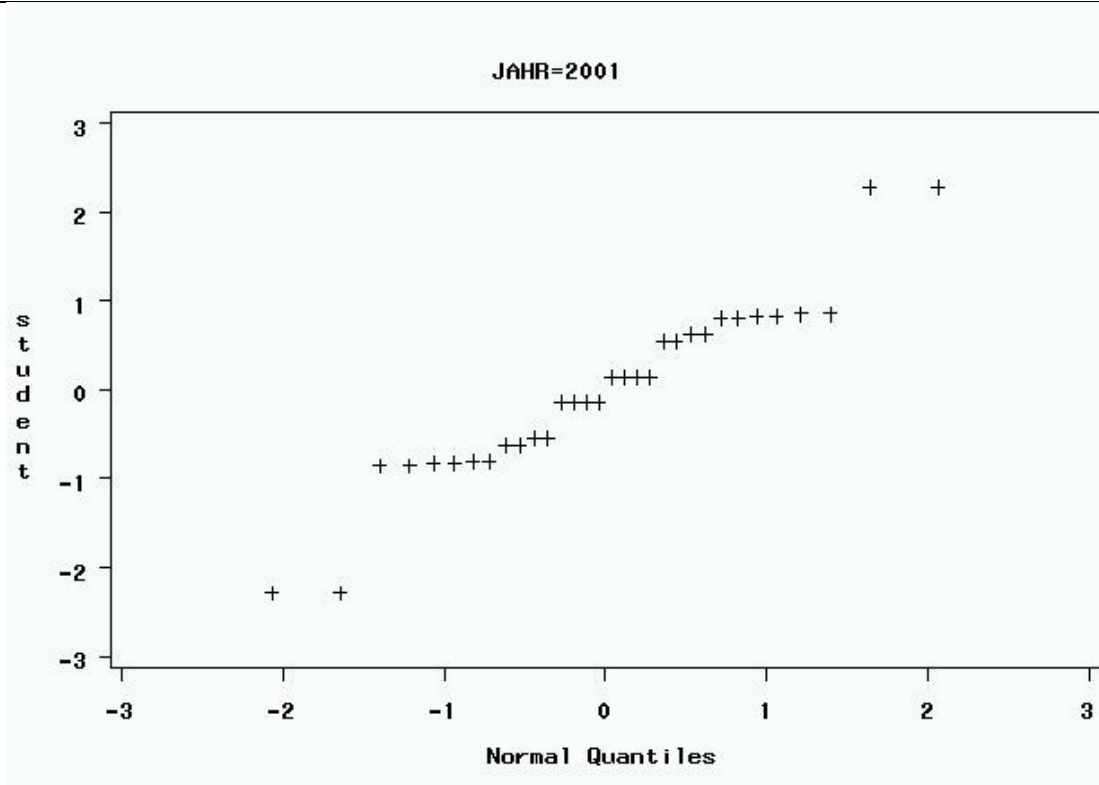


Abb. 4.2: Q-Q-Plot der studentisierten Residuen gegen die Quantile der Normalverteilung vor Transformation

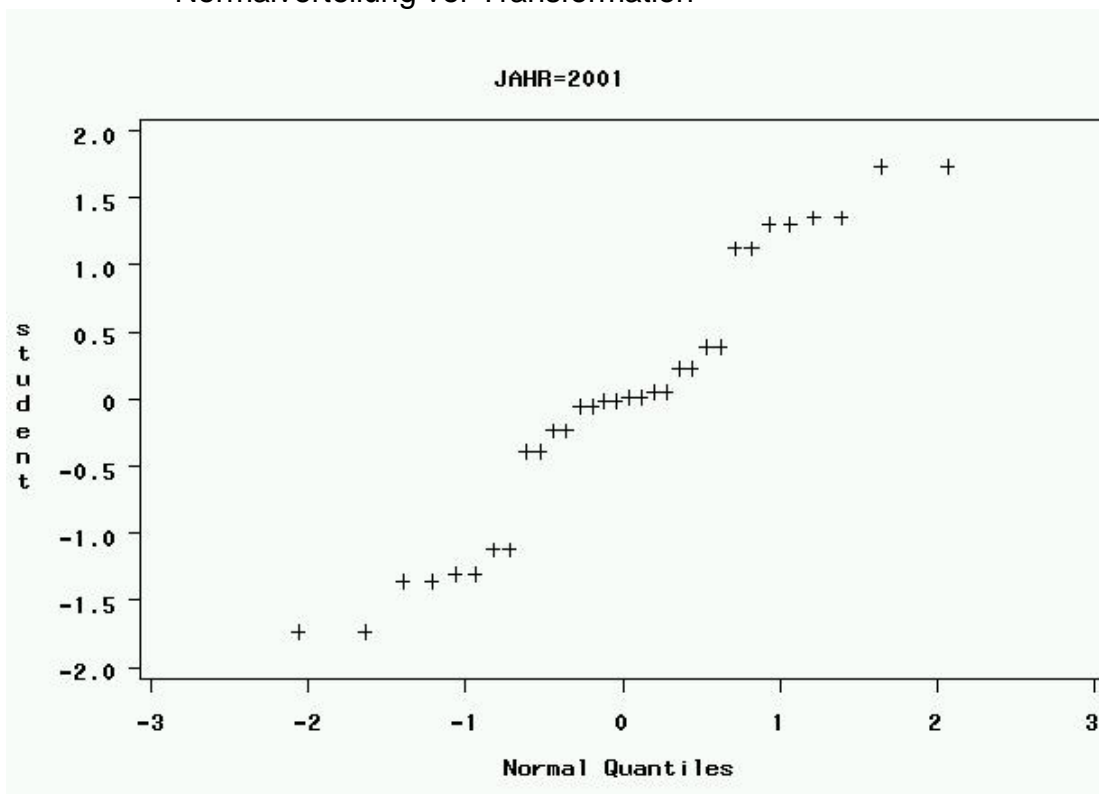
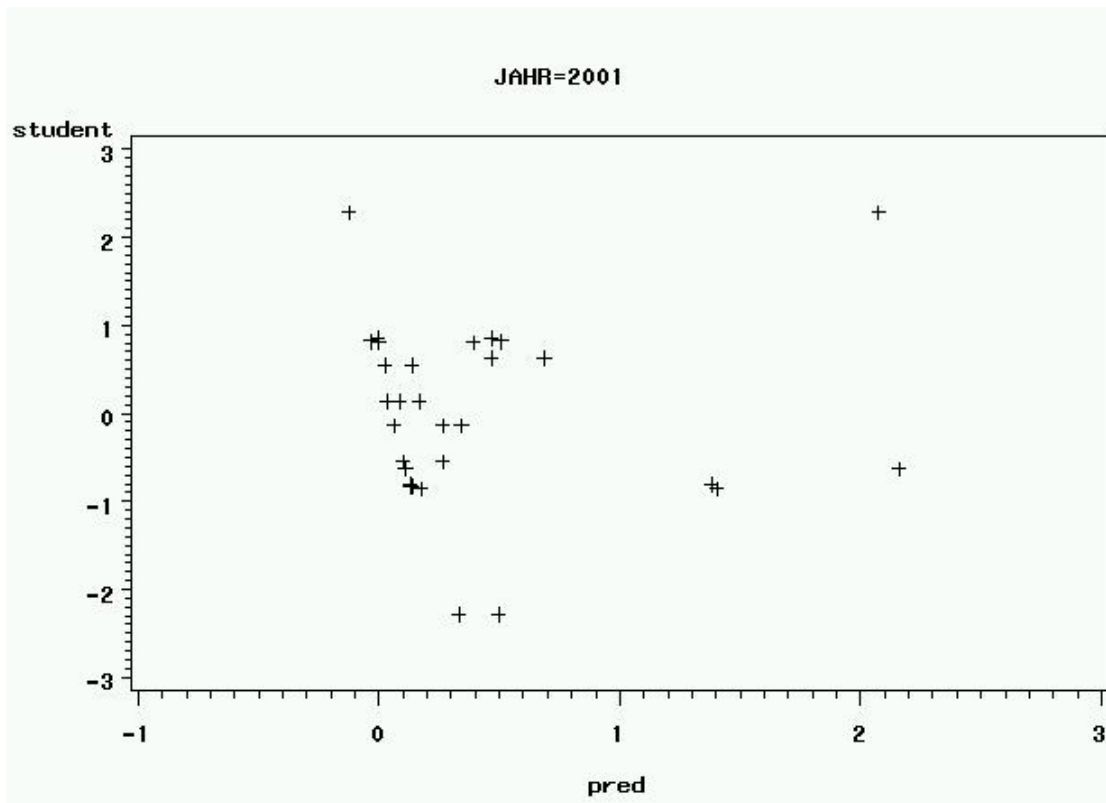
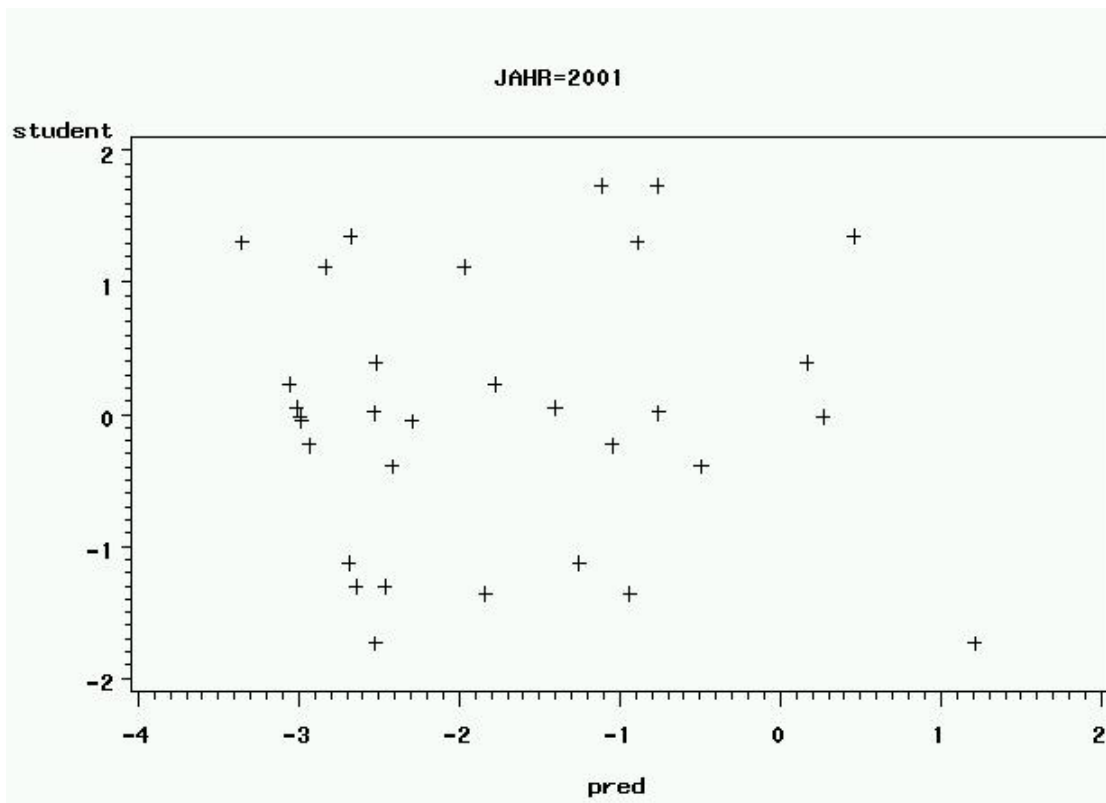


Abb. 4.3: Q-Q-Plot der studentisierten Residuen gegen die Quantile der Normalverteilung nach Log-Transformation



**Abb. 4.4:** Plot der studentisierten Residuen gegen die Erwartungswerte vor Transformation



**Abb. 4.5:** Plot der studentisierten Residuen gegen die Erwartungswerte nach Log-Transformation

Durch die Log-Transformation wird die Anpassung an die Normalverteilung etwas verbessert und die Varianz der Residuen stabilisiert (Abb. 4.5) und wir nehmen deshalb alle weiteren Auswertungen mit den transformierten Daten vor.

Von den festen Effekten sind im Jahr 2001 Bodenbearbeitung, Sorte, die Interaktion Bodenbearbeitung\*Sorte und Fungizid signifikant. Bei den zufälligen Effekten, für die man Varianzkomponentenschätzungen erhält, ist auffällig, dass einige Varianzkomponenten den Wert Null erhalten. Der iterative Schätzalgorithmus der Prozedur MIXED lässt nicht zu, dass negative Varianzkomponenten auftreten. Faktisch bedeutet dieses, dass diese Varianzkomponenten unbedeutend sind und das Modell automatisch um diese reduziert wird. Die Auswertung folgt dann nicht mehr der ursprünglichen Versuchsanlage und kann sich in den F-Werten vom vollständigen Modell und damit der Auswertung nach Proc GLM unterscheiden. Um gleiche Resultate wie mit einer Varianzanalyse mit Proc GLM zu erhalten, ist im Mixed-Code die Option „nobound“ zu wählen, womit äquivalent zum ANOVA-Konzept negative Varianzkomponentenschätzungen zugelassen werden. Dieses sollte aber nur im Fall balancierter Daten mit einfacher Korrelationsstruktur (Compound Symmetry) gewählt werden. Für komplexe gemischte Modelle und unbalancierte Daten empfehlen wir dagegen die Mixed-Voreinstellung ohne „nobound“ (PIEPHO & SPILKE 1999).

Kasten 4.2: Auswertung nach Modell (1) über SAS, Prozedur Mixed, mit nobound

```
Proc mixed data=fusarium nobound; by jahr;
class BB Sorte Fungi WDH BD;
model DON = Wdh BB|Sorte|Fungi/ddfm=kr outpred=pred;
random BB*Wdh Fungi*Wdh BB*Fungi*Wdh BB*Sorte*Wdh; run;
```

Kasten 4.3: Output der einjährigen Auswertung GLM (gekürzt)

```
JAHR=2001

Covariance Parameter
Estimates

Cov Parm           Estimate

BB*WDH              0
Fungi*WDH           0
BB*Fungi*WDH       0.007211
BB*Sorte*WDH       0
Residual            0.2420

Fit Statistics

-2 Res Log Likelihood      46.3
AIC (smaller is better)   50.3
AICC (smaller is better)  50.9
BIC (smaller is better)   50.4

Type 3 Tests of Fixed Effects

Num      Den
```

Effect	DF	DF	F Value	Pr > F
WDH	3	9	1.23	0.3537
BB	1	9	35.80	0.0002
Sorte	1	12	97.57	<.0001
BB*Sorte	1	12	5.70	0.0343
Fungi	1	9	20.43	0.0014
BB*Fungi	1	9	2.14	0.1776
Sorte*Fungi	1	12	4.19	0.0633
BB*Sorte*Fungi	1	12	0.18	0.6785
 Jahr=2002				
Effect	Num DF	Den DF	F Value	Pr > F
WDH	3	3	0.18	0.9054
BB	1	3	23.22	0.0170
Sorte	1	6	58.66	0.0003
BB*Sorte	1	6	0.59	0.4727
Fungi	1	6	21.25	0.0037
BB*Fungi	1	6	0.97	0.3635
Sorte*Fungi	1	6	12.33	0.0126
BB*Sorte*Fungi	1	6	0.43	0.5353

### Mehrjährige Auswertung

Der Versuch ist zweijährig. Es liegt Messwiederholung vor, da in beiden Jahren, die Parzellen an der gleichen Stelle liegen. Zwischen den Messwerten aus verschiedenen Jahren, die an der gleichen Parzelle gewonnen wurden, besteht damit eine Korrelation, sie haben eine Kovarianz. Es ist deshalb eine entsprechende Kovarianzstruktur zu wählen. Bei zeitlich wiederholten Messungen ist das „autoregressive“ Modell [AR(1)] oft ein gut passendes. Das bedeutet, direkt aufeinanderfolgende Jahre sind sich ähnlicher als weiter auseinander liegende Jahre. Da im vorliegenden Versuch nur zwei Jahre geprüft wurden, ist diese Struktur identisch mit der Struktur „Compound Symmetry“.

Ein Modell mit Messwiederholungen sollte grundsätzlich mit der Prozedur Mixed ausgewertet werden. Jede Randomisierungseinheit (Groß-, Mittel-, Kleinteilstücke), auf die eine Messwiederholung angewendet wird, muss als „Subject“ definiert werden. Alle Varianzkomponenten, die unten bereits als Subject auftauchen, können bei der ersten Random-Anweisung gelöscht werden. Groß- und Mittelteilstückfehler werden in einer eigenen Random-Zeile mit der zugehörigen Kovarianzstruktur Compound Symmetry (Type=cs) benannt. Der Restfehler muss in einer Zeile „Repeated“ separat definiert werden. Hier ist als Subject die einzelne Parzelle anzugeben. Wichtig ist in diesem Fall, dass auch die Daten eine Variable „Parzelle“ enthalten.

### Vorgehen für mehrjährige Auswertung

<p>Einjähriges Modell wird um den Faktor Jahr erweitert  &gt; Für alle Effekte zusätzlich eine Interaktion mit dem Jahr in das Modell nehmen</p>
--

Für die Groß-, Mittel- und Kleinteilstücke liegt Messwiederholung vor  
 > Korrelationsstruktur festlegen

Da nur zwei Jahre: CS = AR(1)

Faktor Jahr fix oder zufällig?

Wenn Jahre fix > retrospektive Betrachtung

Wenn Jahre zufällig > Prognose für Zukunft

#### Kasten 4.4: SAS-Code für die mehrjährige Auswertung mit Prozedur MIXED

```
Proc mixed data=fusarium;
class JAHR PARZ BB Sorte Fungi WDH BD;
model log_DON= Wdh BB|Sorte|Fungi/ddfm=kr;
random      Jahr  jahr*wdh  Jahr*BB  Jahr*Sorte  Jahr*Fungi
Jahr*BB*Sorte
Jahr*BB*Fungi Jahr*Sorte*Fungi Jahr*BB*Sorte*Fungi;

random Jahr/subject =BB*Wdh type =cs;
random Jahr/subject =Fungi*Wdh type =cs;
random Jahr/subject=BB*Fungi*Wdh type=cs;
random Jahr/subject=BB*Sorte*Wdh type=cs;

repeated jahr/subject=parz type=cs rcorr;
lsmeans BB*Sorte*Fungi/diff;
run;
```

Nun tritt man in eine Phase ein, in der das Modell gegebenenfalls vereinfacht und optimiert werden muss. Z.B. dann wenn die Auswertung zunächst wegen unendlicher Likelihood abbricht oder der REML-Algorithmus nicht konvergiert. In unserem Fall sieht das so aus:

#### Kasten 4.5: Erster Output mehrjährige Auswertung mit Prozedur MIXED

WARNING: Did not converge.

##### Covariance Parameter Values At Last Iteration

Cov Parm	Subject	Estimate
JAHR		0.3743
JAHR*WDH		0.01607
JAHR*BB		0
JAHR*Sorte		0
JAHR*Fungi		0
JAHR*BB*Sorte		2.93E-20
JAHR*BB*Fungi		0
JAHR*Sorte*Fungi		0
JAHR*BB*Sorte*Fungi		0
Variance	BB*WDH	0.005695
CS	BB*WDH	-0.01234
Variance	Fungi*WDH	0
CS	Fungi*WDH	-0.02491



Variance	BB*Fungi*WDH	0
CS	BB*Fungi*WDH	0.004398
Variance	BB*Sorte*WDH	0.06794
CS	BB*Sorte*WDH	-0.09298
CS	PARZ	0.006838
Residual		0.3515

Das Programm konvergiert in diesem Fall nicht. Man erhält aber einen Statusbericht, wie der Stand der Dinge war, als das Programm abgebrochen hat. Der Nutzer kann hier eingreifen und das Modell vereinfachen. Zufällige Effekte mit negativen Varianzschätzungen (bzw. Varianz = 0), insbesondere Interaktionen höheren Grades sind potentielle Kandidaten, aus dem Modell genommen zu werden. Fixe Effekte mit F-Werten <1 können ebenfalls aus dem Modell herausgenommen werden.

Durch eine Vereinfachung des Modells kann in der Regel letztendlich Konvergenz erreicht werden. Es muss hier bemerkt werden, dass es nicht „Das Modell“ gibt. Vielmehr ist ein sinnvolles Modell mit gutem Fit aus einer Vielzahl von möglichen Modellen zu wählen. Verschiedene Modelle, welche dieselben festen Effekte aufweisen (!), können mit dem AIC-Kriterium verglichen werden. Je kleiner AIC, desto besser beschreibt das Modell die Daten.

Bei den Beispieldaten, war es notwendig, den Mittelteilstückfehler „Bodenbearbeitung\*Fungizid\*Wdh“ auszukommentieren und damit aus dem Modell zu nehmen, um Konvergenz zu erzielen.

Kasten 4.6: Code für die mehrjährige Auswertung mit dem reduzierten Modell

```

title "Mehrjährig Mixed mit Messwiederholung, Transformiert vereinfacht";
Proc mixed data=fusarium;
class JAHR PARZ      BB Sorte Fungi WDH BD;
model log_DON= wdh BB|Sorte|Fungi /ddfm=kr;
random      Jahr  jahr*wdh  Jahr*BB  Jahr*Sorte  Jahr*Fungi
Jahr*BB*Sorte
Jahr*BB*Fungi Jahr*Sorte*Fungi Jahr*BB*Sorte*Fungi;
random Jahr/subject =BB*Wdh type =cs;
random Jahr/subject =Fungi*Wdh type =cs;
*random Jahr/subject=BB*Fungi*Wdh type=cs;
random Jahr/subject=BB*Sorte*Wdh type=cs;
repeated jahr/subject=parz type=cs rcorr;
run;

```

Kasten 4.7: Output für das reduzierte Modell

Covariance Parameter Estimates		
Cov Parm	Subject	Estimate
JAHR		0.2393
JAHR*WDH		0.004935
JAHR*BB		0.002621
JAHR*Sorte		0
JAHR*Fungi		3.93E-19

JAHR*BB*Sorte		0
JAHR*BB*Fungi		0
JAHR*Sorte*Fungi		0
JAHR*BB*Sorte*Fungi		0
Variance	BB*WDH	0
CS	BB*WDH	0.03860
Variance	Fungi*WDH	0
CS	Fungi*WDH	-0.02789
Variance	BB*Sorte*WDH	0.08363
CS	BB*Sorte*WDH	-0.1073
CS	PARZ	0.1433
Residual		0.1181

Kasten 4.8: Fortsetzung Output für das reduzierte Modell

Fit Statistics				
-2 Res Log Likelihood			96.3	
AIC (smaller is better)			114.3	
AICC (smaller is better)			118.4	
BIC (smaller is better)			102.5	
Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
WDH	3	1	0.39	0.7936
BB	1	1.33	47.80	0.0525
Sorte	1	6	330.84	<.0001
BB*Sorte	1	6	9.66	0.0209
Fungi	1	3	48.58	0.0061
BB*Fungi	1	9	1.63	0.2338
Sorte*Fungi	1	9	5.03	0.0517
BB*Sorte*Fungi	1	9	0.20	0.6685

Die Haupteffekte Sorte und Fungizid sind signifikant, bei signifikanten Interaktionen zwischen Sorte und Bodenbearbeitung. Der Haupteffekt Bodenbearbeitung sowie die Interaktion zwischen Sorte und Fungizid sind an der Signifikanzschwelle. Mittelwertvergleiche sollten auf Stufe der signifikanten Interaktionen vorgenommen werden. Dieses wird durch Einfügen der Zeile

```
lsmeans Sorte*BB Sorte*Fungi/pdiff;
```

in den SAS-Code erreicht.

Wir erhalten zunächst die Kleinst-Quadrat-Schätzungen für die entsprechenden Stufenmittel. Es ist zu beachten, dass diese auf der logarithmischen Skala vorliegen. Die Sorte Bandit weist jeweils den höheren DON-Gehalt auf. Mulchen führt bei beiden Sorten zu höherem DON-Gehalt. Fungizidbehandlung führt bei beiden Sorten zu einem geringeren DON-Gehalt.

#### Kasten 4.9: Mittelwerte

Effect	BB	Sorte	Fungi	Least Squares Means			t Value	Pr >  t
				Estimate	Standard Error	DF		
BB*Sorte	Mulch	Bandit		0.2172	0.3697	1.18	0.59	0.6492
BB*Sorte	Mulch	Petrus		-1.7949	0.3697	1.18	-4.86	0.1012
BB*Sorte	Pflug	Bandit		-1.2907	0.3697	1.18	-3.49	0.1467
BB*Sorte	Pflug	Petrus		-2.7156	0.3697	1.18	-7.35	0.0628
Sorte*Fungi		Bandit	behandelt	-1.0862	0.3686	1.17	-2.95	0.1778
Sorte*Fungi		Bandit	unbehandelt	0.01274	0.3686	1.17	0.03	0.9774
Sorte*Fungi		Petrus	behandelt	-2.4481	0.3686	1.17	-6.64	0.0714
Sorte*Fungi		Petrus	unbehandelt	-2.0624	0.3686	1.17	-5.60	0.0870

Außer dem Vergleich „Petrus gemulcht“ gegen „Bandit gepflügt“ und „Petrus behandelt“ gegen „Petrus unbehandelt“ sind alle Mittelwertvergleiche signifikant. Grenzdifferenzen können in einem Data-Step berechnet werden.

```
data diffs;
set diffs;
alpha=0.05;
p=1-(alpha/2);
tTab=TINV(p,df);
LSD=stderr*tTab;
run;
```

Zur besseren Interpretierbarkeit werden die Mittelwerte von der logarithmischen Skala auf die ursprüngliche Skala zurücktransformiert. Hierbei ist lediglich zu beachten, dass die zurücktransformierten Mittelwerte Schätzwerte für den Median, nicht aber das arithmetische Mittel (Erwartungswert) sind (siehe Anhang).

#### Kasten 4.10: Transformierte und rücktransformierte Mittelwerte

Obs	BB	Sorte	Fungi	Estimate	est_org
1	Mulch	Bandit		0.2172	1.24263
2	Mulch	Petrus		-1.7949	0.16614
3	Pflug	Bandit		-1.2907	0.27509
4	Pflug	Petrus		-2.7156	0.06617
5		Bandit	behandelt	-1.0862	0.33751
6		Bandit	unbehandelt	0.01274	1.01282
7		Petrus	behandelt	-2.4481	0.08646
8		Petrus	unbehandelt	-2.0624	0.12715

#### Auswertung mit fixen Jahren

Eine alternative Auswertung ist die mit fixen Jahren. Es ist natürlich zu beachten, dass der Aussagebereich bei fixen Jahren auf die im Versuch geprüften Jahre beschränkt ist. Eine Verallgemeinerung ist nicht zulässig.

Die pflanzenbauliche Interpretation bleibt bei den Beispieldaten unverändert, mit der Ausnahme, dass nun auch der Effekt der Bodenbearbeitung signifikant ist.

#### Kasten 4.11: Programmcode für Auswertung mit fixen Jahren

```
title "Mehrjährig Mixed mit Messwiederholung, Jahre fix";
```

```

Proc mixed data=fusarium;
class JAHR PARZ BB Sorte Fungi WDH BD;
model log_DON= Wdh jahr*wdh BB|Sorte|Fungi|Jahr/ddfm=kr;
random Jahr/subject =BB*Wdh type =cs;
random Jahr/subject =Fungi*Wdh type =cs;
*random Jahr/subject=BB*Fungi*Wdh type=cs;
*random Jahr/subject=BB*Sorte*Wdh type=cs;
repeated jahr/subject=parz type=cs rcorr;
*lsmeans BB*Sorte*Fungi/diff;
run;

```

### Auswertung der Mittelwerte über beide Jahre (Summary measures)

Eine einfache und robuste Methode, Messwiederholungen zu handhaben, ist es, über die Stufen des Messwiederholungsfaktors zu mitteln und dann nach dem Programm für einjährige Daten auszuwerten. Unten sind Programmcode und Output für diese Auswertung angegeben. Sie führt zu einem weitgehend ähnlichen Ergebnis wie die Auswertung mit Messwiederholung. Da wechselnde Ergebnisse zwischen den Jahren nicht berücksichtigt werden, ist hier ebenso wie bei fixen Jahren der Aussagebereich eingeschränkt. Außerdem erlaubt diese Art der Auswertung keine Analyse der Jahreswechselwirkungen.

Kasten 4.12: Programmcode für *Summary measures* -Ansatz

```

proc sort data=fusarium;
by wdh bb fungi sorte;
proc means noprint;
by wdh bb fungi sorte;
var log_don;
output out=mittel mean=log_don;
run;

Proc mixed data=mittel;
class BB Sorte Fungi WDH BD;
model log_DON = Wdh BB|Sorte|Fungi/ddfm=kr outpred=pred;
random BB*Wdh Fungi*Wdh BB*Fungi*Wdh BB*Sorte*Wdh;
run;

```

Kasten 4.13: Output für *Summary measures* -Ansatz

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
WDH	3	3	0.23	0.8709
BB	1	3	52.24	0.0055
Sorte	1	12	237.43	<.0001
BB*Sorte	1	12	6.93	0.0219
Fungi	1	6	19.92	0.0043
BB*Fungi	1	6	1.49	0.2681
Sorte*Fungi	1	12	10.22	0.0077
BB*Sorte*Fungi	1	12	0.40	0.5399

## Daten

JAHR	PARZ	BB	Sorte	Fungi	WDH	DON	
2001	1	Pflug	Petrus		unbehandelt	1	0.05
2001	2	Pflug	Bandit		unbehandelt	1	0.59
2001	3	Mulch	Petrus		unbehandelt	1	0.08
2001	4	Mulch	Bandit		unbehandelt	1	1.32
2001	5	Mulch	Bandit		behandelt	2	0.47
2001	6	Mulch	Petrus		behandelt	2	0.05
2001	7	Pflug	Petrus		behandelt	2	0.05
2001	8	Pflug	Bandit		behandelt	2	0.10
2001	9	Pflug	Petrus		behandelt	1	0.05
2001	10	Pflug	Bandit		behandelt	1	0.06
2001	11	Mulch	Petrus		behandelt	1	0.09
2001	12	Mulch	Bandit		behandelt	1	0.55
2001	13	Mulch	Bandit		unbehandelt	2	1.30
2001	14	Mulch	Petrus		unbehandelt	2	0.08
2001	15	Pflug	Petrus		unbehandelt	2	0.05
2001	16	Pflug	Bandit		unbehandelt	2	0.25
2001	17	Mulch	Bandit		unbehandelt	3	2.30
2001	18	Mulch	Petrus		unbehandelt	3	0.11
2001	19	Pflug	Bandit		unbehandelt	3	0.33
2001	20	Pflug	Petrus		unbehandelt	3	0.05
2001	21	Pflug	Petrus		behandelt	4	0.05
2001	22	Pflug	Bandit		behandelt	4	0.19
2001	23	Mulch	Bandit		behandelt	4	0.53
2001	24	Mulch	Petrus		behandelt	4	0.05
2001	25	Mulch	Bandit		behandelt	3	0.27
2001	26	Mulch	Petrus		behandelt	3	0.10
2001	27	Pflug	Bandit		behandelt	3	0.18
2001	28	Pflug	Petrus		behandelt	3	0.05
2001	29	Pflug	Petrus		unbehandelt	4	0.08
2001	30	Pflug	Bandit		unbehandelt	4	0.21
2001	31	Mulch	Bandit		unbehandelt	4	2.10
2001	32	Mulch	Petrus		unbehandelt	4	0.75
2002	1	Pflug	Petrus		unbehandelt	1	0.06
2002	2	Pflug	Bandit		unbehandelt	1	0.90
2002	3	Mulch	Petrus		unbehandelt	1	0.48
2002	4	Mulch	Bandit		unbehandelt	1	3.66
2002	5	Mulch	Bandit		behandelt	2	1.49
2002	6	Mulch	Petrus		behandelt	2	0.28
2002	7	Pflug	Petrus		behandelt	2	0.06
2002	8	Pflug	Bandit		behandelt	2	0.22
2002	9	Pflug	Petrus		behandelt	1	0.05
2002	10	Pflug	Bandit		behandelt	1	0.29
2002	11	Mulch	Petrus		behandelt	1	0.6
2002	12	Mulch	Bandit		behandelt	1	1.08
2002	13	Mulch	Bandit		unbehandelt	2	2.58
2002	14	Mulch	Petrus		unbehandelt	2	0.26
2002	15	Pflug	Petrus		unbehandelt	2	0.12
2002	16	Pflug	Bandit		unbehandelt	2	0.5
2002	17	Mulch	Bandit		unbehandelt	3	2.09
2002	18	Mulch	Petrus		unbehandelt	3	0.3
2002	19	Pflug	Bandit		unbehandelt	3	0.54
2002	20	Pflug	Petrus		unbehandelt	3	0.16
2002	21	Pflug	Petrus		behandelt	4	0.05
2002	22	Pflug	Bandit		behandelt	4	0.25
2002	23	Mulch	Bandit		behandelt	4	1.1
2002	24	Mulch	Petrus		behandelt	4	0.11
2002	25	Mulch	Bandit		behandelt	3	0.63
2002	26	Mulch	Petrus		behandelt	3	0.1

2002	27	Pflug Bandit	behandelt	3	0.21
2002	28	Pflug Petrus	behandelt	3	0.25
2002	29	Pflug Petrus	unbehandelt	4	0.05
2002	30	Pflug Bandit	unbehandelt	4	0.63
2002	31	Mulch Bandit	unbehandelt	4	4.79
2002	32	Mulch Petrus	unbehandelt	4	0.41

## Beispiel 5: Auswertung eines zweifaktoriellen Dauerversuchs mit dreigliedriger Fruchtfolge

### Beschreibung

Beim Systemversuch des IfZ Göttingen (Daten von H.P. König, Institut für Zuckerrübenforschung, Göttingen) handelt es sich um eine dreifeldrige Fruchtfolge (Zuckerrüben-Winterweizen-Wintergerste), die auf drei benachbarten Teilschlägen mit Fruchtwechsel angelegt ist. Zwischen Gerste und Rüben wird noch jeweils eine Zwischenfrucht zur Nematodenbekämpfung angebaut. Jeder Teilschlag ist in eine Spaltanlage mit vier Wiederholungen aufgeteilt. Großteilstückfaktor ist die Bodenbearbeitung (Pflug, Mulch), Kleinteilstückfaktor ist die Stickstoffdüngung (N0, N1, N2, N3). Die Behandlungen sind ortsfest, die Stickstoffstufen sind den Fruchtarten angepasst (N1 Weizen ist z.B. nicht gleich N1 Zuckerrüben). Der Versuch befindet sich mittlerweile in der 4. Rotation.

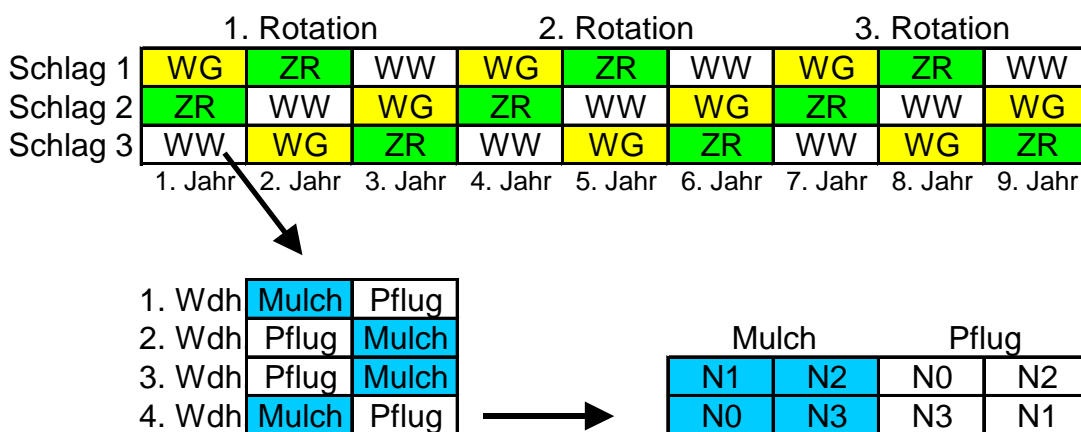


Abb. 5.1: Versuchsplan

### Vorbemerkungen zum Modellaufbau

Bei Dauerversuchen erfolgt meist eine Auswertung getrennt nach Fruchtart. Bei einem Versuchsdesign wie oben kann die Kombination von Schlag und Jahr als zufälliger Umwelteffekt aufgefasst werden. Jeweils nach einer Rotation liegt eine Fruchtart wieder auf dem gleichen Schlag. Diese Messwiederholung ist im Modell durch Wahl einer Korrelationsstruktur entsprechend zu berücksichtigen.

Für Bodenmerkmale kann eine zusammenfassende Auswertung über alle Jahre und alle Fruchtarten sinnvoll sein. In diesem Fall ist allerdings die Vermengung der Faktoren Schlag, Jahr und Fruchtart zu berücksichtigen. Es ist ferner ein Fruchtarteffekt anzunehmen. Eine befriedigende Auswertung ist wegen der drei getrennten Schläge und der angesprochenen Vermengungen nur möglich bei Annahme zufälliger Jahreseffekte.

Der Systemversuch in Göttingen läuft auf drei Schlägen, die unmittelbar benachbart sind. Nach gängiger Praxis würde man Jahr und Schlag als zufällige Faktoren auffassen, wobei der Faktor „Schlag“ so wie der Faktor „Ort“ in einer Versuchsserie zu behandeln ist. Das macht aber nur Sinn, wenn diese Orte (bzw. Schläge) eine zufällige Stichprobe aus einer größeren Grundgesamtheit darstellen. Das ist hier nicht der Fall. Die drei Schläge sind bereits die Grundgesamtheit. Sie sind nur repräsentativ für den spezifischen, vor vielen Jahren gewählten Versuchsstandort. Der Faktor Schlag sollte deshalb als fix betrachtet werden. Die Blöcke (Wiederholungen) innerhalb der Schläge betrachten wir ebenfalls als fix.

Jahre werden in einer Versuchsserie meistens als zufällig angesehen. Sie stellen eine Zufallsstichprobe dar, repräsentieren unter anderem das Klima am Standort und erlauben damit Prognosen für künftige Jahre. Im dargestellten Versuch ist ein Jahr ist in seiner Wirkung nicht ganz zufällig, sondern hängt auch von der Wirkung des Vorjahrs ab (Autokorrelation).

### Modellaufbau

Analog zu PIEPHO et al. (2003a) trennen wir zwischen Behandlungs- und Designeffekten. Die Behandlungseffekte lassen sich durch folgendes Modell beschreiben:

$$\alpha_i + \beta_j + (\alpha\beta)_{ij}$$

wobei:

$\alpha_i$  = Effekt der  $i$ -ten Bodenbearbeitungsvariante

$\beta_j$  = Effekt der  $j$ -ten Düngungsstufe

$(\alpha\beta)_{ij}$  = Interaktion zwischen  $i$ -ter Bodenbearbeitung und  $j$ -ter Düngung

Wir können die Fruchtart formal ebenfalls als Behandlung auffassen. Damit ergibt sich:

$$\alpha_i + \beta_j + \phi_k + (\alpha\beta)_{ij} + (\beta\phi)_{ik} + (\alpha\phi)_{jk} + (\alpha\beta\phi)_{ijk}$$

wobei:

$\alpha_i$  = Effekt der  $i$ -ten Bodenbearbeitungsvariante

$\beta_j$  = Effekt der  $j$ -ten Düngungsstufe

$\phi_k$  = Effekt der  $k$ -ten Fruchtart

Mit der Kurzschreibweise nach PIEPHO et al. (2003a) formulieren wir:

$$B \times N \times F = B + N + F + B \bullet N + B \bullet F + N \bullet F + B \bullet N \bullet F$$

wobei: B = Bodenbearbeitung, N = Stickstoffdüngung, F = Fruchtart

Für die Designeffekte ergibt sich:

$$S / W / \text{Mainplot} / \text{Subplot} = S + S \bullet W + S \bullet W \bullet B + S \bullet W \bullet B \bullet N$$

Wobei:

S	= Effekt eines Schlages
S • W	= Effekt einer Wiederholung innerhalb eines Schlages
S • W • B	= Großteilstückfehler
S • W • B • N	= Parzellen- bzw. Restfehler

Dann werden Behandlungs- und Designmodell jeweils separat mit dem Messwiederholungsfaktor kombiniert.

Wenn die Jahre zufällig sind, so sind ebenfalls alle Interaktionen mit „Jahr“ zufällig. Um Trends zu erfassen, nehmen wir einen fixen Effekt für die Rotation in das Modell auf und betrachten Jahre als geschachtelt in Rotation.



### Mehrjähriges Modell für Behandlungseffekte:

$$\begin{aligned} B \times N \times F \times (R/J) &= B + N + F + B \cdot N + B \cdot F + N \cdot F + B \cdot N \cdot F \\ &+ B \cdot R + N \cdot R + F \cdot R + B \cdot N \cdot R + B \cdot F \cdot R + N \cdot F \cdot R + B \cdot N \cdot F \cdot R \\ &+ B \cdot R \cdot J + N \cdot R \cdot J + F \cdot R \cdot J + B \cdot N \cdot R \cdot J + B \cdot F \cdot R \cdot J + N \cdot F \cdot R \cdot J \\ &+ B \cdot N \cdot F \cdot R \cdot J \\ &+ R + R \cdot J \end{aligned}$$

wobei: R = Rotationseffekt = zeitlicher Trend (fix)  
R•J = Jahr geschachtelt in Rotation (zufällig)

### Mehrjähriges Modell für Designeffekte:

Zur Codierung der Großteilstücke können wir die Stufe der Bodenbearbeitung nutzen, die auf einem Großteilstück liegt. Der Effekt der Blöcke und der Großteilstückfehler lassen sich in eine jahresspezifische, eine rotationsspezifische und in eine über alle Jahre wirkende Komponente zerlegen.

$$\begin{aligned} (S / W / \text{Mainplot} / \text{Subplot}) \times (R/J) \\ &= S + S \cdot W + S \cdot W \cdot B \\ &+ S \cdot R + S \cdot W \cdot R + S \cdot W \cdot B \cdot R \\ &+ S \cdot J \cdot R + S \cdot W \cdot J \cdot R + S \cdot W \cdot B \cdot J \cdot R \\ &+ R + J \cdot R \\ &+ \underline{S \cdot W \cdot B \cdot N \cdot J \cdot R} \end{aligned}$$

wobei: S•W•B•N•J•R = Restfehler

Allerdings ist die Annahme eines rotationsspezifischen Großteilstückfehlers nicht sonderlich sinnvoll. Dieser bedingt, dass sich Beobachtungen aus der gleichen Rotation ähnlicher sind als solche aus verschiedenen Rotationen. Da der Faktor Rotation aber nachträglich auf die Versuchsjahre gelegt worden ist, macht dieses biologisch keinen Sinn. Zwei aufeinanderfolgende Jahre sollten immer gleich korreliert sein, egal zu welcher Rotation sie formal gehören. Deshalb wäre dieser Effekt nachträglich wieder aus dem Modell zu eliminieren. Eine Alternative zur Herleitung des Modells folgt PIEPHO et al. 2003b. Hierbei ist für die Kombination des einjährigen Designmodells mit dem Messwiederholungsfaktor nicht der „Kreuz“- sondern der „pt“-Operator zu verwenden. Dieser sorgt dafür dass ausschließlich eine Kombination mit dem Jahr erfolgt:

$$\begin{aligned} (S / W / \text{Mainplot} / \text{Subplot}) \bullet \text{pt}(R/J) \\ &= S + S \cdot W + S \cdot W \cdot B \\ &+ S \cdot J \cdot R + S \cdot W \cdot J \cdot R + S \cdot W \cdot B \cdot J \cdot R \\ &+ \underline{S \cdot W \cdot B \cdot N \cdot J \cdot R} \end{aligned}$$

wobei: S•W•B•N•J•R = Restfehler

Begründung hierfür ist, dass eine hierarchische Strukturierung der Designeffekte nach R/J innerhalb eines gemischten Modells an dieser Stelle nicht erforderlich ist.

Diese Strukturierung erfolgt erst bei der Wahl der Varianz-Kovarianz-Struktur für die zufälligen Effekte.

Das Modell für die Designeffekte und das für die Behandlungseffekte werden kombiniert und nach fixen und zufälligen Effekten sortiert. Feste Effekte sind zuerst und zufällige nach einem Doppelpunkt aufgeführt. Den Grundsätzen von PIEPHO et al. (2003a, 2003b) folgend erhalten wir:

$$\begin{aligned}
 Y = & R + B + N + F + B \cdot N + B \cdot F + N \cdot F + B \cdot N \cdot F + B \cdot R + N \cdot R + F \cdot R + B \cdot N \cdot R + \\
 & B \cdot F \cdot R + N \cdot F \cdot R + B \cdot N \cdot F \cdot R + S + S \cdot W : J \cdot R + S \cdot J \cdot R + S \cdot W \cdot J \cdot R \\
 & + B \cdot R \cdot J + N \cdot R \cdot J + F \cdot R \cdot J + B \cdot N \cdot R \cdot J + B \cdot F \cdot R \cdot J + N \cdot F \cdot R \cdot J + B \cdot N \cdot F \cdot R \cdot J + S \cdot W \cdot B \\
 & + S \cdot W \cdot B \cdot J \cdot R + \underline{S \cdot W \cdot B \cdot N \cdot J \cdot R}
 \end{aligned}$$

Wir können uns die Kombination von Behandlungs- und Designeffekten auch graphisch darstellen (Abb. 5.2). Wir postulieren zunächst, dass zwischen Behandlungs- und Designeffekten keine Interaktionen vorliegen.

**Abb. 5.2:** Behandlungs- und Designeffekte und Interaktionen zwischen diesen

	B	N	F	BN	BF	NF	BNF
R	RB	RN	RF	RBN	RBF	RNF	RBNF
JR	JRB	JRN	JRF	JRBN	JRBF	JRNF	JRBNF
S							
SJR							
SW							
SWJR							
SWB							
SWBJR							

Es fällt auf, dass einige Effekte miteinander vermengt (confounded) sind. Der Versuch ist nicht orthogonal, da nur ein Drittel der Kombinationen Jahr\*Schlag\*Frucht realisiert wird. So umfasst beispielsweise der Effekt Schlag\*Jahr gleichzeitig die Effekte Frucht\*Jahr und Schlag\*Frucht\*Jahr, da in einem Jahr jede Fruchtart nur auf einem Schlag steht und jeder Schlag jährlich auch nur eine Frucht trägt. Damit ist abschließend eine Modellreduktion um diese vermengten bzw. redundanten Effekte notwendig.

Hierzu folgende Überlegung: Wenn die Frucht als Behandlung aufgefasst wird, so ist die Struktur innerhalb einer Rotation analog zu einem „Lateinischen Quadrat“. Eine Frucht ist in Jahr 1, Jahr 2 und Jahr 3 dieselbe „Behandlung“ da es sich um eine Rotation handelt, also Vorfruchtwirkungen identisch sind. Im Unterschied zu „konventionellen“ Lateinischen Quadraten ist allerdings zu beachten, dass in diesem Fall keine Randomisation der Zeilen und Spalten erfolgt.

	1. Rotation		
Schlag 1	Gerste	Rüben	Weizen
Schlag 2	Rüben	Weizen	Gerste
Schlag 3	Weizen	Gerste	Rüben
	Jahr 1	Jahr 2	Jahr 3

**Abb. 5.3:** Darstellung einer Rotation als Lateinisches Quadrat

Wenn man das Jahr als zufällig betrachtet, sind auch alle Interaktionen mit dem Jahr zufällig. Die aus der Vermengung von Effekten resultierenden Probleme sind dann akzeptabel, da die zufälligen Effekte nicht geschätzt werden müssen (als Mittelwert für zufällige Effekte wird Null postuliert) und weil man nur an den Varianzen interessiert ist. Ein sehr einfaches Modell für unser „Lateinisches Quadrat“ sieht dann so aus:

$$y_{lmn} = \mu + s_n + j_m + f_l + e_{lmn}$$

$s_n$  = Effekt des n-ten Schlages

$j_m$  = Effekt des m-ten Jahres (zufällig, normalverteilt, Mittelwert 0, Varianz  $\sigma_j^2$ ),  
 $\sim(N, 0, \sigma_j^2)$

$f_l$  = Effekt der l-ten Fruchtart

$e_{lmn}$  = Fehler der l-ten Fruchtart im m-ten Jahr auf dem n-ten Schlag  
 (enthält in unserem Fall diverse Interaktionen)

Die unter  $e_{lmn}$  subsummierten Interaktionen Schlag\*Jahr, Schlag\*Fruchtart, Jahr\*Fruchtart, Schlag\*Jahr\*Fruchtart sind miteinander vermengt und nicht separat schätzbar. Die drei zufälligen Effekte Schlag\*Jahr, Jahr\*Fruchtart, Schlag\*Jahr\*Fruchtart können, da man nur die Varianzkomponente schätzen will, zu Schlag\*Jahr\*Fruchtart zusammengefasst werden. Problematisch bleibt der fixe Effekt Schlag\*Fruchtart. Man muss vereinfachend diesen Effekt gleich Null setzen. Das bedeutet, man erwartet keine Interaktion zwischen Fruchtart und Schlag. Das ist ein ähnliches Vorgehen wie in einer Blockanlage, auch hier wird keine Interaktion zwischen Block und Behandlung erwartet. Wenn Schlag\*Fruchtart aus dem Modell herausfällt, bleibt das einfache Modell:

$$y_{lmn} = \mu + s_n + j_m + f_l + fjs_{lmn} + e_{lmn}$$

$s_n$  = Effekt des n-ten Schlages

$j_m$  = Effekt des m-ten Jahres (zufällig, normalverteilt, Mittelwert 0, Varianz  $\sigma_j^2$ ),  
 $\sim(N, 0, \sigma_j^2)$

$f_l$  = Effekt der l-ten Fruchtart

$fjs_{lmn}$  = Interaktion von Schlag, Jahr und Frucht

$e_{lmn}$  = Fehler der l-ten Fruchtart im m-ten Jahr auf dem n-ten Schlag

Für die Formulierung des kompletten Modells gehen wir analog vor. Alle Effekte, die Schlag\*Jahr oder Fruchtart\*Jahr enthalten, werden zu Schlag\*Fruchtart\*Jahr zusammengefasst bzw. umcodiert (Abb. 5.4). Eine Interaktion zwischen Schlag\*Fruchtart wird jedoch zunächst nicht ins Modell aufgenommen.

	B	N	F	BN	BF	NF	BNF
R	RB	RN	RF	RBN	RBF	RNF	RBNF
JR	JRB	JRN	JRF	JRBN	JRBF	JRNF	JRBNF
S							
SJR			SJRF		SJRBF	SJRNF	SJRBNF

Abb. 5.4: Schema der Zusammenfassung von Modellelementen

Damit resultiert folgendes Modell (in Kurzschreibweise):

$$\begin{aligned}
 Y = & R + B + N + F + B \cdot N + B \cdot F + N \cdot F + B \cdot N \cdot F + B \cdot R + N \cdot R + F \cdot R + B \cdot N \cdot R + \\
 & B \cdot F \cdot R + N \cdot F \cdot R + B \cdot N \cdot F \cdot R + S + S \cdot W : J \cdot R + S \cdot F \cdot W \cdot J \cdot R + B \cdot R \cdot J + N \cdot R \cdot J \\
 & + B \cdot N \cdot R \cdot J + S \cdot W \cdot B + S \cdot W \cdot B \cdot R + S \cdot F \cdot W \cdot B \cdot J \cdot R + \boxed{S \cdot J \cdot R \cdot F} + \boxed{S \cdot B \cdot F \cdot R \cdot J} + \\
 & \boxed{S \cdot N \cdot F \cdot R \cdot J} + \boxed{S \cdot B \cdot N \cdot F \cdot R \cdot J} + \boxed{S \cdot W \cdot B \cdot N \cdot J \cdot R}
 \end{aligned}$$

Falls die Schläge eine differenzierte Wirkung der Behandlungen erwarten lassen, so müssen Interaktionen Behandlung\*Schlag ins Modell aufgenommen werden. Das Problem der Vermengung wird dann allerdings sehr massiv.

	B	N	F	BN	BF	NF	BNF
R	RB	RN	RF	RBN	RBF	RNF	RBNF
JR	JRB	JRN	JRF	JRBN	JRBF	JRNF	JRBNF
S	SB	SN	SF	SBN	SBF	SNF	SBNF
SJR	SJRB	SJRN	SJRF	SJRBN	SJRBF	SJRNF	SJRBNF

Abb. 5.4: Schema der Zusammenfassung von Modellelementen bei Interaktionen Schlag\*Behandlung

Ob Interaktionen Schlag\*Behandlung in das Modell gehören, ist zu diskutieren. Im vorliegenden Fall würden die Schläge wie Orte in einer Versuchsserie behandelt, damit sind Interaktionen sinnvoll. Es gibt aber auch Argumente dafür, benachbarte Schläge wie Blöcke zu behandeln und gemäß dem Postulat zu verfahren, dass es keine Interaktionen zwischen Behandlungs- und Blockeffekten gibt.

Das entwickelte Modell sollte mit PROC MIXED ausgewertet werden. Die Prozedur GLM kann zur Prüfung des Modells eingesetzt werden, da GLM einzelne Summenquadrate mit  $SQ=0$  berechnet falls noch Vermengung von Effekten vorliegt, was bei MIXED nicht automatisch der Fall sein muss.

### **Berücksichtigung der Messwiederholung**

Die sich aus der Messwiederholung ergebenden Korrelationen können in SAS Proc MIXED über ein Zusammenspiel der Random- und Repeated-Anweisung mit der Option Subject modelliert werden. Es können unterschiedliche Formen von Korrelationen spezifiziert werden. "Compound Symmetry" bedeutet, alle Jahre werden als untereinander gleichermaßen korreliert angenommen. Eine sinnvollere denkbare Kovarianzstruktur ist AR(1) (Autoregressive Order 1), was bedeutet, dass zeitlich benachbarte Jahre als enger korreliert angenommen werden als entfernte. Wir verwenden hier jeweils die Anweisung „Subject“, um das Subjekt zu definieren auf das sich eine Messwiederholung bezieht. Für Klein- und Großteilstücke wird die Korrelationsstruktur AR(1) angepasst. Für alle übrigen zufälligen Effekte wurde ebenfalls die „Subject“-Schreibweise genutzt. Zu den im einzelnen zur Verfügung stehenden Optionen siehe LITTELL et al. (1996).

In diesem Zusammenhang kommt die Frage auf, ob nicht auch für die zufälligen Effekte Jahr\*Schlag\*Fruchtart und Jahr\*Schlag\*Fruchtart\*Wdh eine Autokorrelation angenommen werden sollte. Prinzipiell ist dieses denkbar. Wenn eine Fruchtart in einem Jahr auf einem Schlag eine positive Interaktion zeigt, hinterlässt sie evtl. mehr Biomasse auf dem Feld, wovon die Nachfrucht auf der gesamten Fläche beeinflusst wird. Praktisch sind diesem Ansatz allerdings Grenzen gesetzt, da die Korrelationen aus den nur drei Schlägen und 12 Wiederholungen nicht so genau geschätzt werden können. Da man ein „overfitting“ vermeiden sollte, bleiben wir hier bei einfacheren Modellen und nehmen nur für Groß- und Kleinteilstücke eine Korrelation an, während für alle anderen Versuchseinheiten ein Modell mit unabhängigen Effekten gelten soll.

Bei der SAS-Syntax kann der Effekt J\*R auch kurz durch J dargestellt werden, da die Jahre von 1993 bis 2001 fortlaufend durchnummeriert sind und damit eindeutig zuzuordnen sind. Verschiedene Modelle, die in den festen Effekten übereinstimmen, können über das AIC-Kriterium verglichen werden. Bei diesem Kriterium wird die Schätzung der Log-Likelihood um einen Strafterm korrigiert, der von der Anzahl der Varianzkomponenten abhängt. Einen niedrigen Wert haben solche Modelle, die einen guten Fit aufweisen aber mit vergleichbar wenig Parametern auskommen.

Im weiteren Verlauf kann das Modell nun evtl. um unbedeutende zufällige (Varianz=0) und nicht signifikante fixe Effekten reduziert werden. Diese Reduktion hat im allgemeinen nur geringen Einfluss darauf, welche fixen Effekte letztendlich als signifikant verbleiben. In gewissem Umfang werden aber die Varianzkomponenten und Korrelationen vom fixen Teil des Modells beeinflusst. Man muss deshalb im Wechsel den fixen und zufälligen Teil des Modells anpassen (WOLFINGER 1993).

Mögliche Erweiterung des Modells wäre die Aufnahme eines Regressionsterms für die Rotationen oder die Jahre.

Kasten: Beispiel SAS-Syntax für Modell mit Korrelationsstruktur „Autoregressiv“  
title "Komplettes Modell REML, „Korrelationsstruktur Compound Symmetry“;

```
Proc mixed data=daten; where jahr>1995;
class R J S N B F W mainplot subplot;
model ngespfl = R B N F B*N B*F
              N*F B*N*F B*R N*R F*R
              B*N*R B*F*R N*F*R B*N*F*R
              S S*W /ddfm=kr ;
random J S*F*W*J B*J N*J B*N*J S*W*B S*J*F S*B*F*J S*N*F*J S*B*N*F*J;
random    J /subject=W*S*Mainplot Type=AR(1);
repeated  J /subject=W*S*Subplot Type=AR(1);
run;
```

## Beispiel 6: Transformation von Prozentzahlen

Es wurde ein Laborexperiment zur Dormanzinduktion bei verschiedenen Rapssorten durchgeführt (S. Gruber, Institut für Pflanzenbau und Grünland, Hohenheim). Die Sorten lagen jeweils in einer genetisch modifizierten Variante und in einer nah-isogenen Variante (unmodifiziert) vor. Eine Versuchsfrage war, ob die genetische Transformation die Dormanz beeinflusst.

Für jede Sorte wurden Samen auf Petrischalen zur Keimung ausgelegt. Mehrere Petrischalen wurden jeweils in einem lichtverschlossenen Behälter platziert. Die Auswertung erfordert ein gemischtes Modell mit zufälligen Effekten für die beiden Randomisationseinheiten "Behälter" und "Petrischale". Am Ende des Experiments wurde der Anteil nicht-gekeimter Samen ermittelt. Die Daten sind in Tab. 6.1 wiedergegeben.

Die Daten wurden nach dem folgenden Modell ausgewertet:

$$y_{ijkh} = \mu + \alpha_i + r_j + (\alpha r)_{ij} + c_{jk} + e_{ijkh}$$

wobei

$y_{ijkh}$  = Anteil gekeimter Samen in  $h$ -ter Wiederholung für  $i$ -te Sorte im  $j$ -ten Experiment und  $k$ -ten Kontainer.

$\mu$  = allgemeiner Effekt

$\alpha_i$  = Haupteffekt der  $i$ -ten Sorte

$r_j$  = Haupteffekt des  $j$ -ten Experiments

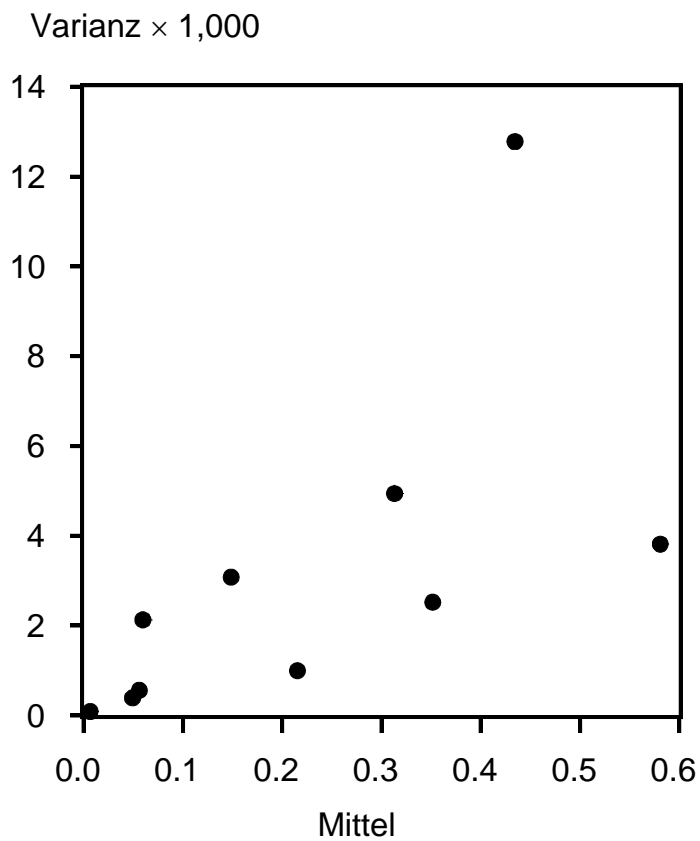
$(\alpha r)_{ij}$  = Wechselwirkung Sorte  $\times$  Experiment

$c_{jk}$  = Effekt des  $k$ -ten Kontainers im  $j$ -ten Experiment

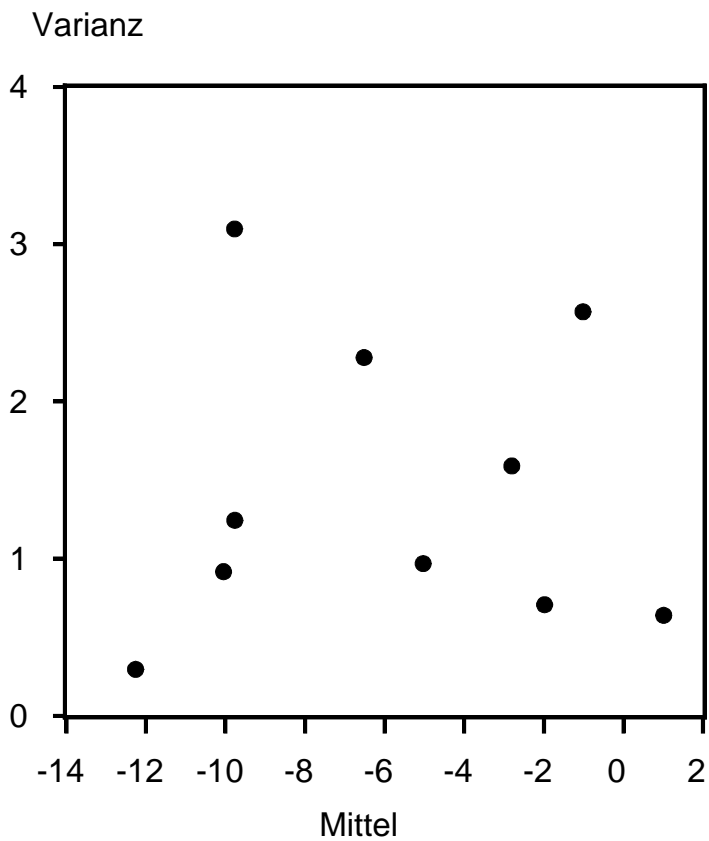
$e_{ijkh}$  = Restfehler (Petrischaleneffekt)

Alle Effekte sind zufällig außer dem Haupteffekt der Sorten ( $\alpha_i$ ) und dem allgemeinen Mittel ( $\mu$ ).

Die Analyse nach einem gemischten Modell basiert auf verschiedenen Annahmen wie Additivität der Effekte, Varianzhomogenität und Normalverteilung. Diese Voraussetzungen sind insbesondere bei Prozentzahlen und Zählwerten oft verletzt. Insbesondere ist die Annahme homogener Varianzen fragwürdig. Das obige Modell würde so angepasst, dass die Restvarianz für den Fehlerterm ( $e$ ) von der Sorte abhängt. Ein Plot der Varianz gegen den Mittelwert zeigt eine klare Varianz-Mittelwert-Beziehung (Abb. 6.1).



**Abb. 6.1:** Varianz-Mittelwert Plot für untransformierte Dormanz-Daten. Varianzen mit REML, Mittelwerte mit LSMEANS. Varianz-Heterogenität signifikant ( $p = 0,0018$ ; LR-Test).



**Abb. 6.2:** Varianz-Mittelwert Plot für transformierte Dormanz-Daten (Gefaltete Exponential-Transformation). Varianzen mit REML, Mittelwerte mit LSMEANS. Varianz-Heterogenität nicht signifikant ( $p = 0,1053$ ; LR-Test) .



Ein sehr gängiges Vorgehen bei Verletzung der Voraussetzungen ist die Transformation der Daten. Im vorliegenden Fall führt eine solche Transformation zur Brechung der Varianz-Mittelwert-Beziehung. Die numerischen Unterschiede der Varianzen sind deutlich reduziert und nicht mehr signifikant (Abb. 6.2). Ähnliche Ergebnisse findet man bezüglich der Normalverteilung der Residuen (nicht gezeigt).

Es stellt sich generell die Frage, welche der verschiedenen möglichen Transformationen zu bevorzugen ist. Besonders praktisch ist es hierbei, eine Familie von Transformationen zu verwenden, die je nach Wahl eines Parameters verschiedene Transformationen zulässt und dann die optimale Transformation aus den Daten zu schätzen. Die bekannteste ist die Box-Cox-Familie von Transformationen.

$$y = \begin{cases} (x^\phi - 1)/\phi, & \phi \neq 0 \\ \log_e(x) & \phi = 0 \end{cases}$$

Diese Transformation hat einen Transformationsparameter  $\phi$ , der aus den Daten geschätzt werden kann. Die Box-Cox Transformation umfasst einige bekannte Transformationen, so z.B.

$$\begin{aligned} y &= x && \text{für } \phi = 1 \\ y &= \sqrt{x} && \text{für } \phi = 0,5 \\ y &= \log(x) && \text{für } \phi = 0 \\ y &= \frac{1}{x} && \text{für } \phi = -1 \end{aligned}$$

Man beachte, dass auch der Fall enthalten ist, dass nicht transformiert wird ( $\phi = 1$ ). Der große Vorteil einer Transformation mit Parameter ist, dass die optimale unter einer Reihe von Transformationen anhand der Schätzung des Transformationsparameters bestimmt werden kann. Man muss also nicht verschiedene Transformationen ausprobieren, sondern man bekommt die optimale in einem Schritt ausgegeben.

Die Box-Cox-Transformation ist für Prozentzahlen allerdings oft keine günstige Transformation zur Erzielung der Voraussetzungen. Sie hat zusätzlich das Problem, dass sie für Prozentzahlen  $x = 0$  generell nicht funktioniert (jedenfalls was das Schätzen des Transformationsparameters anbelangt).

Für Prozentzahlen gibt es eine Reihe von alternativen Transformationen (Piepho, 2003b).

Transformation	Bezeichnung
$y = \log\left(\frac{x}{1-x}\right)$	Logit
$y = x^\phi - (1-x)^\phi$	Gefaltete Power
$y = \left(\frac{x}{1-x}\right)^\phi - 1$	Guerrero-Johnson
$y = \frac{x^\phi - (1-x)^\phi}{x^\phi + (1-x)^\phi}$	Aranda-Ordaz
$y = \sin^{-1} \sqrt{x}$	Winkel-Transformation

Von den obigen Transformationen kann jedoch nur die Winkel-Transformation angewendet werden, wenn  $x = 0$  oder  $x = 1$  auftritt. Diese ist nur für binomialverteilte Daten ohne Überdispersion optimal, und sie hat keinen Transformationsparameter wie die anderen Transformationen, so dass sie nicht dieselbe Flexibilität aufweist.

Mehr Flexibilität bietet die Gefaltete Exponential-Transformation (Piepho, 2003b)

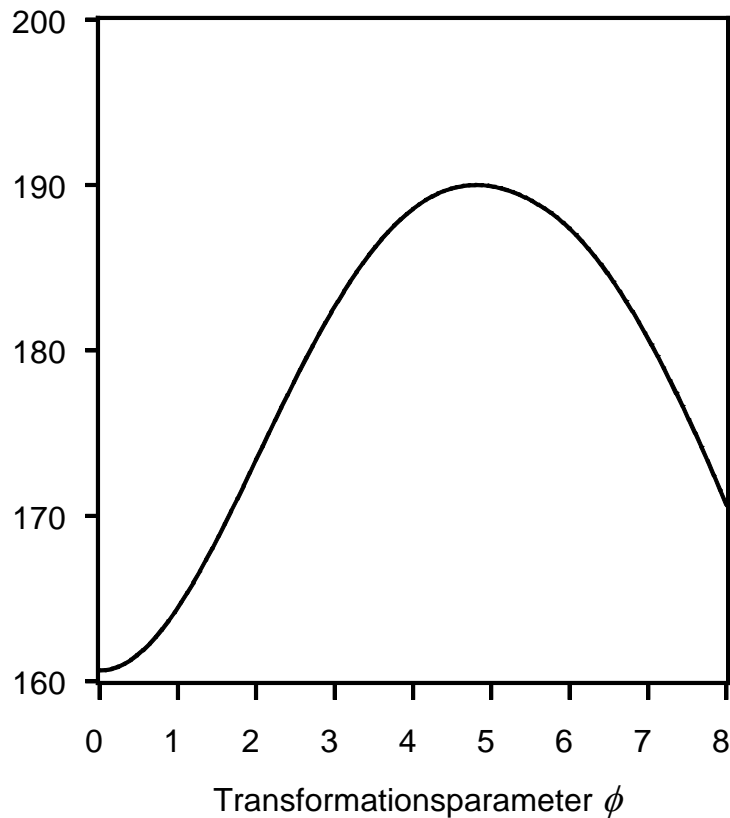
$$y = \begin{cases} \frac{\exp(\phi x) - \exp[\phi(1-x)]}{2\phi} & \phi \neq 0 \\ x & \phi = 0 \end{cases}$$

Diese hat gegenüber anderen Transformationen mit Parameter den Vorteil, dass sie auch bei  $x = 0$  oder  $x = 1$  anwendbar bleibt.

Die Schätzung des Transformationsparameters beruht auf der sog. Likelihood-Funktion. Diese gibt die Wahrscheinlichkeit für die beobachteten Daten in Abhängigkeit von den Parametern an. Man berechnet die Likelihood für verschiedene Werte des Transformationsparameters. Derjenige Wert, der die Likelihood maximiert, ist der optimale Wert für den Transformationsparameter, und zwar in dem Sinne, dass die Wahrscheinlichkeit der beobachteten Daten maximiert wird. In der Praxis maximiert man nicht die Likelihood, sondern deren Logarithmus, die sog. Log-Likelihood, da dieses leichter ist. Das Log-Likelihood-Profil für die Dormanz-Daten ist in Abb. 6.3 gezeigt. Das Maximum liegt bei  $\hat{\phi} = 4,81$ .

Anhand der Log-Likelihood kann man auch prüfen, ob der Transformationsparameter signifikant verschieden von Null ist, ob also die Notwendigkeit zu einer Transformation angezeigt ist. Dieser Test ist hier signifikant ( $\text{LR-}\chi^2 = 58,72, p < 0,0001$ ), also ist die Transformation angezeigt.

Log-Likelihood



**Abb. 6.3:** Log-Likelihood Profil zur Bestimmung des Transformationsparameters der Gefalteten Exponential-Transformation bei den Dormanz-Daten. Maximum bei  $\hat{\phi} = 4,81$ .

Mittelwertvergleiche sind in Tab. 6.2 gezeigt. Die Dormanz unterscheidet sich für einige Paare von ursprünglicher Sorte und nah-isogener Sorte. Der Vergleich von 4b und 4c weist jedoch darauf hin, dass diese Unterschiede stark von der Saatgutvermehrungsumwelt abhängen können. Dies legt ein Nachfolgeexperiment nahe, in dem verschiedene Saatgutvermehrungsumwelten geprüft werden.

**Tab. 6.1:** Zahl dormanter Samen ( $m$ )/Gesamtzahl von Samen ( $n$ ) in Dormanzexperiment mit Raps (Daten von S. Gruber, Universität Hohenheim); Anteil dormanter Samen:  $x = m/n$

Experiment	Container	Sorte						Container	Sorte			
		1a	1b	2a	2b	3a	3b		4a	4b	4c	5
1	1	21/93	64/88	3/98	12/96	5/95	10/97	2	52/95	36/86	35/92	63/91
1	1	25/95	68/93	0/99	9/98	7/99	10/98	2	36/93	44/94	36/96	67/89
1	1	16/94	51/86	0/96	11/99	14/98	5/93	2	50/92	46/90	39/97	65/88
1	1	35/98	55/85	2/100	8/98	21/98	9/96	2	40/93	37/87	44/99	71/88
2	4	8/99	56/94	1/99	4/99	1/10	5/97	3	15/94	30/90	7/99	17/92
2	4	4/96	45/99	1/100	2/97	3/97	4/95	3	27/90	26/91	14/96	--/44
2	4	13/96	53/96	0/94	3/99	5/97	3/96	3	35/95	37/96	12/99	28/94
2	4	15/94	45/95	1/96	5/94	0/96	2/94	3	21/91	37/94	7/97	39/97
3	5	10/99	56/95	0/100	2/100	7/98	5/96	6	20/92	23/88	14/96	15/95
3	5	7/98	46/87	0/98	8/100	2/93	0/90	6	13/90	23/81	11/100	22/90
3	5	9/95	56/94	0/98	1/99	1/96	2/95	6	20/92	15/86	16/97	22/95
3	5	8/96	43/89	0/100	1/97	4/97	2/93	6	20/94	26/93	15/98	33/93

**Tab. 6.2:** Generalisierte Kleinstquadrat-Mittelwerte (LSMEANS) für Dormanz-Daten nach Gefalteter Exponential-Transformation sowie Roh-Mittelwerte (untransformierte Daten). Paarweiser t-Test nach Kenward und Roger (1997).

Gruppe	Sorte <sup>§</sup>	Mittelwerte	
		Transformiert	Untransformiert
1	1a	-6,52 A	0,148
	1b	1,01 B	0,581
2	2a	-12,26 A	0,007
	2b	-9,76 B	0,056
3	3a	-9,77 A	0,060
	3b	-10,05 A	0,049
4	4a	-2,80 A	0,313
	4b	-1,98 A	0,352
	4c	-5,03 B	0,216
5 (Kontrolle)	5	-1,02	0,431

§ a: genetisch modifiziert; b, c: nah-isogene Sorte. Die Sorten a, b und c einer Gruppe stammen jeweils aus verschiedenen Vermehrungsumwelten. Sorten einer Gruppe, die einen Buchstaben gemeinsam haben, sind nicht signifikant verschieden.

SAS-Anweisungen:

```
data d;
input
vers pos dg sorte dorma leba;
x=dorma/leba;
datalines;
1 . 1 1 21 93
1 . 1 1 25 95
1 . 1 1 16 94
1 . 1 1 35 98
2 5 4 1 8 99
2 1 4 1 4 96
2 2 4 1 13 96
2 4 4 1 15 94
3 1 5 1 10 99
3 2 5 1 7 98
3 4 5 1 9 95
3 5 5 1 8 96
1 . 1 2 64 88
1 . 1 2 68 93
1 . 1 2 51 86
1 . 1 2 55 85
2 1 4 2 56 94
2 2 4 2 45 99
2 4 4 2 53 96
```

<b>vers:</b>	Experiment
<b>dg:</b>	Kontainer
<b>dorma:</b>	Zahl dormanter Samen
<b>leba:</b>	Gesamtzahl von Samen

2	5	4	2	45	95
3	1	5	2	56	95
3	2	5	2	46	87
3	4	5	2	56	94
3	5	5	2	43	89
1	.	1	3	3	98
1	.	1	3	0	99
1	.	1	3	0	96
1	.	1	3	2	100
2	3	4	3	1	99
2	1	4	3	1	100
2	6	4	3	0	94
2	4	4	3	1	96
3	1	5	3	0	100
3	2	5	3	0	98
3	5	5	3	0	98
3	6	5	3	0	100
1	.	1	4	12	96
1	.	1	4	9	98
1	.	1	4	11	99
1	.	1	4	8	98
2	3	4	4	4	99
2	3	4	4	2	97
2	2	4	4	3	99
2	6	4	4	5	94
3	2	5	4	2	100
3	3	5	4	8	100
3	5	5	4	1	99
3	6	5	4	1	97
1	.	1	5	5	95
1	.	1	5	7	99
1	.	1	5	14	98
1	.	1	5	21	98
2	2	4	5	1	100
2	6	4	5	3	97
2	5	4	5	5	97
2	3	4	5	0	96
3	1	5	5	7	98
3	3	5	5	2	93
3	4	5	5	1	96
3	6	5	5	4	97
1	.	1	6	10	97
1	.	1	6	10	98
1	.	1	6	5	93
1	.	1	6	9	96
2	6	4	6	5	97
2	4	4	6	4	95
2	5	4	6	3	96
2	1	4	6	2	94
3	3	5	6	5	96
3	3	5	6	0	90
3	4	5	6	2	95
3	6	5	6	2	93
1	1	2	7	52	95
1	2	2	7	36	93
1	3	2	7	50	92
1	4	2	7	40	93
2	1	3	7	15	94
2	2	3	7	27	90
2	3	3	7	35	95
2	4	3	7	21	91

```

3      1      6      7      20      92
3      2      6      7      13      90
3      3      6      7      20      92
3      4      6      7      20      94
1      1      2      8      36      86
1      2      2      8      44      94
1      3      2      8      46      90
1      4      2      8      37      87
2      1      3      8      30      90
2      2      3      8      26      91
2      3      3      8      37      96
2      4      3      8      37      94
3      1      6      8      23      88
3      2      6      8      23      81
3      3      6      8      15      86
3      4      6      8      26      93
1      1      2      9      35      92
1      2      2      9      36      96
1      3      2      9      39      97
1      4      2      9      44      99
2      1      3      9      7      99
2      2      3      9      14      96
2      3      3      9      12      99
2      4      3      9      7      97
3      1      6      9      14      96
3      2      6      9      11      100
3      3      6      9      16      97
3      4      6      9      15      98
1      1      2      10     63      91
1      2      2      10     67      89
1      3      2      10     65      88
1      4      2      10     71      88
2      1      3      10     17      92
2      2      3      10     .      44
2      3      3      10     28      94
2      4      3      10     39      97
3      1      6      10     15      95
3      2      6      10     22      90
3      3      6      10     22      95
3      4      6      10     33      93
;

```

```

%foldedexponential(phimin=0, phimax=8, steps=100, class=vers dg sorte,
model=sorte, stmts=%str(random vers vers*dg vers*sorte;), data=d, response=x);

```

## Beispiel 7: Räumliche Statistik (Nearest-Neighbor-Adjusted BLUP)

Im Rahmen des GABI-Projekts "Genomanalyse der Kühletoleranz" wurde eine große Population von Maislinien in 19 Umwelten in einem sog. Augmented Design geprüft (Dr. T. Presterl, Institut für Pflanzenzüchtung, Universität Hohenheim; Dr. M. Ouzunova, KWS SAAT AG, Einbeck). Hier wird aus der Versuchsserie nur der Ort Arras im Jahr 2001 verwendet. Jeder Versuch war in 9 Blöcke aufgeteilt, wobei in einem Block 90 Linien und 10 Standardsorten angebaut wurden. Somit liegen für die Standards 9 Wiederholungen vor, während die Linien unwiederholt waren. Dieses Design ist ein sog. "augmented design". Die Linien werden für die Auswertung als zufällig betrachtet, da sie aus einer Ursprungskreuzung hervorgegangen sind. Die zufällige Spaltung rechtfertigt die Annahme zufälliger Effekte. Wenn außerdem, wie im vorliegenden Fall, der interessierende Faktor viele Stufen hat (hier 720), so führt die Annahme zufälliger Effekte in der Regel zu besseren Schätzungen als bei Annahme fester Effekte (Searle et al., 1992).

Linieneffekte werden mit der gängigen Schätzmethode für zufällige Effekte, BLUP = best linear unbiased prediction, im Rahmen eines gemischten Modells geschätzt. Die BLUP-Schätzung hat einen geringeren Vorhersagefehler als die beobachteten Linienwerte, vorausgesetzt die Modellannahmen stimmen. Die Effekte für die Standards werden als fest betrachtet.

Wegen der Größe des Versuchs und der unvollständigen Blöcke (100 Versuchsglieder je Block) bietet sich eine geostatistische Auswertung an, um eine Korrektur für die Bodenheterogenität durchführen zu können. Die Korrektur beruht darauf, dass die Ertragsfähigkeit einer Parzelle über die Nachbarparzellen abgeschätzt wird (Nächst-Nachbar-Analyse). Die notwendigen Berechnungen lassen sich sehr einfach über die Anpassung eines gemischten Modells implementieren. Hierzu wurden die Beobachtungen so modelliert, dass Korrelationen zwischen benachbarten Parzellen zugelassen werden, wobei die Korrelation um so enger ist, je näher die Parzellen beieinander liegen. Die Nächst-Nachbar-Korrektur für Bodenheterogenität ergibt sich dann automatisch mit dem BLUP-Verfahren. Die Korrelationsstruktur fungiert hierbei gewissermaßen als Abstandsmaß, so dass eng benachbarte Parzellen einen stärkeren Einfluss auf die Korrektur eines Messwertes haben als weiter entfernte. Abschließend lässt sich die Zusammenfassung der Serie im Rahmen eines gemischten Modells leicht bewerkstelligen (Gilmour et al., 1997), worauf im Rahmen dieser Einführung jedoch nicht eingegangen wird.

Bei der Modellierung ergeben sich verschiedene Optionen. Zum einen ist zu entscheiden, ob der Randomisationsstruktur entsprechende Blockeffekte in das Modell aufgenommen werden sollen. Solche Effekte können aber bei einer geostatistischen Auswertung auch weggelassen werden, wobei dann die durch Blockbildung induzierte Korrelation von Beobachtungen durch die räumliche Korrelationsstruktur modelliert wird. Letzterer Ansatz wird hier verfolgt und der Auswertung ohne geostatistische Korrelationsstruktur gegenübergestellt.

Der Datensatz hat sechs Standards, die von PGL=1001 bis PGL=1006 kodiert sind. Die Linien sind durchlaufend beginnend mit PGL=1 nummeriert. Um für die Standards feste Effekte anzupassen, wird eine Variable GEN eingeführt, welche die



Stufen GEN=1000 für jede Linie und GEN=1001 bis GEN=1006 für die Standards hat.

PGL	GEN
1	1000
2	1000
3	1000
4	1000
.	.
.	.
.	.
.	.
1001	1001
1002	1002
1003	1003
1004	1004
1005	1005
1006	1006

Der feste Teil des Modells hat dann die Form

**MODEL g\_m2=GEN;**

Wobei  $g\_m2$  der Ertrag in g pro  $m^2$  ist. Hierdurch wird für jeden Standard sowie für die Gruppe der Linien, also für jede Stufe von GEN, ein eigener Erwartungswert angepasst. Es gibt somit sieben verschiedene Erwartungswerte.

Die Effekte der Linien werden als zufällig betrachtet, während für die Standards keine zufälligen Effekte angepasst werden. Hierzu wird eine Dummy-Variable  $x$  definiert, die den Wert  $z=0$  für Standards und  $z=1$  für die Linien hat. Der zufällige genetische Effekt wird dann wie folgt definiert:

**RANDOM Z\*PGL;**

Durch die Dummy-Variable wird die genetische Varianz für die Standards auf Null gesetzt, während die Linien eine von Null verschiedene Varianz haben. Der genetische Wert einer Linie setzt sich zusammen aus dem Erwartungswert aller Linien ( $\mu$ ) zuzüglich eines zufälligen Effektes der Linie:

$$g_i = \mu + a_i$$

wobei

$g_i$  = Genetischer Wert

$\mu$  = Erwartungswert über die Linien

$a_i$  = Abweichung der  $i$ -ten Linie von Erwartungswert = genetischer Effekt

Die zufälligen Effekte der Linien ( $a_i$ ) haben einen Erwartungswert von Null. Ziel der Auswertung ist es, BLUPs für die zufälligen genetischen Effekte ( $a_i$ ) bzw. der genetischen Werte ( $g_i$ ) zu erhalten. Beide unterscheiden sich nur um einen Schätzwert für die Konstante ( $\mu$ ), so dass es egal ist, welcher der beiden geschätzt wird.

Zur BLUP-Schätzung im Rahmen eines gemischten Modells muss noch die Blockstruktur des Modells spezifiziert werden. Hierzu können wir zum einen ein der Randomisationsstruktur entsprechendes Modell aufstellen. Wenn die Blöcke des augmented design mit BLOCK und die Parzellen mit PARZ kodiert sind, lautet das Modell in der bei Piepho et al. (2003a) beschriebenen Syntax:

BLOCK/PARZ = BLOCK + BLOCK•PARZ

Nun wurde in einem der neun Blöcke des augmented design (je 100 Genotypen) ein Zweisatzgitter verwendet, um die Fehlerstreuung der Versuchsfläche näher zu untersuchen. Die vollständige Wiederholung dieses Designs wird mit WH kodiert, während unvollständige Blöcke innerhalb einer Wiederholung mit UB bezeichnet werden. Die Erweiterung um diese Faktoren führt zu folgendem Modell (Notation von Piepho et al., 2003a):

BLOCK/WH/UB/PARZ =  
BLOCK + BLOCK•WH : BLOCK•WH•UB + BLOCK•WH•UB•PARZ

Die entsprechenden MIXED Anweisungen für die Berechnung von BLUPs für die Linien (PGL) sind:

```
proc mixed;
class pgl wh gen block ub;
model g_m2=gen block*wh;
random block*wh*ub x*pgl/solution;
run;
```

Falls die Linien als fest betrachtet werden und LS-Means berechnet werden sollen, sind die Anweisungen:

```
proc mixed;
class pgl wh block wh un;
model g_m2=pgl block*wh;
random block*wh*un;
lsmeans pgl;
run;
```

Die alternative Auswertung besteht darin, alle Blockeffekte aus dem Modell fallen zu lassen und stattdessen ein geostatistisches Modell für einen einzigen Fehlerterm anzupassen, z.B. ein Power-Modell der Form

$$\text{Cov}(e_i, e_j) = \sigma^2 \rho^{d_{ij}}$$

wobei

$e_i$  = Fehler auf  $i$ -ter Parzelle

$e_j$  = Fehler auf  $j$ -ter Parzelle

$d_{ij}$  = Euklidische Distanz zwischen zwei Parzellen  $i$  und  $j$

$\sigma^2$  = Varianz

$\rho$  = Korrelationsparameter

Man beachte, dass dieses Korrelationsmodell dasselbe ist, wie es auch bei den Zeitreihendaten in den Beispielen 3 und 4 verwendet wurde. Für den Versuch liegt eine Nummerierung der Zeilen und Spalten vor (ROW, COL), die sich an der räumlichen Reihenfolge der Parzellen orientiert. Daher kann man eine Euklidische Distanz anhand der Zeilen- und Spaltennummern berechnen. Allerdings ist dies nur bei quadratischer Parzellenform eine realistische Distanz. Bei länglichen Parzellen muss man die Nummerierung durch metrische Abstände ersetzen.

Im vorliegenden Fall bestand eine Parzelle aus zwei Maisreihen, die jeweils 0,8 m Abstand hatten (entlang ROW-Richtung). Die Länge der Parzelle war 7 m, einschließlich eines Weges von 0,9 m (COL-Richtung). Nimmt man die Parzellenmitten als Bezugspunkt, so sind zwei Parzellen in ROW-Richtung 1,6 m entfernt, während es in COL-Richtung 7 m sind. Angenommen, die Zeilen- und Spaltennummern sind mit ROW und COL kodiert, so können metrische Koordinaten wie folgt erhalten werden:

$$X = \text{ROW} * 1.6;$$

$$Y = \text{COL} * 7.0;$$

Die Euklidische Distanz zwischen zwei Parzellen  $i$  und  $j$  ist dann gegeben durch (Satz des Pythagoras!)

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

Zur Anpassung sind folgende SAS-Anweisungen notwendig:

```
proc mixed;
class pgl gen;
model g_m2=gen;
random z*pgl/solution;
repeated/sub=int type=sp(pow) (x y);
run;
```

Hieraus ergeben sich BLUPs, die eine Korrektur für Fruchtbarkeitsunterschiede enthalten. Die Korrektur erfolgt vor allem über benachbarte Parzellen, weswegen das Verfahren als NNA-BLUP bezeichnet wird (nearest neighbor adjusted BLUP; Stroup und Mulitze, 1991). Die Korrektur wird über die Korrelationsstruktur vermittelt: je enger die Korrelation zu einer Parzelle, um so höher ist der Einfluss dieser Parzelle auf den BLUP-Wert einer DH-Linie.

Da die Parzellen hier länglich waren, wird zusätzlich ein anisotropes Power-Modell angepasst:

$$\text{Cov}(e_i, e_j) = \sigma^2 \rho_1^{d_{1ij}} \rho_2^{d_{2ij}}$$

wobei

$\rho_1, \rho_2$  = Korrelationsparameter für die beiden Dimensionen (ROW, COL)  
 $d_{1ij}, d_{2ij}$  = Distanzen in den beiden Dimensionen (X, Y)

$$d_{1ij} = x_i - x_j$$

$$d_{2ij} = y_i - y_j$$

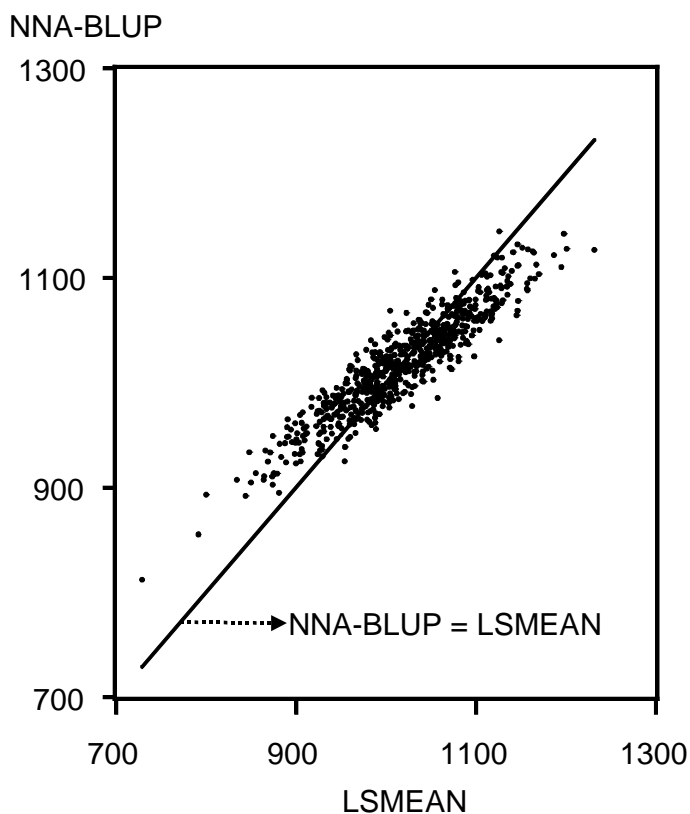
In der REPEATED-Anweisung muss hierfür die Option TYPE=SP(POWA) verwendet werden. Dieses Modell passte jedoch nicht besser als das einfache (isotrope) Power-Modell, so dass letzteres für die Berechnung von NNA-BLUP verwendet wird.

Nun zu den Ergebnissen. Die Analyse nach dem Modell mit festen Linieneffekten liefert folgende Varianzkomponentenschätzungen:

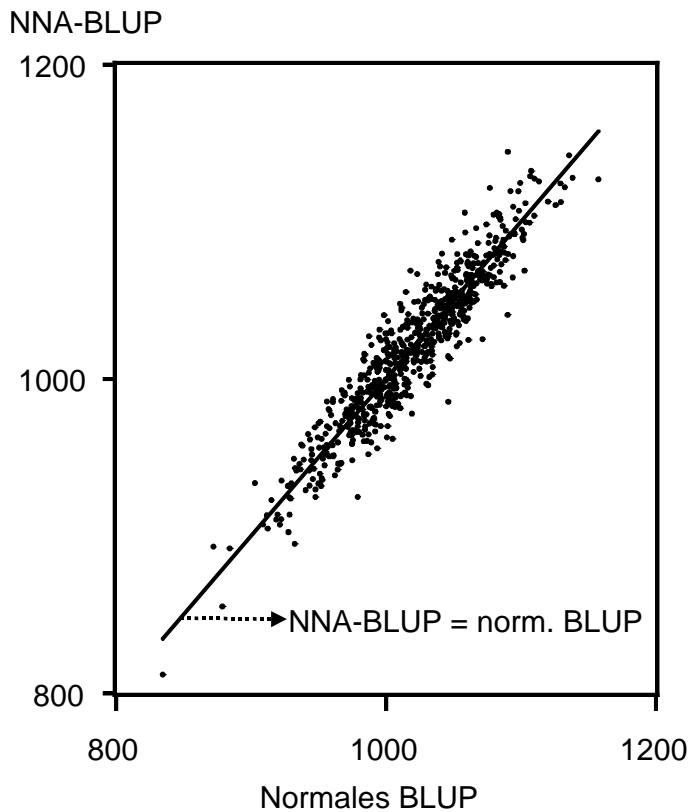
Unvollständige Blöcke im Gitter: 122  
Rest: 1800

Dies Ergebnis deutet darauf hin, dass die Fehlerstreuung sehr kleinräumig strukturiert ist, so dass hier ein relativ großes Potential für eine geostatistische Adjustierung (NNA-BLUP) besteht.

In Abb. 7.1 sind die NNA-BLUPs gegen die LS-Means abgetragen. Durch Vergleich mit der Winkelhalbierenden ( $NNA-BLUP = LS-Means$ ) sieht man, dass die BLUPs relativ zu den LS-means zum Gesamtmittel hin "geschrumpft" sind. Die Schrumpfung ist ein typisches Merkmal für BLUP-Schätzungen. Sie fällt umso stärker aus, je größer der Anteil der nichtgenetischen Streuung an der Gesamtstreuung ist. Diese Eigenschaft ist als erstrebenswert anzusehen, weil sie das widerspiegelt bzw. vorwegnimmt, was der Züchter praktisch oft erlebt, dass nämlich die Nachkommen gegenüber den Eltern zum Mittel zurückfallen. Man erhält somit realistischere Einschätzungen des tatsächlichen genetischen Wertes der Linien.



**Abb. 7.1:** NNA-BLUP (Power, isotrop) versus LSMEAN für GABI-Daten.



**Abb. 7.2:** NNA-BLUP (Power, isotrop) versus normales BLUP (keine Geostatistik) für GABI-Daten.

**Tab. 7.1:** Rangkorrelationen zwischen verschiedenen Schätzverfahren.

Schätzverfahren	Schätzverfahren		
	(1)	(2)	(3)
(1) LSMEANS (BLUE)	1	0.99	0.92
(2) Normales BLUP		1	0.93
(3) NNA-BLUP (Power, isotrop)			1

Die Rangkorrelationen der Schätzungen mit verschiedenen Methoden in Tab. 7.1 zeigen, dass es nicht unerhebliche Unterschiede gibt. Während das normale, auf der Randomisationsstruktur basierende BLUP-Verfahren den LS-Means sehr ähnlich ist, sind beide Verfahren deutlich von NNA-BLUP verschieden, da die Rangkorrelation deutlich von 1 abweicht (siehe auch Abb. 7.2). Es ist hier anhand der Korrelationen allerdings nicht ohne weiteres zu entscheiden, welches Schätzverfahren das bessere ist. Viele Arbeiten zeigen jedoch, dass BLUP oft besser ist als LSMEANS. Aus theoretischer Sicht ist bekannt, dass BLUP die tatsächliche Rangfolge der genetischen Effekte generell besser wiedergibt als LS-Means. Ob die geostatistische Version von BLUP (NNA-BLUP) besser ist als das normale BLUP, ist ebenfalls nicht ohne weiteres klar. Aber auch hier bestehen merkliche Unterschiede. Vorteile sind insbesondere dann zu erwarten, wenn die Blöcke groß sind und viele Prüfglieder enthalten. Im vorliegenden Fall hat ein Block 100 Parzellen, sodass ein Gewinn zu

erwarten ist. In jedem Fall ist NNA-BLUP eine interessante Variante von BLUP, die in der Züchtung und in Genbanken zunehmend an Bedeutung gewinnt.

## Zusammenfassung

Dieser Workshop hat hoffentlich folgendes verdeutlicht:

- Die Auswertung gemischter Modelle mit MIXED ist sehr ähnlich der mit GLM. Es gibt nur ein paar kleine Änderungen im Output wegen des modifizierten Auswertungsverfahrens (REML).
- Richtige Standardfehler und Mittelwertvergleiche werden nur mit MIXED erhalten, nicht mit GLM.
- Gemischte Modelle sind sehr flexibel einsetzbar für die Modellierung in Nicht-Standard-Designs. Voraussetzung: Richtige Randomisation. (Statistiker fragen!)
- Gemischte Modelle sind wichtig für Auswertung von Daten mit zeitlicher und räumlicher Korrelation (Messwiederholungen, repeated measures, Geostatistik)
- Daten-Transformationen sind in gemischten Modellen leicht einsetzbar

**Fazit:** Gemischte Modelle gehören in Handwerkskasten des Pflanzenbauwissenschaftlers.

## Literatur

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

Piepho, H.P. (2003a): An algorithm for a letter-based representation of all-pairwise comparisons. *Journal of Computational and Graphical Statistics*

Piepho, H.P. (2003b): The folded exponential transformation for proportions. *The Statistician*

Piepho, H.P., Büchse, A., and Emrich, K. (2003a): A hitchhiker's guide to the mixed model analysis of randomized experiments. *Journal of Agronomy and Crop Science* **189**, 310-322.

Piepho, H. P., Büchse, A., Richter, C. (2003b): A mixed modelling approach to randomized experiments with repeated measures. *Journal of Agronomy and Crop Science*

Piepho, H.P. und Spilke, J. (1999): Anmerkungen zur Analyse balancierter gemischter Modelle mit der SAS Prozedur MIXED. *Zeitschrift für Agrar informatik* **7**, 39-46.

Schabenberger, O., Pierce, F.J. (2001): Contemporary statistical models for the plant and soil sciences. CRC Press, Boca Raton.

Stroup, W. W. , Mulitze, D. K. (1991): Nearest neighbor adjusted best linear unbiased prediction. *The American Statistician* **45**, 194-200.

Verbeke, G., Molenberghs, G. (1997): Linear mixed models in practice: A SAS-oriented approach. Springer, Berlin.

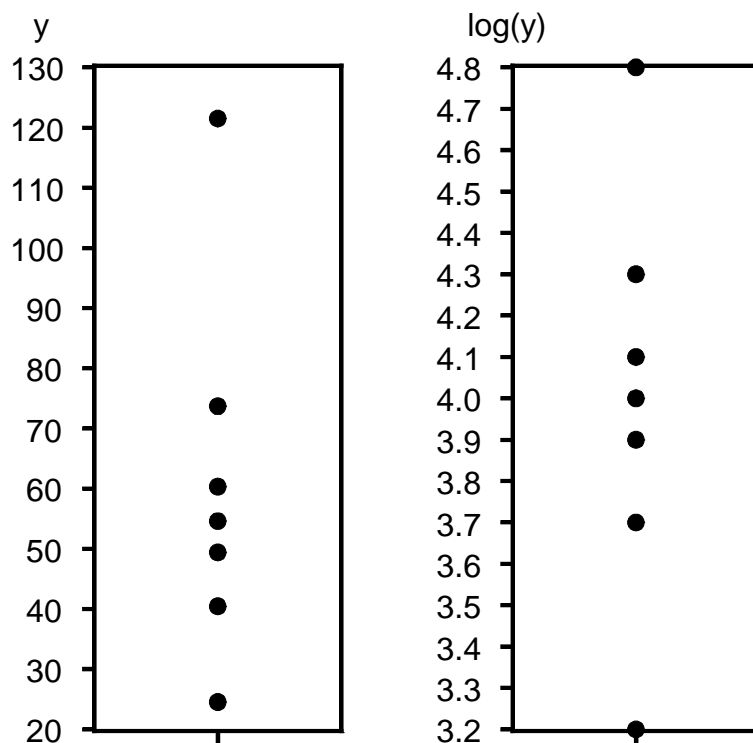
## Anhang: Rücktransformierte Mittelwerte (arithmetische Mittel) sind Mediane

Diese Aussage kann man am besten anhand einer einfachen Stichprobe und des einfachen Stichprobenmittelwertes bzw. Stichproben-Medians veranschaulichen. Analoges gilt für Schätzwerte, die auf einem linearen gemischten Modell beruhen.

**Stichprobenwerte:** Die folgenden sieben hypothetischen Stichprobenwerte sind absichtlich so gewählt, dass sich für die logarithmierten (transformierten) Werte runde Zahlen ergeben.

y	log(y)	
24.533	3.2	
40.447	3.7	
49.402	3.9	
54.598	4.0	
60.340	4.1	
73.700	4.3	
121.510	4.8	
60.647	4.0	arithmetisches Mittel
50.598	4.0	Median

Die Verteilung der transformierten Daten ist symmetrisch, während die untransformierten Daten schief verteilt sind (Abb. A.1).



**Abb. A.1:** Verteilung der transformierten und der untransformierten Daten.

Die schiefen Daten verletzen die Voraussetzung der Normalverteilung, weshalb eine Transformation angezeigt ist. Die transformierten Daten sind dagegen symmetrisch und einigermaßen normalverteilt (die Normalverteilung ist eine symmetrische Verteilung), so wie man es nach einer Transformation ja auch hofft. Wegen der Symmetrie der Verteilung gilt für die log-transformierten Daten, dass



Median = arithmetisches Mittel ( = 4)

Diese Gleichheit von Median und Mittel gilt offensichtlich nicht für die schief-verteilten Original-Daten (Median = 54.598, arith. Mittel = 60.647). Die Rücktransformation des Mittelwertes von der log-Skala ist:

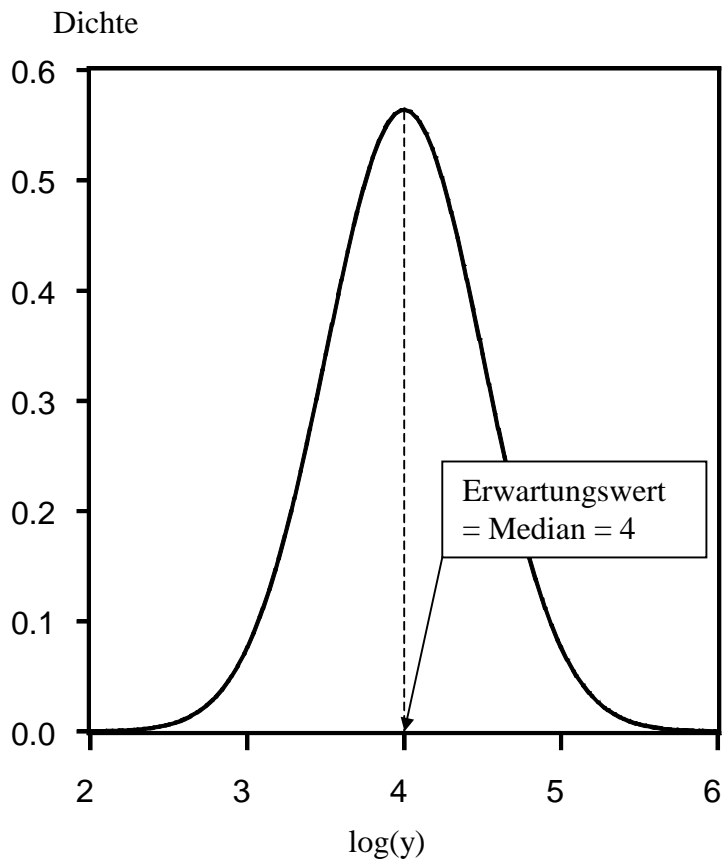
$\exp(\text{Mittelwert log-Skala}) = \exp(4) = 54.598 = \text{Median Original-Skala}$

Dies ist derselbe Wert, der sich durch Rücktransformation des Median ergibt, weil ja Median und arith. Mittel auf der log-Skala übereinstimmen! In diesem Beispiel ist also offensichtlich zutreffend, dass der rücktransformierte Mittelwert dem Median auf der Original-Skala entspricht.

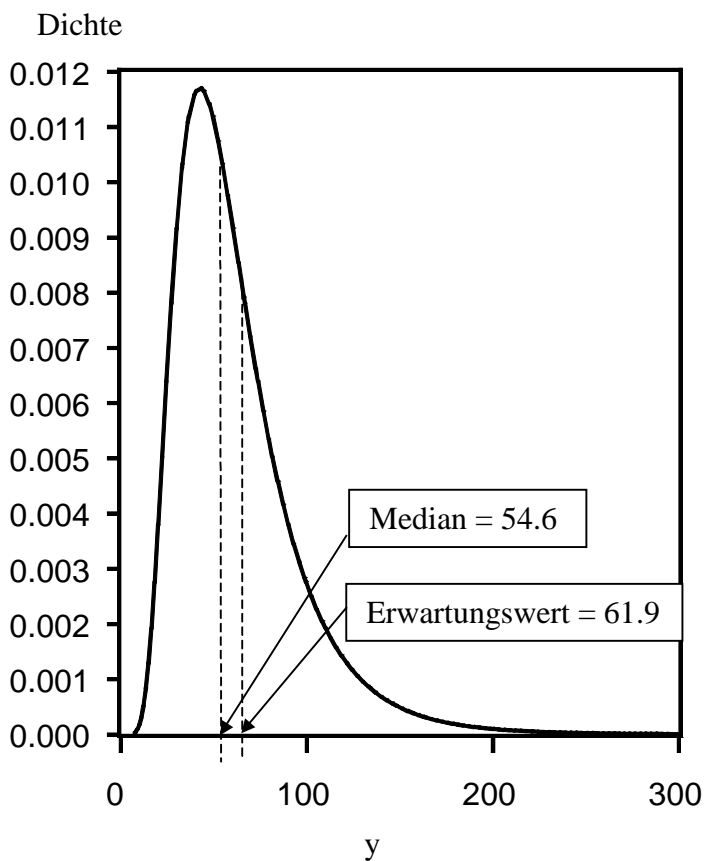
Der Median ist ein Wert, unter dem und über dem jeweils 50% der Daten liegen. Der rücktransformierte Median ist immer gleich dem Median auf der Original-Skala, weil der Median einem einzelnen der beobachteten Werte entspricht (bzw. einem Mittel aus zwei Werten). Diese Eigenschaft des Medians ist unbeeinträchtigt von Daten-Transformationen. Dagegen ist der rücktransformierte arithmetische Mittelwert in der Regel nicht gleich dem arithmetischen Mittel auf der Original-Skala.

**Theoretische Verteilungen:** In der Realität sind die transformierten Daten selbst bestenfalls annähernd symmetrisch verteilt, so dass die obigen Beziehungen nur näherungsweise zutreffen. Aber die Symmetrie gilt immer exakt für die angenommene theoretische Normalverteilung auf der log-Skala. Für diese gilt immer Median = Erwartungswert, wobei der Erwartungswert bei theoretischen Verteilungen das arithmetische Mittel ersetzt. Die rücktransformierte Verteilung ist bei dieser Annahme immer schief (eine sog. Lognormalverteilung), so dass auf der rücktransformierten Skala Median  $\neq$  Erwartungswert. Außerdem ist der rücktransformierte Erwartungswert (Mittelwert) immer gleich dem Median auf der Original-Skala. Dies ist in den Abbildungen A.2 und A.3 veranschaulicht für den Fall einer Normalverteilung der log-transformierten Daten mit Erwartungswert 4 und Standardabweichung 0,5, was etwa dem Beispiel der sieben Stichprobenwerte entspricht. Man sieht, dass Median < Erwartungswert für die Original-Skala, während Median = Erwartungswert für die log-Skala. Es nochmals betont, dass all diese Aussagen an die Gültigkeit der Normalverteilung auf der transformierten Skala gebunden sind. Es ist also für diese Interpretation wichtig, dass man tatsächlich eine Transformation findet, welche die Normalverteilung auf der transformierten Skala zumindest näherungsweise sicherstellt.

**Lineare Modelle:** Für Schätzwerte nach dem linearen Modell ergeben sich analoge Aussagen, wenn man von Normalverteilung auf der transformierten Skala ausgeht. Diese Annahme bedingt wiederum automatisch eine schiefe Verteilung auf der Original-Skala. Die Krux ist, dass wir bei linearen Modellen Schätzwerte der Erwartungswerte (Mittelwerte) leicht berechnen können. Bei Normalverteilung auf der transformierten Skala sind diese immer auch Schätzwerte der Mediane, so dass Rücktransformation auf jeden Fall zu Schätzwerten von Medianen führt, nicht aber zu Schätzwerten von Erwartungswerten.



**Abb. A.2:** Verteilung der log-transformierten Daten (Normalverteilung mit Erwartungswert = 4 und Standardabweichung = 0,5).



**Abb. A.3:** Verteilung der Originaldaten (Log-Normalverteilung).