

SAS/AF-Anwendung zur Planung und Auswertung ein- bis dreifaktorieller Feldversuchsanlagen FELD_VA II

Eckard Moll

Biologische Bundesanstalt für Land und Forstwirtschaft

Stahnsdorfer Damm 81

14532 Kleinmachnow

e.moll@bba.de

Zusammenfassung

Beispielhaft werden aus der SAS/AF-Anwendung mit SAS 9.1 FELD_VA II einige der Feldversuchsanlagen vorgestellt. Die Planung der Versuche umfasst die Konstruktion der Lagepläne und Genauigkeitsaussagen. Für die Lagepläne werden die Prüfglieder modellabhängig zufällig angeordnet. Die Genauigkeitsaussagen für alle ein- bis dreifaktoriellen vollständigen Versuchsanlagen beziehen sich wahlweise auf die Zielgrößen Anzahl der Wiederholungen bzw. Blocks, Mindestdifferenz, Risiko 1. Art und Risiko 2. Art unter Berücksichtigung der Testverfahren multipler t-Test (zweiseitig), Tukey-Test, ein- und zweiseitiger Dunnett-Test.

Die Versuchsauswertung orientiert noch auf die Auswertung der Einzelversuche. Als Testverfahren können zur Varianzanalyse der multiple t-Test, t-Tests zur Bezugsbasis (Bildung linearer Kontraste), Bonferroni-Test, Tukey-Test, ein- und zweiseitiger Dunnett-Test gewählt werden.

Schlüsselworte: SAS 9, Feldversuche, Planung und Auswertung

1 Wieso FELD_VA II?

Die ersten Arbeiten an einem Projekt zur Planung und Auswertung von Feldversuchen – im Wesentlichen vollständiger Blockanlagen - begannen 1993/94 unter SAS 6.03 bzw. SAS 6.04. Dass es dabei zu einer SAS-Anwendung unter einer gemeinsamen Oberfläche kam, ist vor allem den Konsultationen bei Herrn ETZOLD, dem damaligen Leiter der SAS-Geschäftsstelle Berlin, zu danken. Erst eine derartige Oberfläche gestattet es, verschiedene Programme in Abhängigkeit von Eingaben abzuarbeiten.

In SAS 6.08 war die Prozedur Proc MIXED neu. Sie hatte noch nicht die Leistung von heute, so dass die auf der Grundlage von Proc GLM begonnene Programmierung beibehalten wurde. Die bei einigen Versuchsanlagen erforderliche Wichtung der Varianzen passierte auf der Grundlage der gewichteten Verteilungsquantile – z.B. gewichtete t-Quantile.

FELD_VA (MOLL 1996, 1997) ist derzeit eine SAS/AF-Anwendung unter SAS 6.12 (Abb. 1). Eine Übernahme zu SAS 8 war ohne größere Anpassung nicht möglich. Aus diesem Grund wird SAS 6.12 noch für die Nutzung von FELD_VA benötigt.

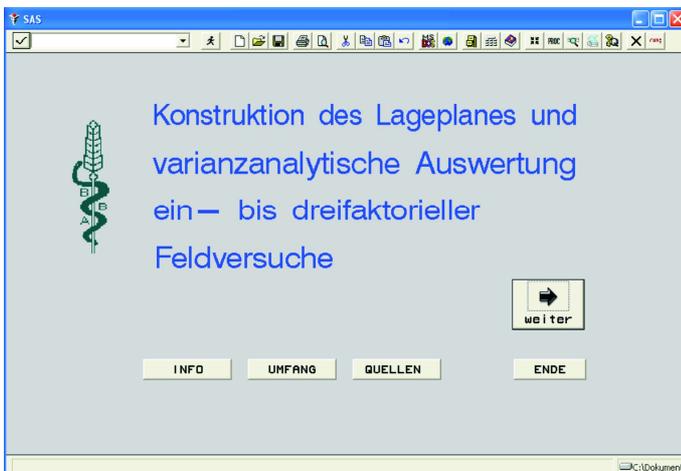


Abb. 1: FELD_VA, Eröffnungsbildschirm

Seit einigen Jahren ist es Standard, für die erforderliche Wichtung der Varianzen die gewichteten Freiheitsgrade (Methode von SATTERTHWAITE oder von KENWARD und ROGER) zu nutzen. Es gab und gibt also mehrere Gründe, FELD_VA zu überarbeiten.

Zu klären war noch die Frage, ob für die Konstruktion der Oberfläche SAS/AF oder SAS/WebAF genutzt werden soll. Da das neue FELD_VA nicht ins Web gestellt werden sollte, war die Entscheidung für SAS/AF ohne HTML-Formate einfach.

Die Konstruktion einer SAS/AF-Oberfläche ist leider nicht ganz so einfach, zumal ein Ansprechpartner wie damals bei SAS im Großraum Berlin fehlt. Unterstützung war (und ist) aber notwendig. Bedanken möchte ich mich für meine SAS-spezifischen Fragen bei Hans-Jürgen BAUMANN¹, Martin BENTNER¹, Michael BÖTTNER¹, Philipp ERHARDT¹, Karoline FOERSTER¹, Anke GRÖNING², Wolf F. LESENER³, Mat-

¹ SAS-Deutschland, Technical Consult, Heidelberg (SAS-Hotline)

² Data Warehouse Beratung, Berlin

³ Humboldt-Universität zu Berlin, Computer- und Medienservice

thias SCHACHNER⁴ und Claudia SCHOO⁵ und für die finanzielle Unterstützung durch die Biologische Bundesanstalt für Land und Forstwirtschaft. Einige der Hinweise und Lösungen werde ich hier erwähnen, weil sie auch für andere von Nutzen sein könnten.

2 FELD_VA II

2.1 Der Eröffnungsbildschirm

Für die grafische Oberfläche wird eine Bildschirmauflösung von 1024 x 768 Pixel zugrunde gelegt.



Abb. 2: FELD_VA II, Eröffnungsbildschirm

Der Eröffnungsbildschirm von FELD_VA II (Abb. 2) zeigt die Wahlmöglichkeiten:

⁴ SAS-Deutschland, Professional Services / SAS Training, Heidelberg

⁵ SAS-Deutschland, Professional Services / Consult, Hamburg

- Hinweise zur Benutzung,
- Beschreibung einer Versuchsanlage,
- Konstruktion eines randomisierten Lageplans,
- Berechnung von Genauigkeitsgrößen und
- Auswertung eines Einzelversuchs.

Die entsprechenden Frames und Programme sind überarbeitet oder wie die Berechnung von Genauigkeitsgrößen neu hinzu gekommen.

2.2 Beschreibung einer Versuchsanlage

Die bessere grafische Gestaltungsmöglichkeit unter SAS 9 im Vergleich zu SAS 6.12 gestattet es, alle wichtigen Informationen zu einer gewählten Versuchsanlage auf dem Bildschirm anzuzeigen. Die Versuchsanlagen sind:

einfaktorische randomisierte Anlage	A - R
einfaktorische randomisierte Blockanlage	A - Bl
zweifaktorielle randomisierte Anlage	(AxB) - R
zweifaktorielle randomisierte Blockanlage	(AxB) - Bl
zweifaktorielle Spaltanlage	(A/B) - Bl
zweifaktorielle Streifenanlage	(A+B) - Bl
dreifaktorielle randomisierte Anlage	(AxBxC) - R
dreifaktorielle randomisierte Blockanlage	(AxBxC) - Bl
dreifaktorielle Spaltanlage	(A/B/C) - Bl
dreifaktorielle zweistufige Spaltanlage	[(AxB)/C] - Bl
dreifaktorielle zweistufige Spaltanlage	[A/(BxC)] - Bl
dreifaktorielle zweistufige Streifenanlage	[A+(BxC)] - Bl
dreifaktorielle Streifen-Spaltanlage	[A+(B/C)] - Bl
dreifaktorielle Spalt-Streifenanlage	[(A+B)/C] - Bl
dreifaktorielle Spalt-Streifenanlage	[A/(B+C)] - Bl



Abb. 3: Versuchsanlagen

Für die ausgewählte Versuchsanlage (Abb. 3) werden die Bezeichnung, das Modell, eine verbale Beschreibung und ein Beispiel angegeben. Beispielhaft werden nachfolgend einige Versuchsanlagen vorgestellt (Abb. 4 a bis e).

einfaktorielle vollständig randomisierte Anlage A-R

Modell: $y_{ij} = \mu + \alpha_i + e_{ij}$

Beschreibung: Alle Prüfglieder, die Stufen des Faktors A, werden mit ihren Wiederholungen den Teilstücken zufällig zugeordnet. Unterschiede zwischen den Prüfgliedern werden mit gleicher Präzision geschätzt.

Beispiel: $a = 6, r = 4$

a)

2	1	1	3
1	6	5	1
2	3	5	4
3	4	5	6
5	4	6	3
6	2	2	4

Teilstücke A_4

b)

2	1	1	3	1	6	5	1
2	3	5	4	3	4	5	6
5	4	6	3	6	2	2	4

Teilstücke A_2

Abb. 4a: Informationen zur einfaktoriellen randomisierten Anlage A - R

einfaktorielle randomisierte Blockanlage A-BI

Modell: $y_{ijk} = \mu + \alpha_i + \beta_k + e_{ijk}$

Beschreibung: Alle Prüfglieder, die Stufen des Faktors A, werden pro Block den Teilstücken zufällig zugeordnet. Die Blockbildung gestattet das Erfassen und damit das Ausschalten systematischer Störgrößen wie beispielsweise systematische Änderungen eines exogenen Faktors (Boden) in einer Richtung. Unterschiede zwischen den Prüfgliedern werden mit gleicher Präzision geschätzt.

Beispiel: $a = 5, r = 6$

						Block
3	5	2	4	1	6	6
4	1	3	5	2	5	5
5	3	1	2	4	4	4
1	4	5	3	2	3	3
3	2	4	1	5	2	2
2	5	3	4	1	1	1

Teilstücke A_5

Abb. 4b: Informationen zur einfaktoriellen randomisierten Blockanlage A - BI

zweifaktorielle Spaltanlage (A/B)-Bl

Modell: $y_{gab} = \mu + \tau_k + \alpha_j + e_g + \beta_1 + (\alpha\beta)_{jg} + e_{gab}$

Beschreibung: Spaltanlagen werden angelegt, wenn technologisch bedingt für die Stufen eines Faktors größere Teilstücke notwendig sind. Die Stufen des Faktors A werden den Großteilstücken und die Stufen des Faktors B innerhalb der Großteilstücke den Kleinteilstücken zufällig pro Block zugeordnet. Folglich treten die beiden Fehlerkomponenten Fehler a der Großteilstücke und Fehler ab der Kleinteilstücke auf. Die Wirkung der beiden Faktoren A und B wird im Allgemeinen mit unterschiedlicher Präzision geschätzt: die Präzision von Faktor A (Großteilstücksfaktor) ist kleiner als die von Faktor B (Kleinteilstücksfaktor).

Beispiel: a = 4, b = 3, r = 6

												Block
32	33	31	11	13	12	45	41	42	22	21	23	6
13	11	12	32	33	31	22	21	23	41	43	42	5
23	21	22	43	42	41	11	13	12	32	31	33	4
42	41	43	11	13	12	31	32	33	23	22	21	3
31	33	32	22	23	21	41	43	42	12	11	13	2
43	42	41	31	33	32	23	22	21	13	12	11	1
Großteilstück A ₄						Kleinteilstück A ₂ B ₁						

Abb. 4c: Informationen zur zweifaktoriellen Spaltanlage (A/B) - Bl

dreifaktorielle zweistufige Streifenanlage [A+(BxC)]-Bl

Modell: $y_{gab} = \mu + \tau_k + \alpha_j + e_g + \beta_1 + \gamma_1 + (\beta\gamma)_{11} + e_{g1} + (\alpha\beta)_{jg} + (\alpha\gamma)_{j1} + (\alpha\beta\gamma)_{jg1} + e_{gab}$

Beschreibung: Diese Anlage ist eine Kombination aus einer einfaktoriellen (Faktor A) und einer zweifaktoriellen (Faktoren B und C) Blockanlage. Beide Anlagen werden über Kreuz in Streifen realisiert. Die drei Fehlerkomponenten Fehler a, Fehler bc und Fehler abc sind zu berücksichtigen.

Beispiel: a = 3, b = 2, c = 3, r = 4

													Block	
														Block
Großteilstück A ₁	1	323	313	322	312	321	311	112	122	111	123	113	121	2
		123	113	122	112	121	111	312	322	311	323	313	321	
		223	213	222	212	221	211	212	222	211	223	213	221	
Großteilstück B ₁ C ₂	3	211	212	223	213	221	222	122	121	112	111	113	123	4
		311	312	323	313	321	322	322	321	312	311	313	323	
		111	112	123	113	121	122	222	221	212	211	213	223	
Kleinteilstück A ₃ B ₂ C ₁														

Abb. 4d: Informationen zur dreifaktoriellen zweistufigen Streifenanlage [A+(BxC)] - Bl

dreifaktorielle Streifen-Spaltanlage [A+(B/C)]-Bl

Modell: $y_{gij} = \mu + \tau_i + \alpha_j + e_{ij} + \beta_1 + e_{a1} + \gamma_1 + (\beta\gamma)_{11} + e_{a1} + (\alpha\beta)_{1a} + e_{a1} + (\alpha\gamma)_{1a} + (\alpha\beta\gamma)_{a11} + e_{gij}$

Beschreibung: Streifen-Spaltanlagen werden häufig für mehrfaktorielle Versuche eingesetzt, wenn die technologischen Vorteile der Streifen- und der Spaltanlage genutzt werden sollen. Diese Anlage ist eine Kombination aus einer einfaktoriellen Blockanlage (Faktor A) und einer zweifaktoriellen Spaltanlage, die über Kreuz in Streifen zusammengeführt werden. Als Fehlerkomponenten treten auf: Fehler a, Fehler b, Fehler bc, Fehler ab und Fehler abc. Als Faktoren A und B sollten diejenigen gewählt werden, deren Stufen voraussichtlich die größten Differenzen hervorrufen. Im Allgemeinen wird die Wirkung des Faktors C mit höherer Präzision als die der Faktoren A und B geschätzt.

Beispiel: a = 3, b = 2, c = 3, r = 4

		Block						Block					
Großteilstück A ₁	1	322	321	323	311	313	312	223	221	222	213	211	212
		122	121	123	111	113	112	123	121	122	113	111	112
		222	221	223	211	213	212	323	321	322	313	311	312
Mittelteilstück A ₃ B ₂													
Großteilstück B ₂	3	322	321	323	311	313	312	223	221	222	213	211	212
		122	121	123	111	113	112	123	121	122	113	111	112
		222	221	223	211	213	212	323	321	322	313	311	312
Mittelteilstück B ₂ C ₃													
		Kleinteilstück A ₃ B ₂ C ₂											

Abb. 4e: Informationen zur dreifaktoriellen Streifen-Spaltanlage [A+(B/C)] - Bl

2.3 Konstruktion einer Versuchsanlage

Die Versuchsanlage, zu der ein randomisierter Lageplan erstellt werden soll (s.o., vgl. Abb. 3), wird gewählt. Der Anzahl der Faktoren entsprechend sind die Anzahl der Stufen der Faktoren und die Anzahl der Wiederholungen in entsprechende Masken einzugeben. Vorzugeben ist noch die Datei, in die der Lageplan geschrieben werden soll.

Veranschaulicht wird das an Hand einer einfaktoriellen randomisierten Blockanlage A – Bl (s. Abb. 4b). Beispielhaft werden gewählt:

- Anzahl Stufen des Faktors A: 5
- Anzahl Wiederholungen / Blocks: 6

Der Inhalt der Ausgabedatei ist:

Der nachfolgende Lageplan basiert auf den vorgegebenen Werten:

$$a = 5$$

$$\text{Blocks} = 6$$

Die erste Ziffer ist dem Faktor A und die zweite den Blocks zuzuordnen, d.h. eine Spalte entspricht einem Block.

Lageplan für eine einfaktorielle randomisierte Blockanlage A-B1
LAGEPLAN

51	12	53	34	25	26
31	42	23	44	55	16
41	32	43	24	45	56
11	52	13	14	15	46
21	22	33	54	35	36

Zu erkennen ist, dass innerhalb der Blocks die Prüfglieder von 1 bis 5 zufällig angeordnet sind.

Da das die erste hier vorgestellte Ausgabe von Ergebnissen auf eine Datei ist, wird darauf hingewiesen, dass die standardseitig von SAS ausgegebenen Seitenumbrüche sich beim Druck solcher Dateien als sehr lästig erweisen können. Die Option⁶

```
options formdlim=' ';
```

ersetzt den Seitenumbruch durch eine Leerzeile. Die Option `pagesize =` sollte hoch angesetzt werden.

2.4 Berechnen der Genauigkeitsgrößen

2.4.1 Genauigkeitsgrößen, Tests und Vergleiche von Mittelwerten

Die Versuchsanlage (s.o., vgl. Abb. 3) wird gewählt. In Anhängigkeit von der Zielgröße, dem Test, dem Vergleich der Mittelwerte und der Versuchsanlage sind einige Eingaben vorzunehmen.

Als Genauigkeitsgrößen (Zielgröße) können berechnet werden:

- r : Anzahl Wiederholungen / Blocks
- d : Mindestdifferenz
- α : Signifikanzniveau, Risiko 1. Art
- β : Risiko 2. Art .

Die Berechnung erfolgt für die Tests

- multipler t-Test
- Tukey-Prozedur
- Dunnett-Prozedur (einseitig)
- Dunnett-Prozedur (zweiseitig) .

⁶ eine der erwähnten Hilfen der SAS-Hotline

Die Berechnung der Genauigkeitsgrößen für die o.g. Tests kann für die Hauptwirkungen der Faktoren A, B oder C und zusätzlich für den multiplen t-Test und die Tukey-Prozedur im dreifaktoriellen Fall für die Kombinationswirkungen der AB-, AC- und BC-Mittelwerte vorgenommen werden. Nachfolgend wird die Berechnung am Beispiel der Mindestdifferenz für eine zweifaktorielle Spaltanlage (A/B) - Bl gezeigt.

2.4.2 Berechnung der Mindestdifferenz für eine zweifaktorielle Spaltanlage (A/B) - Bl

Die Freiheitsgrade zum Vergleich der A-Mittelwerte einer zweifaktoriellen Spaltanlage (A/B) - Bl (s. Abb. 4c) sind $FG = (a - 1) \cdot (r - 1)$. Die Mindestdifferenz d zum Vergleich von A-Mittelwerten (Mittelwerte des Großteilstücksfaktors) berechnet sich bekanntlich für

$$\text{t-Test:} \quad d = (t_{1-\alpha/2; FG} + t_{1-\beta; FG}) \cdot \sqrt{\frac{2 \cdot s_{\text{Fehler a}}^2}{b \cdot r}}$$

$$\text{Tukey-Prozedur:} \quad d = (q_{1-\alpha; FG, a} / \sqrt{2} + t_{1-\beta; FG}) \cdot \sqrt{\frac{2 \cdot s_{\text{Fehler a}}^2}{b \cdot r}} .$$

Analog würde das auch für die Dunnett-Prozedur gehen, wenn nicht die optimale Versuchsplanung für diesen Test besagt, dass der Standard im Allgemeinen häufiger zu wiederholen ist. Das berücksichtigen die Freiheitsgrade zum Vergleich von A-Mittelwerten: $FG = (a - 2) \cdot (r - 1) + r_0 - 1$, wobei r die Wiederholung der Prüfglieder und r_0 die Wiederholung des Standards A_0 ist. Es gilt: $r_0 \geq r$. Wird die unterschiedliche Wiederholung berücksichtigt, so berechnet sich die Mindestdifferenz d zum Vergleich von A-Mittelwerten einer zweifaktoriellen Spaltanlage (A/B) - Bl für die zweiseitige Dunnett-Prozedur

$$d = \left(|d|_{1-\alpha; FG, a-1, \lambda_1, \dots, \lambda_{a-1}} + t_{1-\beta; FG} \right) \cdot \sqrt{\frac{s_{\text{Fehler a}}^2}{b \cdot \frac{r \cdot r_0}{r + r_0}}} \quad \text{mit} \quad \lambda_i = \sqrt{\frac{r}{r + r_0}}, \quad i = 1, \dots, a-1 .$$

Zum Vergleich von B-Mittelwerten (Mittelwerte des Kleinteilstücksfaktors) sind die Freiheitsgrade $FG = a \cdot (b - 1) \cdot (r - 1)$ und unter Berücksichtigung eines Standards B_0 mit r_0 ($r_0 \geq r$) Wiederholungen $FG = a \cdot (b - 2) \cdot (r - 1) + a \cdot (r_0 - 1)$.

Die Mindestdifferenz d berechnet sich ähnlich:

t-Test:
$$d = (t_{1-\alpha/2; FG} + t_{1-\beta; FG}) \cdot \sqrt{\frac{2 \cdot s_{\text{Fehler ab}}^2}{a \cdot r}}$$

Tukey-Prozedur:
$$d = (q_{1-\alpha; FG, b} / \sqrt{2} + t_{1-\beta; FG}) \cdot \sqrt{\frac{2 \cdot s_{\text{Fehler ab}}^2}{a \cdot r}}$$

Dunnett-Prozedur:
$$d = (d_{1-\alpha; FG, b-1, \lambda_1, \dots, \lambda_{b-1}} + t_{1-\beta; FG}) \cdot \sqrt{\frac{s_{\text{Fehler ab}}^2}{a \cdot \frac{r \cdot r_0}{r + r_0}}}$$
 mit $\lambda_i = \sqrt{\frac{r}{r + r_0}}$ (zweiseitig) $(i = 1, \dots, b-1)$

2.4.3 Zahlenbeispiel

Betrachtet wird der Vergleich der A-Mittelwerte in einer zweifaktoriellen Spaltanlage (A/B) - Bl mit Hilfe der Tukey-Prozedur. Mit den nachstehenden Werten erhält man für die Zielgröße 'Anzahl Wiederholungen / Blocks'

$a = 5 \quad b = 4 \quad d = 1.78 \quad s_{\text{Fehler a}}^2 = 2.8 \quad \alpha = 0.05 \quad \beta = 0.25$

$r = 6$

Berechnet man die anderen Zielgrößen mit dieser Wiederholung, so erhält man:

a	b	r	d	$s_{\text{Fehler a}}^2$	α	β	Zielgröße
5	4	6		2.8	0.05	0.25	$d = 1.777$
5	4	6	1.78	2.8		0.25	$\alpha = 0.0494$
5	4	6	1.78	2.8	0.05		$\beta = 0.2483$

Die Genauigkeitsgrößen (Zielgrößen) werden in Abhängigkeit der Werte für die anderen Größen berechnet: $d = d(r, \alpha, \beta, s^2)$, $\alpha = \alpha(r, d, \beta, s^2)$, $\beta = \beta(r, d, \alpha, s^2)$ und $r = r(d, \alpha, \beta, s^2)$. Bei der Dunnett-Prozedur sind immer zwei Wiederholungen, die für den Standard und die für die anderen Prüfglieder zu berücksichtigen. Dem entsprechend werden bei der Zielgröße 'Anzahl Wiederholungen / Blocks' zwei im Allgemeinen verschiedene Wiederholungen berechnet. Mit den Werten $a = 5$, $b = 4$, $d = 1.78$, $s_{\text{Fehler a}}^2 = 2.8$, $\alpha = 0.05$ und $\beta = 0.25$ erhält man für den Vergleich der A-Mittelwerte mit Hilfe der zweiseitigen Dunnett-Prozedur $r = 6$ und $r_0 = 2$ für den Standard. Das

bedeutet, dass der Standard in jedem Block zweimal wiederholt werden sollte. Für die Planung der Versuchsanlage würde man in einem solchen Fall nicht von $a = 5$, der Standard tritt genauso oft auf wie die anderen Prüfglieder, sondern von $a = 6$ ausgehen.

2.4.4 Berücksichtigung mehrerer Varianzen

Entscheidend für den berechneten Wert der Zielgröße ist die Varianz. Nicht immer ist aber nur eine Varianz zu berücksichtigen (Tab. 1). Beim Vergleich der AB-Mittelwerte in einer dreifaktoriellen Spaltanlage (A/B/C) - Bl sind das zwei: $s_{\text{Fehler a}}^2$ und

$s_{\text{Fehler ab}}^2$. Mit $s_{AB}^2 = \frac{s_{\text{Fehler a}}^2}{r \cdot b \cdot c} + \frac{(b-1) \cdot s_{\text{Fehler ab}}^2}{r \cdot b \cdot c}$ und den Freiheitsgraden

$$FG = \frac{s_{AB}^4}{\frac{s_{\text{Fehler a}}^4}{(a-1) \cdot (r-1) \cdot (r \cdot b \cdot c)^2} + \frac{(b-1)^2 \cdot s_{\text{Fehler ab}}^4}{a \cdot (b-1) \cdot (r-1) \cdot (r \cdot b \cdot c)^2}}$$

berechnet sich für die Tukey-Prozedur zum Vergleich der AB-Mittelwerte die

Minstdifferenz $d = (q_{1-\alpha; FG, ab} / \sqrt{2} + t_{1-\beta; FG}) \cdot \sqrt{\frac{2 \cdot s_{AB}^2}{r}}$.

Tabelle 1: Bei den Versuchsanlagen und Mittelwertvergleichen zu berücksichtigende Varianzen

		s_{Fehler}^2
A-R	A	
A-Bl	A	

		$s_{\text{Fehler a}}^2$	$s_{\text{Fehler b}}^2$	$s_{\text{Fehler ab}}^2$
(AxB)-R	A			
	B			
(AxB)-Bl	A			
	B			
(A/B)-Bl	A			
	B			
(A+B)-Bl	A			
	B			

		$S_{\text{Fehler a}}^2$	$S_{\text{Fehler b}}^2$	$S_{\text{Fehler ab}}^2$	$S_{\text{Fehler ac}}^2$	$S_{\text{Fehler bc}}^2$	$S_{\text{Fehler abc}}^2$
(AxBxC)-R	A						
	B						
	C						
	AB						
	AC						
	BC						
(AxBxC)-Bl	A						
	B						
	C						
	AB						
	AC						
	BC						
(A/B/C)-Bl	A						
	B						
	C						
	AB						
	AC						
	BC						

		$S_{\text{Fehler a}}^2$	$S_{\text{Fehler b}}^2$	$S_{\text{Fehler ab}}^2$	$S_{\text{Fehler ac}}^2$	$S_{\text{Fehler bc}}^2$	$S_{\text{Fehler abc}}^2$
[(AxB)/C]-Bl	A						
	B						
	C						
	AB						
	AC						
	BC						
[A/(BxC)]-Bl	A						
	B						
	C						
	AB						
	AC						
	BC						
[A+(BxC)]-Bl	A						
	B						
	C						

	AB						
	AC						
	BC						
[A+(B/C)]-B1	A						
	B						
	C						
	AB						
	AC						
	BC						
[(A+B)/C]-B1	A						
	B						
	C						
	AB						
	AC						
	BC						
[A/(B+C)]-B1	A						
	B						
	C						
	AB						
	AC						
	BC						

2.4.5 Ausgabe des Ergebnisses

Die Ausgabe des berechneten Wertes der Zielgröße erfolgt auf dem Bildschirm. Der gleichzeitig verfügbare Button 'Druck' ermöglicht auch die Ausgabe des Ergebnisses mit den dazu gehörenden Eingabewerten sofort auf den Drucker. Beispielhaft wird das in Abb. 5 gezeigt. Für die Druckerausgabe wird ein eigener Stil kreiert:

```
proc template;
  define style Styles.feldva;
    parent = Styles.Printer;
  ...
```

FELD_VA II						
Versuchsanlage: zweifaktorielle Spaltanlage (A/B) - BI						
Berechnung des Mindeststichprobenumfangs						
Für die Tukey-Prozedur zum Vergleich der A-Mittelwerte						
Eingabedaten:						
a	b	d	s ² (Fehler	alpha	beta	
(Anz. Stufen	(Anz. Stufen	(Mindest-	a)			
Faktor A)	Faktor B)	differenz)				
5	4	1.78	2.8	0.05	0.25	
Ergebnis:						
Anzahl mindestens notwendiger Blocks (Wiederholungen): 6						

Abb. 5: Beispiel für den Druck des Ergebnisses

2.4.6 Umfang

Für die vier Zielgrößen r , d , α und β , drei Testprozeduren – drei deshalb, weil bei der Dunnett-Prozedur die ein- oder zweiseitige Fragestellung mit Hilfe einer Macrovariablen gesteuert wird – und den zu vergleichenden Mittelwerten der Haupt- und Kombinationswirkungen (nur für t-Test und Tukey-Prozedur im dreifaktoriellen Fall) umfasst im Frame Genauigkeit die Berechnung 660 Source-Programme und 404 Source-Programme zur Ausgabe der Informationen auf den Drucker.

2.5 Auswertung eines Einzelversuchs

2.5.1 Eingaben und Wahlmöglichkeiten

Zuerst wird die Versuchsanlage (s.o., vgl. Abb. 3) gewählt. Die Daten können im SAS- oder Text-Format vorliegen. Im SAS-Format müssen die Variablen, die das Prüfglied kennzeichnen: Faktor A, Faktor B, Faktor C und Block, mit A, B, C und Block bezeichnet sein. Beim Text-Format ist die Reihenfolge wichtig. Die ersten Spalten (ohne Überschrift) dienen der Kennzeichnung des Prüfglieds. Sie werden programmseitig bei einer dreifaktoriellen Anlage in Blocks zu den Variablen A, B, C

und Block. Dann können mehrere numerische Spalten folgen. Für beide Formate gilt, dass aus den numerischen Spalten, die nicht der Kennzeichnung des Prüfglieds dienen, mehrere als auszuwertende Merkmale ausgewählt werden können. Für diese sollte die Möglichkeit der verbalen Beschreibung (z.B.: Ertrag in kg/ha) genutzt werden.

Wenn Datendatei und auszuwählende Merkmale feststehen, werden verbale Bezeichnungen für die Prüffaktoren eingetragen. Die Blocks können zufällig oder fix gewählt werden. Das hat Auswirkungen auf die Konfidenzschätzung der Mittelwerte und auf das Varianzanalysemodell.

Der Wert für α , für das Konfidenzniveau $1 - \alpha$ und dem Signifikanzniveau, ist mit 0.05 voreingestellt. Er ist veränderbar.

Als Tests stehen zusätzlich zu Varianztabelle

- multipler t-Test
- t-Tests mit Bezugsbasis⁷
- Bonferroni-Test
- Tukey-Prozedur
- Dunnett-Prozedur (zweiseitig)

zur Wahl.

Die Datei, in die die Ergebnisse geschrieben werden soll, ist vorzugeben.

2.5.2 Grundstruktur von Proc Mixed für t-Test, Bonferroni-Test und Tukey-Prozedur

Am Beispiel einer zweifaktoriellen Spaltanlage (A/B) - Bl ist die Grundstruktur von Proc Mixed

```
<ODS-Anweisungen>
PROC MIXED NOBOUND;
  CLASS a b block;
  MODEL merkmal = <block> a b a*b / DDFM=KENWARDROGER;
  RANDOM >block< block*a;
  LSMEANS a b a*b / PDIF ADJUST=BON CL ALPHA=&alpha;
  LSMEANS a b a*b / PDIF ADJUST=Tukey CL ALPHA=&alpha;
RUN;
QUIT;
```

⁷ Vorschlag von Herrn Dr. Falk Krüger, Landesamt für Verbraucherschutz, Landwirtschaft und Flurneuordnung des Landes Brandenburg

Ist der Block-Faktor fix, so steht er in der Model-Anweisung; ist er zufällig, dann in der Random-Anweisung. Da nur eine dieser Möglichkeiten in Frage kommt, wurde die unterschiedliche Kennzeichnung für den Block-Faktor gewählt.

Die erste LsMeans-Anweisung liefert die Testergebnisse für den multiplen t-Test und den Bonferroni-Test, die zweite für den multiplen t-Test und die Tukey-Prozedur. Die Speicherung aller Ergebnisse erfolgt mit Hilfe der ODS-Anweisungen in SAS-Dateien, die für die Ausgabe in die Ausgabedatei (Text-Format) aufbereitet werden.

In der Ausgabedatei stehen der Name der Datendatei, die verbalen Bezeichnungen der Faktoren und des jeweilig auszuwertenden Merkmals, das Modell der Versuchsanlage, die Schätzungen der Mittelwerte (LsMeans) einschließlich Konfidenzintervalle, die Varianztabelle der fixen Effekte, die Schätzungen der Varianzkomponenten der zufälligen Effekte und die Ergebnisse der Tests, die am Beispiel der Tukey-Prozedur für eine zweifaktorielle Spaltanlage (A/B) - BI demonstriert werden:

Tukey-Testprozedur

Grenzdifferenzen im balanzierten Fall (!)
für die Vergleiche der

A-Effekte	:		8.306005021
	-	AB-Mittelwerte auf gleicher B-Stufe:	14.561433931
B-Effekte	:		5.4716123309
	-	AB-Mittelwerte auf gleicher A-Stufe:	11.999848249
AB-Effekte	-	(nicht auf gleicher A- bzw. B-Stufe):	14.561433931

Dem folgen paarweise die Testergebnisse für den Vergleich der A-, B- und AB-Mittelwerte – auch unter Berücksichtigung signifikanter Wechselwirkung. Diese Vergleiche werden auf Grund des für den Druck des Artikels zur Verfügung stehenden Formats nicht vorgestellt. Abschließend wird eine Kurzform der Signifikanzdarstellung ausgegeben:

Tukey-Testprozedur

Methode der Verbindungslinien (V)

Mittelwerte, die durch die selbe Linie verbunden werden, sind untereinander nicht signifikant.

```

A      MEAN_A V      (HSD = 14.561433931)
                        A B      MEAN_AB V
2 150.31307 |
1 158.15307 | |
3 158.69514 |
                        2 2 144.35776 |
                        2 3 145.94419 | |
                        3 2 150.97486 | | |
                        2 5 153.43688 | | |
B      MEAN_B V      2 1 153.91177 | | |
                        2 4 153.91473 | | |
2 150.6305 |
3 155.87481 | |
5 156.98837 |
1 157.29235 |
4 157.81609 |
                        1 2 156.55889 | | |
                        3 1 157.33593 | | |
                        1 5 157.50156 | | |
                        1 4 157.69945 | | |
                        1 3 158.37607 | | |
                        3 5 160.02667 | |
                        1 1 160.62936 |
                        3 4 161.83409 |
                        3 3 163.30417 |

```

2.5.3 Dunnett-Prozedur (zweiseitig)

Wird als Testverfahren die Dunnett-Prozedur gewählt, dann sind folgende LsMeans-Anweisungen (hier für eine zweifaktorielle Anlage) in die Grundstruktur von Proc Mixed einzufügen (s.o.):

```

LSMEANS a / PDIFF ADJUST=DUNNETT DIFF=CONTROL("StandardA")
                        CL ALPHA =&alpha;
LSMEANS b / PDIFF ADJUST=DUNNETT DIFF=CONTROL("StandardB")
                        CL ALPHA =&alpha;
LSMEANS a*b/ PDIFF ADJUST=DUNNETT
                        DIFF=CONTROL("StandardA" "StandardB") CL ALPHA =&alpha;

```

Da es möglich sein muss, ein oder mehrere Stufen eines Faktors als Standard zu deklarieren, findet der Nutzer auf dem Frame Einzelversuch entsprechende Button, die die Zuordnung der Stufe(n) als Standard ermöglichen. Wird eine derartige Zuordnung nicht vorgenommen, gilt jeweils die erste Stufe als Standard. Da die Standards aus mehreren Stufen eines Faktors gebildet werden können, wird aus diesen Stufen eine neue Stufe (z.B. StandardA) gebildet. Da sich im Vergleich zu den bereits genannten Tests (s.o.) die Anzahlen der Stufen der Faktoren ändern können, werden neben den Mittelwerten und deren Konfidenzintervallen auch die Varianztabelle der fixen Effekte, die Schätzungen der Varianzkomponenten der zufälligen Effekte neu berech-

net. Die Grenzdifferenzen werden wie bei der Tukey-Prozedur ausgegeben. Die paarweisen Vergleiche werden ähnlich denen der Tukey-Prozedur begrenzt für die Vergleiche mit den Standards in die Ausgabedatei geschrieben.

2.5.4 t-Tests mit Bezugsbasis

Die Testmöglichkeit 't-Tests mit Bezugsbasis' führt zum Testen linearer Kontraste mit Hilfe des t-Tests. An die Stelle der LsMeans-Anweisungen für die Tests (s.o.) treten nun Estimate-Anweisungen. Um das zu veranschaulichen seien die Stufen des Faktors A beispielhaft A1, A2 und A3. Wenn keine andere Bezugsbasis festgelegt wird (auch dafür stehen auf dem Frame Einzelversuch entsprechende Button bereit), wird die erste Stufe als Bezugsbasis angenommen: A1 = BB_A . Die Estimate-Anweisungen lauten dann für die Vergleiche A2 - BB_A und A3 - BB_A:

```
ESTIMATE 'A2-BB_A' A -1 1 0 /alpha=0.05 cl divisor= 1;
ESTIMATE 'A3-BB_A' A -1 0 1 /alpha=0.05 cl divisor= 1;
```

Die Ergebnisse entsprechen natürlich denen des multiplen t-Tests.

Seien die Stufen des Faktors B beispielhaft B1, B2, B3 und B4 und als Bezugsbasis werden B2 und B3 (= BB_B) gewählt, dann lauten die Estimate-Anweisungen für diese Kontraste:

```
ESTIMATE 'B1-BB_B' B 2 -1 -1 0 /alpha=0.05 cl divisor= 2;
ESTIMATE 'B4-BB_B' B 0 -1 -1 2 /alpha=0.05 cl divisor= 2;
```

Derartige Estimate-Zeilen hängen von der Anzahl der Stufen und der gewählten Bezugsbasis ab. Sie müssen folglich variabel in Proc Mixed eingefügt werden. Das bedeutet, dass sie vorher „zusammen zu basteln“ und in Dateien abzulegen sind. Diese Dateien werden dann aufgerufen.

Den letzten Schritt, das Schreiben der einzelnen Teile in eine Textdatei, verdeutlichen die folgenden Anweisungen.

```
filename est1 "c:\est1.txt";
data includeA;
  file est1;
  merge inca1 inca2 inca3;
  by i;
  drop i;
  put ESTIMATE coll--col&aa divisor;
run;
```

In Proc Mixed werden an Stelle der Estimate-Anweisungen die Befehle

```
%INCLUDE est1 ; * für Vergleiche Faktor A ;
%INCLUDE est2 ; * für Vergleiche Faktor B ;
```

gesetzt. Beim Ausführen dieser Prozedur werden dann die entsprechenden Estimate-Anweisungen eingefügt. Richtig aufwändig wird das Erstellen der Estimate-Zeilen, wenn Wechselwirkungen zu berücksichtigen sind.

Da Textdateien angelegt wurden, müssen sie wieder gelöscht werden:

```
data _null_;
  fname="est1";
  rc=filename(fname,"C:\est1.txt");
  if rc = 0 and fexist(fname) then rc=fdelete(fname);
  rc=filename(fname);
run;
```

Für beispielsweise drei Stufen A1, A2 und A3 mit der Bezugsbasis BB_A = A1 und fünf Stufen B1, B2, B3, B4 und B5 mit der Bezugsbasis BB_B = B1 und B2 erhalten wir als Ergebnis (ohne Standardfehler, Freiheitsgrade und Überschreitungswahrscheinlichkeit verkürzte Darstellung):

Vergleich	Differenz der Mittelwerte	Test	Konfidenz- intervall untere Grenze	Konfidenz- intervall obere Grenze
A 2-BB_A	-7.8400	signifikant	-14.4642	-1.2158
A 3-BB_A	0.5421	n.s.	-6.0821	7.1663
B 3-BB_B	1.9134	n.s.	-1.4342	5.2609
B 4-BB_B	3.8547	signifikant	0.5071	7.2022
B 5-BB_B	3.0269	n.s.	-0.3206	6.3745

3 Ausblick

FELD_VA II hat im Vergleich zum Vorgänger ein größeres Leistungsangebot, was alleine schon durch die Berechnung von Genauigkeitsgrößen auffällt, eine dem heutigem Stand entsprechende Wichtung der Varianzen und auch die Berücksichtigung fehlender Werte mit der in Proc Mixed verwendeten LsMeans-Schätzung (REML-Schätzung) der Mittelwerte.

Die bereits vorliegenden Programme haben eine beträchtliche Zahl erreicht und sind auf versteckte Fehler hin zu untersuchen.

Liegen stärker unbalancierte Daten vor, so sollen sie entsprechend ausgewertet werden. Die Auswirkungen auf die Varianzanalyse, insbesondere auf die Zerlegung der Quadratsummen, und auch die Zulässigkeit einiger Tests, das bedeutet das Ausweichen auf andere Testverfahren, sind zu berücksichtigen.

Gehen wir davon aus, einen Einzelversuch als Teil einer Versuchsserie anzusehen, so zeichnet sich auch für diesen Bereich ein u.U. aufwändiger Weg ab.

Literatur

MOLL, E. (1996): Die SAS-Anwendung FELD_VA - Konstruktion des Lageplanes und die varianzanalytische Auswertung ein- bis dreifaktorieller Feldversuche
Berichte aus der Biologischen Bundesanstalt für Land- und Forstwirtschaft, Heft 14, 1996, 43 S.

MOLL, E. (1997): FELD_VA - Konstruktion des Lageplanes und varianzanalytische Auswertung ein- bis dreifaktorieller Feldversuche
1. Konferenz der SAS-Benutzer in Forschung und Entwicklung (KSFE), Berlin, Konferenzbeiträge, S. 254-255