



Stand: 11.06.2009

## PIAFStat-Verfahren für die „Hohenheim-Gülzower Serienauewertung“

**Name:** MW

**Label:** Bestimmung der Mittelwerte

**Macros:** %lsm

**Funktionen:** PFAD; TYP; MEAN; DATEN2; VORGABEN; OPTION; MC\_LSM; ADD; DUBL\_CHECK; ZWEIDEUTIG1; GRAFIK; REGRESSION; RESIDUEN; VERSUCHSNIVEAU; VK\_ANPASSER, VK\_KONS

**Konzeption:** Michel, Piepho

**Umsetzung SAS, PIAFStat:** Zenk

**SAS-Macros:** Möhring

Seit Version vom 21.9.2007 wurden insbesondere folgende Neuheiten in das Programm eingefügt:

- Erweiterung der Option ,04 – mehrfaktoriell (j/n)' um Stufenmittel und Behandlungseffekt
- Erweiterung der Option ,05 – Reduktion' um Unteroption O7B – Versuche je Sorte insgesamt
- Nutzung von Erfahrungswerten für Phi (externe Datei PHISart.xls)
- Erweiterung der Plausi-Tabelle:
  - relativierte Residuen nach vollst. Modell -> rRes\_Mod = outpredict (\*100/GesMW)
  - vereinfachte Residuen Sorte\*Umwelt -> Res\_su = diff- (MW(diff)by s)

Seit Version vom 3.5.2006 wurden neben kleinen Fehlerbeseitigungen folgende Neuheiten in das Programm eingefügt:

- Abspeicherung der Ergebnisse auch in Access (siehe auf Ergebnis-Pfad nach)
- Bildung der Dateinamen mit Zusatz von [A1] und beliebiger Ergänzung (&Z; default=1)
- Plausibilitätstabelle (erscheint nicht im LIST-Fenster, sondern wird nur in Excel bzw. in der Access-DB als Tabelle Plausi\_[A1]\_&Z abgelegt)
- Ergänzung der Tabelle ‚Sorten-MW‘ um Öko-Stabilitätsparameter (Öko-se und b), dies ist in Option O24A auszuwählen, wird aber nur bei großen Datenmengen berechnet (Es müssen Sorten im Dataset sein, die ein n gesamt größer als 19 besitzen.)
- Ausgabe der Grafik Sortenmittelwerte mit Intervallen zu den paarweisen Vergleichen

## 1. Ziel und Leistung des Verfahrens

Ziel der Auswertung ist eine optimale Einschätzung des mittleren (Ertrags-) Potentials einer Sorte für eine Region (=Zielgebiet). Dieser Schätzwert ist unter Berücksichtigung der anderen wertbestimmenden Sorteneigenschaften die beste Grundlage für die Sortenwahl für ein neues Jahr, dessen konkrete Spezifik (Witterung, Befallssituation ...) im Vorherein nicht bekannt ist

Dabei können alle verfügbaren Daten auch bei unvollständigen Datenstrukturen, also z.B. die den LSV zeitlich vorgelagerten WP, EUSV und OV einbezogen werden. Außerdem können benachbarte standortkundlich ähnliche Anbauggebiete bzw. Boden-Klima-Räume mit einfließen, wobei diese ein optimal verringertes Gewicht erhalten. Für ein vorzugebendes Zielgebiet werden (mehrjährige) Sorten-Mittelwerte in folgender Weise bestimmt: die MW sind auf eine einheitliche Ebene /Vergleichsbasis /mittleres absolutes Ertragsniveau so projiziert, dass sie unmittelbar und optimal miteinander vergleichbar sind (Adjustierung). Eine Relativierung muss außerhalb dieses Verfahrens erfolgen.

Als Ergebnis-Tabellen werden die Sorten-Mittelwerte sowie Versuchs-Mittelwerte, Anzahl Versuche je Sorten in Anbaugebieten und Gewichte der Anbauggebiete je Sorte in optionalen Sortiervarianten angeboten. Alle Ergebnis-Tabellen können in Excel abgespeichert werden.

Anforderungen:

- Regionalspezifik berücksichtigen
- robuste und stabile MW-Schätzung
- Objektivität und Unabhängigkeit
- optimiert, effizient, schnell, handhabbar
- fehlerkritisch

## 2. IT- Voraussetzungen und – Hinweise

(Siehe hier auch Info\_VK.pdf!)

- Das Verfahren ist eigenständig lauffähig
- Bei Verwendung einer A-Serie aus PIAF empfehlen wir, die ADS auf der Aggregationsebene A (einfaktorielle Versuche) bzw. AB (zweifaktorielle Versuche) auszugeben. Sind aber keine Mittelwerte (arithmetisch oder adjustiert) in PIAF zurückgelesen worden, kann auch die ADS auf der Aggregationsebene P ausgegeben werden. Ein interner Verfahrensschritt bildet vor dem weiteren Programmablauf temporär arithmetische Mittelwerte.
- In die ADS brauchen nur die zu analysierenden Merkmale und gegebenenfalls ihre zugehörigen Standardfehler ausgegeben werden.
- Im Verfahren MW wird auf die durch das Verfahren VK erzeugte Datei **varianzk.sas7bdat** zugegriffen. Diese Datei muß im Ergebnispfad (z.B. G:\Hoh\_meth\Ergebnisse\Kulturart) als SAS-Datei abgelegt sein. Der Dateiname kann durch den Anwender aber verändert worden sein. Die Möglichkeit der Veränderung des Dateinamens ist im Verfahren gegeben.
- Zur Unterstützung des Verfahrens kann auf externe Excel-Dateien zugegriffen werden.

Excel-Datei	Erläuterung	siehe auch:
obermenge.xls	externes Dataset	Info_vk.pdf oder <b>O3</b> (unten)
reg.xls	Excel-Datei zur Zuordnung der Versuchsorte zu Anbauregionen	Info_vk.pdf oder <b>O9</b> (unten)
phistart.xls	Erfahrungswerte von Phi	Neuerungen ab Juni 2008.doc

reglim.xls	Excel-Datei zur Einschränkung der Daten je nach den für ein Zielgebiet relevanten Regionen	nächster Absatz
gruppen.xls	für Darstellungszwecke in Ergebnistabellen (nicht unbedingt notwendig)	Info_vk.pd
sorti.xls	Excel-Datei des derzeit aktuellen LSV-Sortimentes, Voraussetzung der Sortierung und der Grafik	übernächster Absatz und <b>O11</b> und <b>O18</b>

- Die Einschränkung der Daten je nach den für ein Zielgebiet relevanten Regionen wird notwendig, wenn für jedes ZAG andere NAG relevant sind. Geschaffen wurde ein Auswahlmechanismus, der ohne Manipulation der Daten oder reg-Datei auskommt. Die Information sind dazu vorher auf dem Ordner für externe Daten in einer Excel-Mappe mit dem derzeit festen Namen "reglim" einzutragen. reglim.xls hat folgendes außer dem Namen zwingende Format
  - Wenn die Mappe existiert, dann muss sie für das auszuwertende ZAG eine Tabelle enthalten, deren Name die Nummer des ZAG ist (gemäß der Macrovariablen im Verfahren).
  - Diese Tabelle enthält eine Spalte mit der Überschrift r, unter der die Nummern aller für das ZAG relevanten Regionen (d.i. seine relevanten NAG zuzüglich seiner selbst) aufgelistet sind. Soll keinerlei solche Einschränkung vorgenommen werden, so ist lediglich keine reglim-Datei auf den Ordner für externe Daten zu stellen. Das Verfahren passt sich automatisch an. Möchte man einige ZAG einschränken, andere nicht, empfiehlt sich eine reglim-Datei, in der sich auch für die nicht einzuschränkende ZAG eine Tabelle befindet, welche eben alle Nummern der in den Daten enthaltenen Regionen enthält.
- Soll die Sortierungsoption ‚1.Priorität – PG-Nr.‘ (O19) genutzt werden (was aber nicht zwingend erforderlich ist!), wird eine externe Excel-Datei sorti.xls benötigt. Diese Datei kann zum Beispiel das derzeit aktuelle Sortiment des LSV des Zielgebietes enthalten. Sie sollte im Datenpfad (z.B. G:\Hoh\_meth\Daten\Kulturart) abgelegt sein. Das Beispiel einer solchen Datei wird den Verfahren beigelegt (<http://www.biomath.de> → Download). Wichtig: notwendige Tabelle im Tabellenblatt ‚Tabelle1‘ ablegen. In der ersten Zeile müssen die Spaltenüberschriften stehen. Folgende Informationen und Spaltenüberschriften muss diese Tabelle enthalten (Spaltenüberschrift→ Erläuterung):
  - S → Sortenschlüssel, BSA-Kennnummer entsprechend PIAF (z.B. WW 025360)
  - PGNr → Prüfgliednummer bzw. anderweitig definierte numerische Sortierungsnummer

#### Leistung der Funktionen:

Funktion	Leistung
PFAD	Pfade festlegen
TYP	Automatische Generierung des Typs im Datensatz
MEAN	Komprimierung doppelter Datensätze (z.B. bei Intensitätsstufe II und III)
DATEN2	Vorbereitung des Datensatzes, Datenbasis, und deren Kontrolle
VORGABEN	Phi und alpha können vorgegeben werden
OPTION	gewählte Optionen werden im Output ausgegeben
ADD	Automatische Generierung der Additionskonstante
DUBL_CHECK	Prüfung des Datensatzes auf Dubletten
ZWEIDEUTIG1	Prüfung der BSA-Kennnr. und der Sortennamen auf Eindeutigkeit
MC_LSM	Macro-Aufruf zur Berechnung der Mittelwerte
GRAFIK	Grafik nach Berechnung der Sorten-Mittelwerte
REGRESSION	Ökovalenzberechnung
RESIDUEN	Ausgabe von Residuen in Plausi-Tabelle
VERSUCHSNIVEAU	Tabelle LSV-Niveaus
VK_KONS	VK nach konservativer Methode umwandeln
VK_ANPASSER	Anpassung der VK an vorhandene Regionen

### 3. Optionen

- O0: Kulturart:

Es erfolgt die Auswahl einer Kulturart, die bearbeitet werden soll. Ist diese Kulturart derzeit nicht aufgeführt, dann kann sie in der Funktion PFAD im Deklarationsteil mit aufgenommen werden.

Es wird eine Macrovariable (&Kulturart) zur Kulturart gebildet. Diese wird im Verfahren genutzt, um z.B. die Pfad-Struktur zu vervollständigen oder um das kulturart-spezifische Anbaugebiet anzusprechen zu können.

- O1: Pfad zu den Ergebnis - Dateien:

Hier kann die in der Funktion **PFAD** vorgegebene Verzeichnisstruktur temporär für einen SAS-Durchlauf verändert werden. Durch Veränderung der Angaben in der Funktion PFAD wird die Verzeichnisstruktur dauerhaft geändert.

- O2: Pfad zu den externen Dateien:

Entsprechend der Option O1 werden hier Veränderungen der Pfadangaben für die externen Dateien vorgenommen.

- O3: Datenbasis

In diesem Verfahren wird im Normalfall auf externe Daten (obermenge.xls), die nach Nutzung des Verfahrens DATEN abgespeichert wurden, zugegriffen. Es kann aber auch ausschließlich auf eine bestehende ADS einer A-Serie zugegriffen werden, ohne das Verfahren Daten vorzuschalten.

- O4: mehrfaktoriell (ja/nein):

Ist diese Option deaktiviert, wird von einer einfaktoriellen Auswertungsserie ausgegangen. Weiteres ist dabei nicht zu berücksichtigen.

Bei Aktivierung der Option kann eine zweifaktorielle Auswertungsserie ausgewertet werden. In diesem Fall muß der Anwender sich für eine der Unteroptionen entscheiden:

- Einzel-Intensitätsstufe – es ist eine Stufe zu wählen (z.B.:2 )
- Stufenmittel – hier wird angegeben, das Mittel welcher Stufen ausgewertet werden soll (z.B.: 1+2 oder 1+2+3 oder 2+3) (z.Z. noch nicht realisiert !)
- Behandlungseffekt – hier erfolgt die Auswertung von Stufendifferenzen (z.B. 2-1 oder 2+3-1) (z.Z. noch nicht realisiert !)

- O5: Reduktion des Datensatzes:

Der Anwender hat die Möglichkeit, die Datenbasis hinsichtlich Anzahl einzubeziehender Jahre und Sorten zu reduzieren. Das empfiehlt sich z.B. bei Testläufen, um die Rechenzeit zu reduzieren oder bei Szenarien, wenn z.B. das Ergebnis eines sehr langjährigen Datensatzes mit dem eines zeitlich abgekürzten Datensatzes verglichen werden soll.

Hier ist es auch möglich, die Mittelwertberechnung für ein spezielles Auswertungsjahr vorzunehmen. Dann ist anstelle der Unteroption Startjahr die Unteroption Auswertungsjahr zu aktivieren und im rechten Eingabeblock das gewünschte Jahr einzugeben.

Alle gewählten Einstellungen sind im Output nachzuvollziehen.

- O8: Zusatz zu Dateinamen:

Hier ist es dem Anwender möglich, einen Zusatz zu den Ergebnis-Datei- Namen festzulegen. Default hierfür ist 1.

Dieser Zusatz wird in einer Makrovariablen Z abgespeichert. Neben der Merkmalsbezeichnung wird allen Ergebnisdateien dieser Zusatz zugefügt (Beispiel: S\_MW\_ERTR86DT\_1.xls). Dadurch wird das Abspeichern mehrerer Mittelwerttabellen unterschiedlicher Merkmale und Zielgebiete möglich, ohne das versehentlich ein Überschreiben der Ergebnis-Dateien erfolgt.

- O9: Regionen:

Im Normalfall werden Regionen verwendet, um die objektiv abgestufte Wichtung von Ziel- und Nachbargebieten zu gewährleisten. Wird diese Option ausgeschaltet, werden alle Orte einer gemeinsamen Region (,100') zugeordnet. Dies entspricht der Strategie einer Großraum-Verrechnung, aber ohne abgestufte Wichtung der Anbaugebiete innerhalb des Großraumes. Z.B. könnte damit das Ergebnis zweier Rechengänge a) mit und b) ohne abgestufte Wichtung verglichen werden. (Um dies durchführen zu können, müssen aber auch dementsprechend zwei Dateien zu den Varianzkomponenten bereitstehen – also ist vorher das Verfahren VK sowohl mit aktivierter als auch mit deaktivierter Option ,Regionen' notwendig. Zudem ist im Verfahren MW daran zu denken, beim Verrechnen mit Region 100 das Zielgebiet zu deaktivieren oder auch auf 100 zu setzen.)

Beachte: Beide Rechengänge haben einen unterschiedlichen fachlichen Anspruch, unterschiedliche Aussagebereiche und Repräsentativität. Die Präzision der Ergebnisse ist im Standardfehler (SE) also auch nur bedingt vergleichbar.

Z.Z. erfolgt die Zuordnung der Regionen im Dataset durch Mergen mit externen Stammdaten aus der Datei reg.xls.

Die dem Datensatz zugeordneten Regionen werden im Output ausgegeben und können so kontrolliert werden. Bei Vergabe einer Region ,0' müssen die Stammdaten in reg.xls hinsichtlich Vollständigkeit und Übereinstimmung der ortid mit den Stammdaten von PIAF überprüft werden denn die ,0' zugeordneten Versuchsdaten bleiben unberücksichtigt! In diesem Sinne kann die Verwendung der Region ,0' in reg.xls aber auch gezielt genutzt werden.

- O10: Varianzkomponenten

Hier wird dem Anwender die Möglichkeit gegeben, den Namen der heranzuziehenden Datei ,varianzk\_[A1]\_&Z' zu verändern.

- O11: Sortenreihenfolge

Wird es gewünscht, in den Ergebnistabellen die Sortierung der Sorten nach einer aktuellen Sortimentsreihenfolge vorzunehmen (O19), ist eine externe Excel-Datei (sorti.xls) notwendig. Sorten, die nicht in dieser Datei aufgeführt werden, erhalten eine einheitliche PG-Nummer (500). Durch Aktivierung der Option O11 wird die externe Datei eingelesen, dabei kann der Anwender gegebenenfalls den Namen der externen, einzulesenden Datei verändern. Ist diese Option deaktiviert, erhalten alle Sorten eine PG-Nummer (500).

Zusätzlich enthält diese externe Datei Angaben, die für die Gestaltung der Grafik notwendig sind (Sorten der BB und Sorten, die in der Grafik dargestellt werden sollen). Wird die Option O11 ausgestellt, wird auch automatisch die Option O29 (Grafik) deaktiviert.

- O12: Dataset abspeichern:

Mit dieser Funktion kann ein Dataset im Excel-Format abgespeichert werden, der exakt die Daten für die Berechnung der Mittelwerte enthält. Das bietet sich an, wenn die Berechnungen später noch einmal wiederholt werden sollen (z.B. mit anderem phi).

- O13: phi

Hier kann eine Transformation der Urdaten angewiesen werden. Ein Transformationserfordernis besteht, wenn die Modellvoraussetzungen (z.B. Additivität) nicht hinreichend erfüllt sind (Bei Prozent-Merkmalen wie z.B. Anteil lagernder Pflanzen bei Mais) trifft das i.d.R. zu. Dies kann im Verfahren 'phi' geprüft werden, in dem dann auch das optimal zu verwendende phi ausgegeben wird. Der ggf. verwendete Transformationsparameter phi muss identisch mit dem phi im Verfahren ,Bestimmung der Varianzkomponenten' sein.

Das verwendete phi wird im Output ausgewiesen.

- O14: Zielgebiet

Hier wird dem Anwender die Möglichkeit gegeben, das Zielgebiet anzugeben. Diese Angabe muß dem in der Datei reg.xls enthaltenen numerischen Schlüssel für das Zielgebiet übereinstimmen.

Das Zielgebiet, für das die Mittelwerte errechnet werden, wird im Output ausgewiesen.

- O15: Gewichtung nach SE (ja/nein)

Wenn für die Mittelwerte aus den Einzelversuchen SE (Standardfehler dieser Mittelwerte) vorliegen, empfehlen wir die Wichtung. Bei Verwendung der Gültzower Verfahren zur Einzelversuchsauswertung werden die SE automatisch erzeugt, in PIAF gespeichert und stehen ohne Mehraufwand für dieses Serien-Verfahren zur Verfügung.

Die unterschiedliche Schätzgenauigkeit der Mittelwerte in Abhängigkeit von der jeweiligen Wiederholungsanzahl im Versuch bzw. des einzelnen Prüfgliedes und vom versuchsspezifischen Fehler (s%) wird dann automatisch optimal (in Abhängigkeit von der Relation der VK zueinander) berücksichtigt. Die Schätzgenauigkeit der Serien-Mittelwerte steigt durch die SE-Wichtung (i.d.R. aber nur leicht). Bei stabilen guten Serien ist die Auswirkung auf die Serien-Mittelwerte meist nur geringfügig, in kleinen Serien mit großer Differenzierung der Versuchsqualität – wie bei ‚kleinen Arten‘ oft gegeben – können sich allerdings relevantere Unterschiede einstellen. Dies sollte als Verbesserung angesehen werden.

**Voraussetzung für diese Option ist, dass auch im Verfahren VK mit Gewichtung nach SE gerechnet wurde.**

Ob mit Gewichtung gerechnet wird oder nicht, wird im Output ausgewiesen.

- **O16: Übersicht und Kontrolle:**

Diese Option erzeugt eine Übersicht über die einbezogenen Versuche und ihre Aufteilung auf Jahre, Orte, Anbauggebiete und Länder. Fehlende Zuordnungen Ort-Anbauggebiet werden erkannt und können in reg.xls ergänzt werden.

- **O17: Dubletten-Check**

Es werden Versuche erkannt, die im Dataset z.B. durch das Zusammenführen von PIAF- und externen Daten doppelt vorhanden sind.

- **O18: Sortierung der Sorten**

Für die in Option O23 ausgewählten Ergebnistabellen (Zeile=Sorte) wird durch diese Option die Sortierung der Sorten festgelegt. Sie ist bei allen Ergebnis-Tabellen dann identisch.

Als Möglichkeiten können gewählt werden:

**1. Priorität - PG-Nummer** (entspricht der PGNr in der Datei sorti.xls, siehe O11)

Das Ausschalten dieser Option ist möglich! Dann wird nur nach 2. Priorität sortiert. Bei Aktivierung erhalten die Sorten, die in der Datei sorti.xls – also im aktuellen Sortiment – nicht enthalten sind, erhalten eine PG-Nummer 500.

Deren Sortierung erfolgt dann nach:

**2. Priorität - ‚alphabetisch‘ oder ‚BSA-Kenn-Nummer‘ oder ‚Prüfhäufigkeit‘**

(genau eine der Optionen 2. Priorität ist möglich)

- **O23: Ergebnis-Tabellen**

Als Ergebnis-Tabellen können ausgegeben werden:

- |                       |   |
|-----------------------|---|
| - Sorten-MW           | → Mittelwerte der Sorten im Zielgebiet                      |
| - Zusatz Öko-se und b | → zusätzliche Spalten öko-se und b in Sorten-MW-Tabelle     |
| - Versuchs-MW         | → Mittelwerte der einbezogenen Versuche                     |
| - N in Anbaugebieten  | → Anzahl der einbezogenen Versuche je Sorte und Anbaugebiet |
| - Gewichte            | → Gewichtung der Anbaugebiete je Sorte                      |
| - Plausibilität       | → Tabelle zur Unterstützung von Plausibilitätsprüfungen     |

Jede zur Ausgabe ausgewählte Ergebnistabelle wird im Output dargestellt und zusätzlich in je einer Exceldatei bzw. in der ACCESS-DB im Ergebnisfad abgespeichert. Im Dateinamen/Tabellennamen wird das verrechnete Merkmal und ein Zusatz (&Z) mit angegeben. Der Name der Excel-Datei ist vom Anwender veränderbar. Neben den Excel-Ergebnistabellen werden auch die zugrunde liegenden SAS-Datasets abgespeichert, um in einem späteren Verfahren auf diese errechneten Werte zugreifen zu können.

## Standard- Vorgaben zu den Ergebnis- Tabellen:

Kurzbezeichnung	Default-aktiviert?	nutzt Option ‚Sortierung der Sorten‘?	Vorgabe für Dateinamen	Excel-Datei- Name	SAS-Datei- Name
Sorten-MW	ja	ja	S_MW	S_MW _Merkmal_&Z	S_MW _Merkmal_&Z
Versuchs-MW	ja	nein	V_MW	V_MW _Merkmal_&Z	V_MW _Merkmal_&Z
N in Anbaugebieten	ja	ja	N_AG	N_AG _Merkmal_&Z	N_AG _Merkmal_&Z
Gewichte	nein	ja	G_AG	G_AG _Merkmal_&Z	G_AG _Merkmal_&Z
Plausi	ja	nein	Plausi	Plausi _Merkmal_&Z	Plausi _Merkmal_&Z

### • O29: Grafik

Sind in der externen Hilfsdatei sorti.xls die Informationen BB (Sorten der Bezugsbasis, mit B gekennzeichnet) und Diagramm (Sorten, die in der Grafik dargestellt werden sollen) enthalten, können die Grafiken für die absoluten und für die relativen Sortenmittelwerte in eine rtf-Datei aufgegeben werden.

Ist die Option O11 ‚Sortenreihenfolge‘ deaktiviert, kann die Option O29 ‚Grafik‘ nicht aktiviert werden!

Auszug aus der Datei sorti.xls:

PGNr	Name	gr	s	BB	Diagramm
1	Akteur	E	WW 02998	B	1
2	Cetus	E	WW 03176	B	2
3	Magister	E	WW 03197	B	3
4	Skagen	E	WW 03382		4
	usw.				

### Ausblick:

- ein- und mehrjährige Auswertung und Tabellierung kombinieren
- mehrere externe Datensätze verknüpfen
- GD für Sortengruppen (1-jährig LSV, 2-jährig ... oder 4-7 Versuche; 8-12 Versuche; >12 Versuche)

### 4. Standard- Output

- Anzahl Datensätze, die zur Verrechnung bereit stehen
- Übersicht zur Anzahl Versuche je Jahr und Ort - Kontrolle der Zuordnung zu den Anbaugebieten (AG)
- Übersicht zu den Class-Merkmalen im Gesamt-Datensatz
- Ergebnis des Dubletten-Checks
- Ausgabe der Ergebnisse von No-Name-Check, Namen-Check und BSA-Kennnummern-Check. Diese Ausgaben sind nicht optional zu wählen sondern erfolgen generell!

(diese Punkte entsprechen dem Output des Verfahrens VK → siehe: Info\_VK.pdf)

- Ergebnistabellen

Die Abkürzungen im Tabellenkopf bedeuten:

MW	-	adjustierter Mittelwert der Sorten im Zielgebiet
se	-	Standardfehler des Mittelwertes
N ges	-	Anzahl Versuche mit dieser Sorte in der Summe aller Anbaugebiete
N ZG	-	Anzahl Versuche mit dieser Sorte im Zielgebiet

### **Mittelwerte der Sorten im Zielgebiet 4**

für das Merkmal ERTR86DT Intensitätsstufe: 2

Auswertungszeitraum: 2002 bis 2005

Mindestanzahl Versuche je Sorte im Zielgebiet: 1

Wichtung nach Versuchspräzision: ja gewählter Transformationsparameter: Phi = 1

Sorten					MW	se	N ges	N ZG
1	Enorm	WW	02803	E	90.0	1.1	26	18
2	Akteur	WW	02998	E	90.5	1.2	22	12
3	Privileg	WW	03080	E	91.4	1.2	19	14
7	Pegassos	WW	01969	A	94.9	1.4	20	1
8	Cubus	WW	02787	A	96.3	1.1	44	18

usw.

#### Versuchs- Mittelwerte

für das Merkmal ERTR86DT Intensitätsstufe: 2

AR / Ort / Jahr / VNR				Versuchsmittel
5	Vi pperow	2003	108	59.9
		2004	3	88.3
		2005	4	74.1
			959	87.7 usw.
4	Neuhof 1	2003	103	87.7
		2004	104	105.2
			102	100.3
		2005	1	95.1
	Bi estow	2003	109	98.6

usw.

#### Anzahl Versuche in Anbaugebieten

für das Merkmal ERTR86DT Intensitätsstufe: 2

Sorten					AG		
					4	5	3
1	Enorm	WW	02803	E	18	8	.
2	Akteur	WW	02998	E	12	7	3
3	Privileg	WW	03080	E	14	4	1
7	Pegassos	WW	01969	A	1	19	.
8	Cubus	WW	02787	A	18	26	.

usw.

#### Gewicht eines Anbaugebietes in der Summe der Versuche

für das Merkmal ERTR86DT Intensitätsstufe: 2

Summe der Gewichte = 1 (0 bis 1 für 0% bis 100%)

Sorten	ag_3	ag_4	ag_5
--------	------	------	------



1	Enorm	WW	02803	E	0.00	0.75	0.25
2	Akteur	WW	02998	E	0.09	0.63	0.28
3	Pri v i l e g	WW	03080	E	0.07	0.76	0.17
7	Pegassos	WW	01969	A	0.00	0.10	0.90
8	Cubus	WW	02787	A	0.00	0.61	0.39

usw.

#### 4. Methodische Hintergründe und Diskussion (MICHEL)

Ziel der Auswertung ist eine optimale Einschätzung des mittleren (Ertrags-) Potentials einer Sorte für eine Region (=Zielgebiet). Dieser Schätzwert ist unter Berücksichtigung der anderen wertbestimmenden Sorteneigenschaften die beste Grundlage für die Sortenwahl für ein neues Jahr, dessen konkrete Spezifik (Witterung, Befallssituation ...) im Vorherein nicht bekannt ist.

Für diese Schätzung sollen alle geeigneten Reserven mit optimaler Gewichtung einbezogen werden:

- möglichst vieljährige Daten unter Einbeziehung der den LSV vorgelagerten WP und EUSV
- Versuche aus benachbarten, agrarökologisch ähnlichen Gebieten

Diese Versuche erhalten für die Schätzung allerdings nicht das gleiche Gewicht wie Versuche, die im Zielgebiet liegen. Der Grad dieser Wichtungsabstufung wird automatisch in der Weise vorgenommen, dass der Vorhersagefehler im Ziel-Anbaugebiet minimiert, also die Sortenbewertung im Ziel-Anbaugebiet optimiert wird. D.h. die Gewichte sind so gewählt, dass ein Nachbargebiet ein umso höheres Gewicht erhält, je ähnlicher es dem Zielgebiet ist. Wird aus den Daten andererseits keine Ähnlichkeit festgestellt, werden die Daten dieses Nachbargebietes automatisch nicht einbezogen (Wichtung=0), ohne dass sie eigens aus dem Datensatz gestrichen werden müssten. Die Gewichte werden maßgeblich durch die genetische Korrelation der Anbaugebiete bestimmt; diese ist ein Maß für die Ähnlichkeit der Gebiete, das implizit für die Berechnung der optimalen Gewichte verwendet wird. Grundlage dafür bildet die Schätzung der Varianzkomponenten im zeitlich vorgelagerten Verfahren ‚VK‘, dessen Infotext bezüglich methodischer Hintergründe ebenfalls relevant ist.

Bundesweit führt diese Methode zu gleitend überlappenden Anbaugebieten. So nutzt D-Nord z.B. Daten aus D-Süd und Ostholstein, Ostholstein wiederum Daten aus D-Nord, nicht aber aus D-Süd usw.

nachfolgend in loser Abfolge Notizen zu methodischen Problemen, Fragen und Diskussionen im Zusammenhang mit der Schätzung adjustierter Mittelwerte:

#### Modelle zur Auswertung beraterrelevanter landwirtschaftlicher Sortenversuche

→ Methodenvergleich zur Adjustierung bei unbalancierten Daten

- absolute oder relative Differenzen
- kleinste Quadrate = FITCON
- gemischtes Modell mit gewichteten kleinsten Quadraten und REML-Schätzung (Maximum-Likelihood-Methode) → **hier verwendet !**
  - nutzt alle Informationen mit optimaler Gewichtung
  - ist ‚Methode der Wahl‘ für derartige Fragestellungen
  - ist Grundeinstellung in SAS-Prozedur MIXED
  - steht in PIAFStat zur Verfügung

#### weiteres zur Gewichtung der Anbaugebiete

Die Wichtung eines Nachbar-Anbaugebietes resultiert nicht ausschließlich aus der genetischen Korrelation. Sie lässt sich gedanklich aus 2 Komponenten herleiten:

- a) genetische Korrelation zum Zielgebiet: je größer, desto stärker gewichtet  
für alle Sorten gleich  
bei gleicher Besetzung der Sorten in den AG wäre nur  
a) relevant

b) Prüfumfang im Anbaugebiet:

je größer, desto stärker gewichtet  
für die Sorten/ Jahrgänge differenziert !!!

- relatives Verhältnis von  $N_{(\text{Nachbargebiet})}$  zu  $N_{(\text{Zielgebiet})}$ :  
sind z.B. in einem Nachbargebiet deutlich mehr Versuche als im Zielgebiet, so steigt die Wichtung des Nachbargebietes und es gibt eine konkurrierende Wirkung von a) und b);  
zu beachten ist, dass z.B. der Präzisionsschub z.B. von  $N=1$  auf  $N=4$  ungleich höher ist als z.B. von 4 auf 7 (degressive Zunahme der Präzision mit steigender Versuchszahl)  
Dies führt dazu, dass ein Nachbargebiet mit nur einem Ort/Jahr insgesamt zwar gering, je Ort aber relativ hoch gewichtet wird.
  - absolutes  $N$ :  
bei steigender absoluter Versuchs-Anzahl im Zielgebiet ist die Schätzung für das Zielgebiet zunehmend aus „eigener Kraft“ möglich und sinnvoll. Daher wird die Eigengewichtung bei zunehmender Anzahl Versuche je Sorte (z.B. bei langjährig geprüften Sorten) automatisch steigen, unabhängig davon, ob die Datenbasis der Nachbargebiete ebenso steigt.
- eine ausgewogene Besetzung der Anbaugebiete mit möglichst  $\approx 3$  Orten ist sehr wichtig!  
Der Ansatz, z.B. von einem Nachbargebiet nur den nächstliegenden Ort zu verwenden würde zu einer z.T. sehr hohen Gewichtung dieses Ortes führen, weil an ihn die Anforderung gestellt ist, ein ganzes Anbaugebiet zu repräsentieren. Das erscheint dann im Ergebnis z.T. nicht ganz nachvollziehbar, liegt aber an extremen Vorgaben des Anwenders.

Nähere Information: siehe Infotext zum Verfahren ‚VK‘

### ausgegebene Standardfehler SE

In diesem Verfahren werden die Schätzwerte für die Sorten bestimmt. Zusätzlich werden Standardfehler für den Vergleich der Sorten zum Mittel angegeben. Die angegebenen Standardfehler sind also keine Standardfehler für das absolute Niveau eines Sortenschätzwertes, sondern sie sind ausschließlich vergleichsbezogen. In diesem Sinne wären auch die aus den SE leicht ableitbaren Vertrauensintervalle zu interpretieren. SE und ‚Vertrauensintervalle für paarweise Vergleiche‘ sind ein sehr geeignetes Instrument zur Fehlerkritischen Ergebnisbewertung und Eigenkontrolle.

### Unterschiede der adäquaten SAS-Prozeduren gegenüber Tabellenkalkulation, Datenbank-abfrage, PIAF-Tabellierung etc.:

- diverse Rechen- und Entscheidungsregeln entfallen
  - z.B. WP: bis 15 Orte alle gleich gewichtet, darüber keine zusätzliche Wichtung
  - Umgang mit mehr oder weniger Fehlwerten
  - Umgang mit Versuchen unterschiedlicher Präzision
  - Umgang mit mehreren Versuchen an  $O \times J$  (WP-OV+LSV)
  - Umgang mit verschiedenen besetzten Jahren oder Orten
  - Wegfall einer starren Bezugsbasis, Sortimente brauchen nicht formell begründet werden, brauchen nur beratungsrelevante Sorten enthalten
- weniger aufwändig und anfällig bei komplexen Daten
- nur mit entsprechenden Prozeduren reproduzierbar, also nicht leicht nachrechenbar
- optimierte Schätzung für Grundgesamtheit
- einfach für den Nutzer
- Automatisierbar

### Adjustierung = Adjustment = Bereinigung ; adjusted means = adjustierte Mittelwerte

Wird bei Vorhandensein von Störfaktoren ein Parameter zunächst ohne Berücksichtigung dieser Effekte geschätzt und anschließend der Schätzwert so korrigiert, dass die Störeffekte auf den korrigierten Schätzwert keinen Einfluss haben, so spricht man von B. (Biometrisches Wörterbuch)

(auf eine einheitliche Ebene /Vergleichsbasis/mittleres Ertragsniveau projizieren)

### Warum ist komplette Balanciertheit nicht realistisch?

Sortenversuche sind:

- keine dreijährige Serie
  - sondern ein Kontinuum
  - und dabei ein Selektionsprozess der jährlich neu gespeist wird
- in diesem Prozess werden Sorten reduziert und Orte erweitert (WP1 bis LSV)

→ Balanciertheit ist immer ein Artefakt, eine künstlich zugeschnittene (reduzierte) Datenstruktur unter Verzicht auf relevante Daten

### Optimierung des SAS-Codes:

Durch Optimierung des SAS-Programms wurden Durchlauf und akzeptable Rechenzeiten auch bei extrem großen Datensätzen abgesichert:

- notest
- ddfm = residual
- subject- Variable
- nicht mit Sorte gekreuzte Umwelteffekte formal fix gesetzt
- nicht mit Sorte gekreuzter Umwelteffekt nur im höchsten Interaktionslevel aufgenommen.

### Gründe für die Trennung von Varianzkomponentenschätzung und Mittelwertbestimmung:

- VK mit langjährigerem Datensatz als Mittelwertschätzung
- Minimierung der Rechenzeit; für ‚VK‘ und ‚MW‘ gesondert optimiert
- zeitliche Trennung von VK- und Mittelwertschätzung
  - VK in Nebensaison bestimmen
  - VK ggf. nur periodisch schätzen und mehrjährig nutzen

### Bildung von Relativwerten

Relativwerte werden nicht innerhalb des Verfahrens gebildet. Die ausgewiesenen absoluten MW sind auf eine einheitliche Ebene /Vergleichsbasis /mittleres (Ertrags-)Niveau projiziert, sodass sie unmittelbar und optimal miteinander vergleichbar sind (Adjustierung).

Diese Absolutwerte können außerhalb des Verfahrens (z.B. in der ausgegebenen Excell-Tabelle) unproblematisch relativiert werden, indem eine sachlich begründete Auswahl von Sorten getroffen wird, deren Mittelwert die Bezugsbasis (100%) bilden soll. Dabei ist es nicht erforderlich, dass diese Sorten in allen Versuchen standen, was z.B. bei einer 6-jährigen Verrechnung kaum realisierbar wäre. In der Regel wird die Auswahl der Sorten aber so getroffen werden, dass diese in allen Jahren wenigstens enthalten waren und insgesamt häufig geprüft wurden. Abgestimmte Verrechnungssortimente von WP bis LSV und über die Gebiete bieten sich an.

### Vorteile der Einbeziehung von vorgelagerten Versuchen (WP ...) und mehr als 3 Jahren

- in MV Verdopplung der relevanten Datenbasis (in MV gab es viele WP! )
- Verbesserung der Klima-Repräsentanz durch z.B. 6-jährige Auswertung
- Verbesserung der regionalen Repräsentanz durch zusätzliche Standorte (BSA, Züchter-WP, ...)
- bessere Gesamtpräzision (wir sind bei unserer Datenstruktur jetzt nach dem ersten bis zweiten LSV-Jahr so schätzsicher wie nach alter Methode erst nach dem dritten)
- frühzeitigere Einstufung junger Sorten
- objektivere Einschätzung älterer, im aktuellen Sortiment nicht mehr geprüfter Sorten im Vergleich zur aktuellen Bezugsbasis

### weitere Diskussionspunkte zur Verrechnung unbalancierter Daten:

### Messen → Rechnen → Schätzen

Häufig wird den modernen Modellen vorgehalten, es würden Werte geschätzt, während die herkömmlichen Methoden gemessene Werte ausweisen:

Bei alten wie bei neuen Modellen werden ausschließlich gemessene Werte verarbeitet. Gemessene Werte sind nur die Einzelparzellenwerte. Mittelwerte im Versuch und erst recht über mehrere Versuche (Orte und Jahre) sind immer berechnete, nie gemessene Werte. Alte und neue Methoden unterscheiden sich allerdings in den Rechenalgorithmen. Die hier diskutierten neuen Modelle ermöglichen es, gemessene Werte auch von Versuchen einzubeziehen, die in den Sortimenten nicht völlig übereinstimmen, sondern nur orthogonale Überschneidungsböcke besitzen. Die Mittelwerte (Ismeans) werden natürlich nicht geschätzt, sondern ebenfalls nach exakt reproduzierbaren Rechenalgorithmen berechnet.

Der bei den Modellen häufig verwendete Ausdruck „Schätzen der Sortenrelationen“ meint etwas ganz anderes, das für alte wie neue Modelle gleichermaßen zutrifft (siehe nachfolgenden Punkt „Stichprobe und Grundgesamtheit“)!

### Selbstverständnis und Ziel der Auswertung (Stichprobe und Grundgesamtheit)

Die Theorie des landwirtschaftlichen Versuchswesens geht davon aus, dass die Versuchsergebnisse eine Stichprobe mit exakt zu berechnenden Mittelwerten darstellen, mit denen Parameter der Grundgesamtheit geschätzt werden sollen. Es ist a priori klar, dass diese Schätzung nicht „richtig“, „wahr“ oder „exakt“ im absoluten Sinne ist, sondern nur mehr oder weniger gut.

Bei der Ertragsauswertung von Serien von Sortenversuchen heißt das konkret: Man will die Ertragsrelation einer Sorte in einer Region schätzen: das mittlere Ertragspotential über alle Schläge und denkbaren Jahre (Grundgesamtheit). Die Stichprobe ist nur Mittel zum Zweck: eigentlich interessiert ja nicht ein Mittelwert von z.B. Gülzow/Tützpatz/Vipperow aus 2002/2003/2004 (Stichprobe). Unser Auftrag ist eine Praxis-Empfehlung für ein neues Jahr und keine rückwirkende Empfehlung für Versuchstationen.

Auf die Grundgesamtheit bezogen, sind auch vollständig balancierte Daten extrem lückig: was sind z.B. 4 Orte aus 3 Jahren bezogen auf alle Schläge und denkbaren Jahre in einer Region? Wenn also zusätzliche Jahre, Wertprüfungen etc. hinzugenommen werden und dadurch die Unbalanciertheit der Stichprobe steigt, so sinkt trotzdem die Lückigkeit auf die Grundgesamtheit bezogen.

### fixer oder zufälliger Ansatz für Jahre und/oder Orte (Umwelten, Versuche)

Beim fixen Modell sollten die Umwelten in der Weise ausgewählt sein, dass genau ihr Mittelwert eine übergeordnete Region repräsentiert. Hier wird folgerichtig jeder Versuch gleich gewichtet, unabhängig von der Präzision:  $SE = f_{(s\%; N)}$ . Zu beachten ist:

- für lückig besetzte Sorten ist dann kein Mittelwert sinnvoll / möglich
- aber auch der komplette Ausfall eines Versuches macht bei diesem Ansatz die Mittelwertbildung sinnlos, selbst wenn die Sorten in den verbleibenden Versuchen vollständig balanciert sind! Grund: Die Region ist nicht mehr repräsentiert.
- in allen Jahren müssen die gleichen Orte vollständig sein (siehe vorheriger Pkt.)
- bei S x Umwelt - Wechselwirkungen ist Mittelwertbildung nicht sinnvoll

Letzteres ist i.d.R. der Fall und auch die vorherigen Restriktionen sind real kaum einzuhalten. Dieser Ansatz kann bei anbautechnischen Versuchen im Einzelfall sinnvoll sein, im Sortenwesen kaum.

Beim hier verwendeten zufälligen Modell werden die einbezogenen Versuche als Stichprobe einer Region angesehen. Ziel sind Schätzwerte für die Region (siehe Grundgesamtheit), die Orte sind nur Mittel zum Zweck und (theoretisch) austauschbar. Das Fehlen eines Ortes ist (nur) ein gradueller Informationsverlust was die Repräsentanz der Serie nicht prinzipiell in Frage stellt. Auch die Wechselwirkungen S\*O wird als zufällige Stichprobe der typischen Interaktionsbreite in der Region angesehen. Der optimale Schätzwert für die Region ist Ergebnis einer Wahrscheinlichkeitsmaximierung. Bei extremen Ereignissen → extremen Wechselwirkungen oder Niveauunterschieden muss hinterfragt werden, ob einzelne Versuche die Grundgesamtheit repräsentieren (einbeziehen) oder nicht (nicht einbeziehen, gesondert auswerten).

### orthogonale / balancierte Daten

Orthogonale Kernstrukturen sind natürlich ausgesprochen wichtig für die Schätzgüte (Biometrisches Wörterbuch: *„Daten, die bezüglich eines Modells orthogonal sind, erleichtern die statistische Auswertung und Deutung der Versuchsergebnisse“*). Nur absolut orthogonale Auswertungen als exakt anzusehen, ist aber irreführend. Wenn die Stichproben-Daten orthogonal sind, wird bestenfalls die Stichprobe exakt gemittelt, die Grundgesamtheit wird trotzdem nur mehr oder weniger gut geschätzt.

Die Orthogonalitätsforderung für das mehrjährige Mittel ist schon dann nicht erfüllt, wenn in zwei Jahren die Versuchsanzahl unterschiedlich war!

Der Sortenversuch aus mehrjähriger Sicht - und nur diese ist beratungsrelevant – ist von Natur aus unbalanciert. Jedes Jahr kommen neue Sorten in die Versuche und alte werden herausgenommen. Die orthogonale Struktur ist ein Artefakt – eigentlich bereits erzeugte, finanzierte, ausgewertete, vollwertige Daten aus der Region aus zusätzlichen Jahren (Jahre sind ja immer der größte Mangelfaktor) werden ignoriert. Eine im Sprachgebrauch der „orthogonal Auswertenden“ *einjährig* geprüfte Weizensorte ist in D-Nord vorab bereits 11-mal in 4 Jahren geprüft worden. Nach dem ersten LSV-Jahr hat sie bereits 20 Versuche aus 5 Jahren und steht zur eingeschränkten Empfehlung an! In MV müssten wir ca. 50 % der beratungsrelevanten Daten ausschließen, um orthogonale Kerne auszuschneiden. Speziell bei Sorten im frühen Stadium würde sogar die weit überwiegende Anzahl Daten verworfen werden.

Man sollte z.B. bei einer einzubeziehenden WP mehr über die hinzugewonnenen Informationen begrüßen, als über die zwangsläufig entstehenden Lücken in einer künstlichen Tabellenstruktur diskutieren.

Die Verrechnung nach orthogonalen Mustern benötigt im Sortenwesen künstliche Strukturen, die der Sachfrage nicht gerecht werden (Sachfragen ordnen sich dann methodischen Fragen unter):

- Bei dreijährigen Auswertungen wird Schluss gemacht. Sonst wird es zu lückig. Welch ein Verlust! Wir wollen möglichst viele Orte, aber bei Jahren reichen uns 3 ?
- orthogonale einjährige Serien: Innerhalb des Jahres Orthogonalität anzustreben, ist prinzipiell sinnvoll, auch diesem Verfahren kommt das zugute. Aber: der Sortenversuch ist keine einjährige, sondern eine mehrjährige Problematik. In einjährigen Serien zu denken, führt vom Ziel weg. Es ist z.B. nahezu sinnlos, orthogonale Orientierungsversuche anzulegen, die von vornherein weder mit den vorgelagerten WP noch mit den nachgelagerten LSV gemeinsam ausgewertet werden sollen.
- Das parallele Ausweisen von 2-, 3- usw. -jährigen orthogonalen Mittelwertsspalten ist eine überflüssige Aufblähung. Warum soll für eine 4-jährig geprüfte Sorte noch ein zweijähriges Mittel ausgewiesen werden? Nur damit sie orthogonal mit zweijährig geprüften verglichen werden kann? Welche Relevanz für das Folgejahr liegt in einem zweijährigen Mittelwert einer 4-jährigen Sorte. Jede Sorte sollte so gut wie möglich eingeschätzt werden, eine 4 jährige also genauer als eine zweijährige. Dann wird auch automatisch der Vergleich zwischen beiden am wahrscheinlichsten bezüglich der Erwartung für das noch unbekannte Folgejahr.

#### Anmerkungen zu direkten (orthogonalen) und indirekten (adjustierten) Vergleichen:

vorab der Wortlaut zweier relevanter Gerichtsurteile aus dem Sortenzulassungsverfahren:

Urteil vom 2.2.84 –BverwG 3 C 75.82

*Der wertende Vergleich mit den zugelassenen Sorten setzt nicht voraus, dass diese mit der zu entscheidenden Sorte orthogonal mitangebaut wurden.*

*Grundlage für die Feststellung der aktuellen wertbestimmenden Eigenschaften der zum Vergleich herangezogenen Sorten sind die fortgeschriebenen Sortenlisten.*

Urteil vom 16.1.86 – BverwG 3 C 66.84

*Wurde eine entgegengehaltene Sorte ausnahmsweise orthogonal mitgeprüft, so haben ihre Werte aus der Prüfung Vorrang vor denen der BSL.*

Bei unbalancierten Daten basieren paarweise Sortenvergleiche nicht immer ausschließlich auf direkten Vergleichen (beide Sorten hätten dann nur in identischen Versuchen gestanden). Es fließen stattdessen auch Versuche ein, in denen nur eine von beiden gestanden hat. Voraussetzung dafür sind ‚Brücken‘, z.B. Verrechnungsblöcke etc., die den indirekten Vergleich über andere Sorten gestatten. Je größer und stabiler diese ‚Brücken‘ sind, desto mehr gewinnen auch indirekte Vergleiche an Aussagekraft. Die extreme Aussage: ‚Nur direkte Vergleiche sind wertbar‘ muss dahingehend abgeschwächt werden, dass direkte wertvoller als indirekte Vergleiche sind.

Dieses Prinzip wird durch dieses Verfahren optimiert. Für Sorten, die nur in isolierten Versuchen ohne Brücken gestanden haben, erfolgt keine Serien-Schätzung. Häufig geprüfte Sorten mit vielen direkten Vergleichen haben geringere Standardfehler der Differenzen. Der Adjustierungsprozess basiert vorrangig auf direkten Vergleichen und zusätzlich auf Brücken über indirekte Vergleiche.

Die Sortenbewertung, -empfehlung und -wahl ist ohne indirekte Vergleiche nicht denkbar. Wenn nicht im Rechengang, so doch überschlägig oder subjektiv-intuitiv.

1. Beispiel: Die Fortschreibung der BSL mit dem Vergleich etablierter Sorten zu Neuzulassungen wäre z.B. undenkbar.
2. Beispiel: Auch die Einordnung von Neuzulassungen basiert nicht ausschließlich auf dem Vergleich zu den VRS, sondern auch auf dem indirekten Vergleich über die VRS zu etablierten Sorten (s.o. Urteile BverwG).
3. Beispiel: Wenn folgende Signifikanzen bestehen:  $A > B$  und  $B > C$ , dann gilt auch  $A > C$ , selbst wenn A und C nur indirekt verglichen werden können. Dies gilt sogar dann, wenn die ersten beiden Aussagen über viele Versuche erhärtet sind, aber in einem einzelnen gemeinsamen Versuch von A und C die Überlegenheit von A nicht feststellbar war.
4. Beispiel: Wenn von drei Sorten (A, B und C) die Sorte C seltener als A und B geprüft wurde, so wäre nicht einzusehen, dass für den Vergleich A zu B nur Versuche genutzt werden, an denen auch C stand. Genau dieser Informationsverlust entsteht aber beim ‚Schneidern‘ von orthogonalen Strukturen

Fazit: Jede Sorte sollte so gut wie möglich geschätzt werden, dann werden auch die paarweisen Vergleiche optimal.

Natürlich bleiben die Grundaussagen bestehen:

- je größer die Anzahl Versuche (Umwelten) je Sorte, desto genauer
- je mehr direkte Vergleich, desto genauer
- je größer die orthogonalen Kerne (hier Brücken), desto genauer

### Spezifische Sorte \* Umwelt – Wechselwirkungen vernachlässigt ?

Ein Argument gegenüber dem Rechnen mit unbalancierten Daten ist häufig, dass spezifische Sorte\*Umwelt - Wechselwirkungen nicht adäquat berücksichtigt werden. Es ist aber ohnehin so, dass durch Mittelwertbildung spezifische Reaktionen oft verschleiert werden. Dies ist kein Problem der Balanciertheit, sondern der Mittelwertbildung schlechthin. Beim Ansatz mit fixen Umwelten war eine Mittelwertbildung bei sign. Wechselwirkungen generell nicht vorgesehen. Da Wechselwirkungen bei großen Serien immer enthalten sind, dürfte man nach diesem Ansatz praktisch gar nicht über Versuche mitteln!

Serien-Mittelwerte können also die Bewertung der Sorten in spezifischen Umwelten (z.B. nur Auswinterungs-Versuche zusammenfassen, Auswertung eines einzelnen Auswuchsversuches, Bewertung der Jahresreaktion z.B. in einem extremen Trockenjahr durch das einjährige Mittel etc.) nicht ersetzen, sondern nur ergänzen. Verzichtbar ist das Serienmittel zur Einschätzung der mittleren Ertragsrelationen zwischen den Sorten aber nicht. Wenn die Auswerter es nicht ausweisen, überlässt man dieses Problem dem Praktiker, der dies dann gefühlsmäßig oder per Taschenrechner selbst macht, und das wird keinesfalls besser. Versuche, die absolut nicht zur Grundgesamtheit passen, sind ggf. nicht einzubeziehen bzw. gruppiert auszuwerten. Das ist allerdings genau abzuwägen, denn z.B. Trockenstress tritt mit gewisser  $\pm$  geringer Wahrscheinlichkeit ein und gehört damit ggf. zur Grundgesamtheit. Wenn einzelne Sorten in solchen Versuchen fehlten, ist das natürlich ein Informationsverlust, genauso wie es ein Informationsverlust ist, wenn eine ganze Zulassungsgeneration noch nicht in der Reaktion auf Extremtrockenheit eingeschätzt werden kann. Das ist aber kein ursächliches Problem der Balanciertheit, sondern der begrenzten Zahl Versuche, Jahre etc. Dies spricht gerade für eine möglichst langjährige Auswertung und nicht für die Eingrenzung der Daten nur aus Balanciertheitsgründen.

### Rechnen mit unbalancierten Daten

Inwieweit das Rechnen mit unbalancierten Daten zu Problemen führt, hängt von der Art und Weise des Umgangs mit den Daten ab. Es kann nicht bestritten werden, dass unbalancierte Daten Probleme mit sich bringen können. Insbesondere ist bei extremen Differenzen im Ertragsniveau zwischen den Versuchen zu überprüfen, ob diese Versuche zu einer Grundgesamtheit gehören und ob die Additivitätsunterstellung hinreichend erfüllt ist. Aber Analysen (Piepho und Michel, 2001) zeigen ganz deutlich, dass bei adäquater Modellbildung (!) die Vorteile bei weitem überwiegen.

Wo liegt der eigentliche Unterschied zwischen beiden Methoden? Beide verrechnen reale, gemessene Daten und beide schätzen die Grundgesamtheit. Eigentlich unterscheiden sie sich nur in Rechenalgorithmen und -regeln. Das eine ist mit dem Taschenrechner oder Tabellenkalkulation, das andere nur noch mit Matrizenrechnung nachzurechnen. Beide sind aber im Ergebnis exakt reproduzierbar.

Ist dieses leichte Nachrechnen - können wirklich so wichtig?

## Nachrechenbarkeit

Nachrechenbarkeit wird oft fälschlich mit Reproduzierbarkeit synonym gesetzt. Reproduzierbarkeit bedeutet aber nicht, dass ein Ergebnis mit Excel oder Taschenrechner leicht nachrechenbar ist. Reproduzierbarkeit lässt sich nicht an einer Rechenmethode messen sondern meint die Übertragbarkeit in die Praxis. Sie misst sich also an dem unbekannten 'wahren' Wert in der Grundgesamtheit.

Die Vorstellung, alle Rechengänge müssten mit dem Taschenrechner leicht nachzuvollziehen sein, würde die LDS Deutschlands von jeglicher Modernisierung abkoppeln (*andere Staaten sind inzwischen weit voraus; auch das BSA könnte keine BSL herausgeben; oder wie würde z.B. unsere Architektur aussehen, wenn die Forderung bestünde, jede Statik mit Taschenrechnern leicht nachrechnen zu können*). Verbesserungen werden nach unserer Erfahrung sowohl in den Fachkreisen als auch in der Praxis schnell akzeptiert und als Vorteil anerkannt. Im Einzelfall „kritisiert“ natürlich ein Beteiligter schon mal, dass nun gerade seine Sorte z.B. bei 2-jähriger Verrechnung ohne WP besser abgeschnitten hätte als bei 5-jähriger mit WP.

Auch in vorangegangener Software (z.B. EFDAS) gab es schon nicht einfach nachzurechnende Adjustierungen, z.B.:

- Gitteranlagen (Intra- oder Interblockschätzung)
- Fehlstellenausgleich (leider ohne Berücksichtigung der Fehlwerte in der VA)
- Standardausgleich in Langparzellenanlagen

Alle diese Adjustierungen können nachgerechnet und exakt reproduziert werden, nur eben nicht mit dem Taschenrechner oder Excel-Tabellenkalkulationen.

## Praxisakzeptanz

Die Ergebnisse müssen natürlich praxisrelevant und nachvollziehbar sein, nur dann können sie auch dem Praktiker vermittelt werden. Die Argumentation, dass der Praktiker seine Wahl nicht nur auf den nächstgelegenen Versuchsstandort und nicht auf einjährige Daten begründen sollte, sondern auf eine für die Region erhärtete Zusammenfassung, muss ohnehin immer erfolgen – hier ist Überzeugungsarbeit ein ‚Dauerauftrag‘.

Akzeptanz sollte sich aber nicht daraus herleiten, dass (1) alle Werte selbst ausgerechnet werden könnten oder (2) das gewisse formale Regeln eingehalten werden (z.B. nur und genau 3-jährige Mittel).

Nach unserer Erfahrung besteht in der Umstiegsphase ein Akzeptanzproblem überwiegend bei den Dienststellen selber, erst in zweiter Linie bei Züchtervertretern und kaum bei den Praktikern. Entscheidend ist die eigene Erfahrung und Argumentation, dass die Empfehlungen stabiler und sicherer werden. Mit fehlender Akzeptanz hatten wir bei Praktikern nie Probleme, im Gegenteil: Praktikern ist nicht plausibel zu machen, dass nur balancierte Daten einfließen dürfen und wertvolle regionale Versuche aus formalen Gründen nicht einbezogen werden. Züchtervertreter haben es natürlich schwerer mit dem Nachrechnen, sie können nur noch beurteilen, ob die Werte unter Berücksichtigung von WP etc. in plausiblen Grenzen bleiben. In den ersten Jahren wurde aus dieser Richtung in einigen Einzelfällen die Methode angezweifelt - in der Regel dann, wenn ihre Sorte durch WP-Einbeziehung schlechter wurde, also ohne konkretes fachliches Argument. (Am ehesten stehen sich die Dienststellen also noch selbst im Wege.)

Die Einführung und der Transfer in die Praxis setzt natürlich die eigene Akzeptanz und produktive Mitwirkung voraus, sonst sind wie überall Fehlentwicklungen vorprogrammiert.

## gemessene Werte und arithmetische Mittelwerte als Schätzer

Folgende Betrachtungen sollen nicht überexakt werden, sondern nur noch einmal untersetzen, dass die Denkweise „*Arithmetischen Mittelwerten haftet ‚Wahrheit‘ an, die bei Unbalanciertheit und Adjustierung verloren geht.*“, irreführend ist.

Häufig wird fälschlicherweise die Kategorie ‚Mittelwert‘ synonym mit dem arithmetischen Mittel gesetzt und diesem das Attribut ‚wahr‘ zugesprochen. Aber auch die Schätzformel für arithmetische Mittel ist eine reine Rechenanweisung, abstrakt und nicht logischerweise und automatisch richtig/wahr/exakt. Wirklich gemessen ist im Versuchswesen nur ein Parzellenwert.

Ein Mittelwert ist bereits ein Schätzer für den Parameter der Lage und als solcher nur mehr oder weniger gut, niemals aber ‚wahr‘. Es gibt viele andere Rechenanweisungen für die Mittelwertschätzung: Median-/Zentralwert, Modalwert, harmonisches Mittel, geometrisches Mittel, quadratisches Mittel .... arithmetisches Mittel. Letzteres ist nicht immer der beste Schätzer. Arithmetische Mittel

setzen z.B. Normalverteilung und Verhältnisskaliertheit voraus. Wer's exakt nehmen möchte: für Bonituren wären arithmetische Mittel oft nicht adäquat. In anderen Fällen wird das arithmetische Mittel erst nach Transformation zur geeigneten Methode (z.B. Prozentzahlen).

## Literatur

- Atkinson, A.C. 1985. Plots, Transformations and Regression. Oxford University Press, New York
- Beese, G., G. Barthelmes, G. Hartmann, U. Jentsch, V. Michel. 2005. Sortenprüfung jetzt effizienter. Bauernzeitung 46/21, 27-28
- Gilmour, A.R., B.R. Cullis, S.J. Welham, R. Thomson. 1999. ASREML Reference Manual. NSW Agriculture Biometrie Bulletin No.3. NSW Agriculture, Locked Bag 21, NSW, 2800, Australia, 210ff.
- Piepho, H.-P., V. Michel. 2001. Überlegungen zur regionalen Auswertung von Landessortenversuchen. Informatik, Biometrie und Epidemiologie in Medizin und Biologie 31/4, 123-139
- Michel, V., H.-P. Piepho. 2001. Ertragsauswertung der Sortenversuche in Mecklenburg-Vorpommern. Sommertagung der AG Landwirtschaftliches Versuchswesen der Biometrischen Gesellschaft
- Michel, V., 2003. Neuster methodischer Stand bei der Versuchsauswertung in Mecklenburg-Vorpommern. VII. Rapskolloquium, Schleswig-Holstein/Mecklenburg-Vorpommern
- Michel, V., Pienz, G. 2006. Bericht Landessortenversuche Winterweizen, [www.agrarnet-mv.de](http://www.agrarnet-mv.de)
- Möhring, J., A. Büchse, H. Piepho, V. Michel, J. Rath, F. Laidig. 2004. Gesundsparen ohne Nachteile. DLG-Mitteilungen 6(2004)
- Press W.H. 1989. Numerical recipes in Pascal. Cambridge University Press
- Roßber, D., Michel, V., Graf, R., Neukampf, R. 2006. Definition von Boden-Klima-Räumen für die Bundesrepublik Deutschland, . Nachrichtenblatt des Deutschen Pflanzenschutzdienstes. zur Zeit im Druck
- SAS Institute, Inc., SAS/STAT *User's Guide*, version 8, Cary, NC, USA

## Berichte

- Bericht Prüfsystem Winterweizen 24.12.2003. doc
- Prüfsystem Winterweizen 19.03.04. doc
- Prüfsystem Winterweizen 24.08.04.doc
- Prüfsystem WW 24.08.04\_Michel.doc