

Die Hohenheim-Gülzower-Serienauswertung als bundesweites Basisverfahren im regionalisierten Sortenwesen

Volker Michel, Andrea Zenk, Jens Möhring¹, Andreas Büchse¹, Hans-Peter Piepho¹

Einleitung

Landessortenversuche sind die Grundlage regionaler Sortenempfehlungen in Deutschland. Eine Schwerpunktaufgabe besteht in der möglichst frühzeitigen und schätzgenauen regionalen Ertrageinstufung von Sorten. Gerade die Bedeutung der Frühzeitigkeit steigt bei derzeit beschleunigtem Sortenwechsel. Die hergebrachte regionale Ertragsauswertung der meisten Bundesländer bezog Versuchsergebnisse der Landessortenversuche (LSV) beschränkt auf die jeweilige Länderdienststelle (LDS) und die letzten drei Jahre ein. Diese Beschränkung auf LSV, auf orthogonale Sortenvergleiche innerhalb eines LSV-Jahrganges sowie auf eine Regelprüfzeit von 3 Jahren ist im Sinne der o.g. Zielstellung nicht prinzipiell begründbar und nicht optimal. Im Interesse der Frühzeitigkeit und Genauigkeit ist es sinnvoll, die Datenbasis weiter zu fassen.

Für eine ausreichende Genauigkeit wird ein entsprechender Versuchsumfang benötigt. Aufgrund der Sparzwänge wurde und wird die Anzahl der Versuche z.T. deutlich reduziert, insbesondere bei den weniger bedeutsamen Pflanzenarten. Über das neue biometrische Verfahren der von der Universität Hohenheim und der Landesforschungsanstalt für Landwirtschaft und Fischerei Mecklenburg-Vorpommern (LFA) entwickelten ‚Hohenheim-Gülzower Serienauswertung‘ soll die regionale Auswertung verbessert werden.

Dabei wird die Datenbasis in folgender Weise erweitert:

- Durch die umfassende Einbeziehung der vor der Zulassung durchgeführten Wertprüfungen (WP) des BSA und anderer Officialprüfungen in die regionale Auswertung sollen unter Nutzung der neuen Verfahren frühzeitiger verlässliche regionale Aussagen möglich werden.
- Es werden mehr als drei Versuchsjahre einbezogen. Dies hat neben der allgemeinen Erweiterung des Datenumfanges besonders auch den Vorteil, dass dadurch die i.d.R. große Sorte*Jahr-Interaktion berücksichtigt und damit die klimatische Repräsentativität verbessert wird.
- Versuche aus benachbarten Bundesländern im standortkundlichen Überlappungsbereich werden einbezogen.

Optimierungen durch die Hohenheim-Gülzower-Serienauswertung

Die Auswertung langjähriger Sortenversuche beinhaltet eine Vielzahl systemimmanenter Besonderheiten, die in ihrer Kompliziertheit und Komplexität durch herkömmliche Auswertungen z.B. mit Methoden der Datenbankabfrage und Tabellenkalkulation nicht bzw. erst nach willkürlicher, umfangreicher Datenreduktion sachlich adäquat bewältigt werden können. In der Hohenheim-Gülzower-Serienauswertung wurden die Realitäten der Prüfsysteme in gemischten Modellen (PROC MIXED in SAS) abgebildet. Dadurch konnten unter Anderem folgende Besonderheiten des Sortenwesens berücksichtigt und optimal verwertet werden:

Unbalanziertheit:

- Ergebnisse der den LSV vorgelagerten Versuchen (WP, EUSV u.a.) können trotz großer Unbalanziertheit einbezogen werden.
- Trotz der bei langjährigen Daten großen Unbalanziertheit können viele (alle fachlich noch relevanten) Versuchsjahre einbezogen werden.
- Bezüglich einer Sorte schwach besetzte Jahre (z.B. WP-Jahrgänge) werden mit optimalem Gewicht in die Gesamtserie einbezogen.
- Ergebnisse von unterbesetzten Orten (z.B. Züchterstandorte) werden mit optimalem Gewicht in die Gesamtserie einbezogen.
- Bei parallelen Versuchen am gleichen Ort und Jahr (z.B. LSV und WP) werden die mehrfach und die nur einfach geprüften Sorten jeweils optimal in die Serie einbezogen.
- Die Gesamtauswertung erfordert keine absolut durchgängigen Verrechnungssorten oder orthogonalen Blöcke.

¹ Universität Hohenheim, Fachgebiet Bioinformatik

Regionalität: Ergebnisse aus benachbarten Regionen im standortkundlichen Überlappungsbereich können mit optimal abgestuftem Gewicht in die Auswertungen für ein Zielgebiet einfließen.

Differenzierte Versuchspräzision:

- Versuche mit erhöhtem Versuchsfehler (z.B. s%), soweit grundsätzlich ‚brauchbar‘, können mit optimaler Gewichtung einbezogen werden.
- Versuche mit unterschiedlicher Wiederholungsanzahl können mit jeweils optimaler Gewichtung einbezogen werden.
- Prüfglieder mit abweichender Wiederholungsanzahl im Einzelversuch (Fehlstellen) können mit optimaler Gewichtung einbezogen werden.
- Prüfglieder mit unterschiedlichem Standardfehler im Einzelversuch aufgrund unbalanzierter Blockstrukturen, geostatistischer Bodenausgleichsrechnung u.a. versuchsmethodischer Ansätze können mit optimaler Gewichtung einbezogen werden.

Nichtadditives Datenverhalten: wird durch eine entsprechende Transformation der Daten berücksichtigt.

In ihrer Gesamtheit bewirken diese Optimierungen die Minimierung des Vorhersagefehlers für Sortenleistungen in einem neuen Jahr auf Schlägen im Zielanbauggebiet.

Überlappende überregionale Auswertung und abgestufte Gewichtung

Die Einbeziehung von Versuchen aus agrarökologisch ähnlichen Nachbar-Anbaugebieten in die Auswertung für ein Ziel-Anbauggebiet ist eine Möglichkeit den Vorhersagefehler zu verringern. Zu diesem Zweck haben die Länderdienststellen für das Sortenwesen Deutschland in Boden-Klima-Räume und kulturartspezifische Anbaugebiete unterteilt. Diese Regionalisierung orientiert sich nicht mehr an politischen Grenzen, sondern an standortkundlichen Gegebenheiten und deren Relevanz für sortenspezifische Reaktionen.

Wertet man mehrere Anbaugebiete gemeinsam aus, so erhöht sich die Anzahl der Versuche und die Genauigkeit für diesen ‚Großraum‘ steigt. Andererseits führt die Einbeziehung von Versuchen aus Nachbargebieten zu einer gewissen Verzerrung für das Zielgebiet, da sich im Nachbargebiet möglicherweise eine andere Sortenrangfolge zeigt als im Zielgebiet. Ist diese Verzerrung größer als der Genauigkeitsgewinn durch die erhöhte Versuchsanzahl, so wäre die Einbeziehung der benachbarten Anbaugebiete kontraproduktiv. Ist die Verzerrung dagegen kleiner, so ist die Aussage aus den gemeinsam ausgewerteten Anbaugebieten trotz einer gewissen Verzerrung besser als die alleinige Auswertung des Zielgebietes mit zu wenig Versuchen.

Um diesen Kompromiss zwischen ausreichend vielen Versuchen und der Verzerrung optimal zu gestalten, werden die Ergebnisse aus Nachbargebieten geringer gewichtet, als Ergebnisse des Zielgebietes. Der Grad dieser Wichtungsabstufung wird automatisch in der Weise vorgenommen, dass der Vorhersagefehler im Ziel-Anbaugebiet minimiert wird, also die Sortenbewertung im Ziel-Anbaugebiet optimiert wird. Die Gewichte werden maßgeblich durch die genetische Korrelation der Anbaugebiete bestimmt. Diese ist ein Maß für die Ähnlichkeit der Sortenrelationen zwischen Gebieten. Dabei wird der Vorhersagefehler je Sorte im Zielgebiet, der sich als Summe von Prüfgenauigkeit und Verzerrung ergibt, minimiert. 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.

Bundesweit führt diese Methode zu gleitend überlappenden Anbaugebieten. So werden z.B. für das Anbaugebiet D-Nord (überwiegend MV) Daten aus D-Süd und Ostholstein, für das Anbaugebiet Ostholstein wiederum Daten aus D-Nord, nicht aber aus D-Süd genutzt usw.

Die seit 2005 von der LFA praktizierte Verrechnung mit diesen Modellen hat bestätigt, dass durch die Einbeziehung der WP und benachbarter Anbaugebiete der Vorhersagefehler erheblich verringert werden kann. Damit wurden Effizienzreserven erschlossen und die frühzeitige Bewertung neuer Sorten verbessert.

Die Hohenheim-Gülzower-Serienauswertung ist gleichzeitig ein geeignetes Instrument für die Evaluierung der konkreten Größe und Begrenzung von Anbaugebieten und vor allem der Überlappungsräume.

Das System der flexibel überlappenden Auswertungsräume wird nur dann langfristig funktionieren, wenn jedes Anbaugebiet eine verlässliche Anzahl an Versuchen pro Jahr beisteuert. Zur Absicherung dieses Erfordernisses werden derzeit überregionale und regionale Vereinbarungen zwischen den beteiligten Partnern (Länder, Bundessortenamt, Züchtungswirtschaft) getroffen.

Was bedeutet dies für den Landwirt?

Wenn Informationen aus benachbarten Anbaugebieten herangezogen werden, so steigt im Mittel der Abstand (geografisch bzw. standortkundlich) zum eigenen Betrieb. Das könnte zu dem Schluss führen, dass die Verlässlichkeit der Sortenempfehlung für den konkreten Betrieb deswegen geringer ist. Dies ist aber eine Fehleinschätzung, die aus der häufig zu beobachtenden intuitiven Überschätzung des Verzerrungsfaktors und Unterschätzung des Genauigkeitsfaktors her rührt. Die Aussagekraft eines einzelnen Versuches wäre selbst dann gering, wenn dieser auf Flächen des eigenen Betriebes liegen würde. Neben der selbst innerhalb kleiner Räume relativ hohen Differenziertheit der Böden sind insbesondere die unvorhersehbaren Bedingungen des Folgejahres dafür ausschlaggebend, dass aussagekräftige Sorteneinschätzungen sich erst durch umfangreiche Versuchsserien über Orte und Jahre erarbeiten lassen.

Beschluss der Agrarministerkonferenz vom 7.10.04 zu diesen Methoden

„... Die kombinierte Auswertung aller Daten aus Wertprüfung und Landessortenversuchen auf Basis der so genannten „Hohenheimer Methodik“ sichert die optimale Nutzung der Daten für die Sortenberatung und die Fortschreibung der beschreibenden Sortenlisten. Außerdem gewährleistet das System, dass Landwirten und Saatgutwirtschaft erste – auch regionalisierte – Sorteninformationen schneller zur Verfügung stehen und die LDS Informationen zur anbaugebietsspezifischen Auswahl von Sorten für die Landessortenversuche erhalten.“

Zur optimalen Umsetzung der Modelle wird im Beschluss der Agrarministerkonferenz auf die Nutzung des gemeinsamen Planungs-, Informations- und Auswertungssystems für das Feldversuchswesen (PIAF) verwiesen.

Die PIAFStat-Verfahren der Hohenheim-Gülzower Serienauswertung

Die Verfahrensgruppe Hohenheim-Gülzower-Serienauswertung wurde in Kooperation der Universität Hohenheim und der LFA in 3 PIAFStat-Verfahren realisiert (Tab. 1). Dabei haben die Verfahren ‚PHI‘ und ‚VK‘ eine vorbereitende Funktion für die Optimierung der Schätzung der Sortenleistungen. Sie müssen nicht jährlich und nicht im Hauptauswertungszeitraum ablaufen. Das Verfahren ‚MW‘ erfüllt dann die zentrale Aufgabe des Verfahrenskomplexes – die optimierte Schätzung der Sortenleistungen und deren Schätzgenauigkeit.

Tab. 1: Die PIAFStat-Verfahren der Hohenheim-Gülzower-Serienauswertung

Reihenfolge Abarbeitung	Kurzbezeichnung in PIAFStat	Langbezeichnung Label in PIAFStat
1	PHI	Optimale Datentransformation
2	VK	Bestimmung der Varianzkomponenten
3	MW	Bestimmung der Mittelwerte

Das Verfahren ‚PHI‘ – Optimale Datentransformation

In den Serienauswertungen der Länderdienststellen wird die Erfüllung der Modellvoraussetzungen derzeit i.d.R. nicht geprüft, sondern als gegeben vorausgesetzt. Das Verfahren ‚PHI‘ behebt dieses Defizit durch die Suche nach einer optimalen Transformation der Daten.

Das Verfahren PHI gewährleistet insbesondere hinsichtlich des Komplexes Additivität / Varianzhomogenität / Normalverteilung, dass gegebene Daten durch Transformation optimal an die

Modelle angepasst werden. Damit werden Verzerrungspotentiale auch bei der Mittelwert-schätzung reduziert. Es sei darauf hingewiesen, dass die Datentransformation alleine nicht garantiert, dass alle Modellvoraussetzungen optimal erfüllt sind. Zur Bestimmung einer optimalen Datentransformation wird in diesem Verfahren PHI der Transformationsparameter (ϕ bzw. φ) geschätzt. Hierzu wurde die Box-Cox-Transformation gewählt, deren herausragen-der Vorteil die stufenlose Anpassung ist.

Für die Box-Cox-Transformation gilt:

$$y = \begin{cases} \frac{y^\varphi - 1}{\varphi} & \text{für } \varphi \neq 0 \\ \log(y) & \text{für } \varphi = 0 \end{cases}$$

wobei der Transformationsparameter φ nach der Maximum-Likelihood-Methode geschätzt wird. Als Ausgabe wird der optimale Parameter ‚phi‘ ausgegeben und eine Graphik erstellt (Abb. 1). Diese zeigt die Güte des Modells (die sog. Log-Likelihood – loglik) als Kurve in Abhängigkeit vom Transformationsparameter (phi). Der optimale Wert von phi (maximaler Log- Likelihood) weist die beste Anpassung aus. Zur Signifikanzprüfung und Ausweisung eines Vertrauensinter- valls für den geschätzten Parameter wird eine horizontale Linie angezeigt. Die Werte auf der x- Achse, bei welchen die Linie die Log-Likelihood-Kurve schneidet, sind die 95%-Vertrauens- grenzen für phi. Interessant ist insbesondere, ob der Wert 1 für den Transformationsparameter (gleich untransformiert bzw. Additivität) sich vom optimalen Wert unterscheidet.

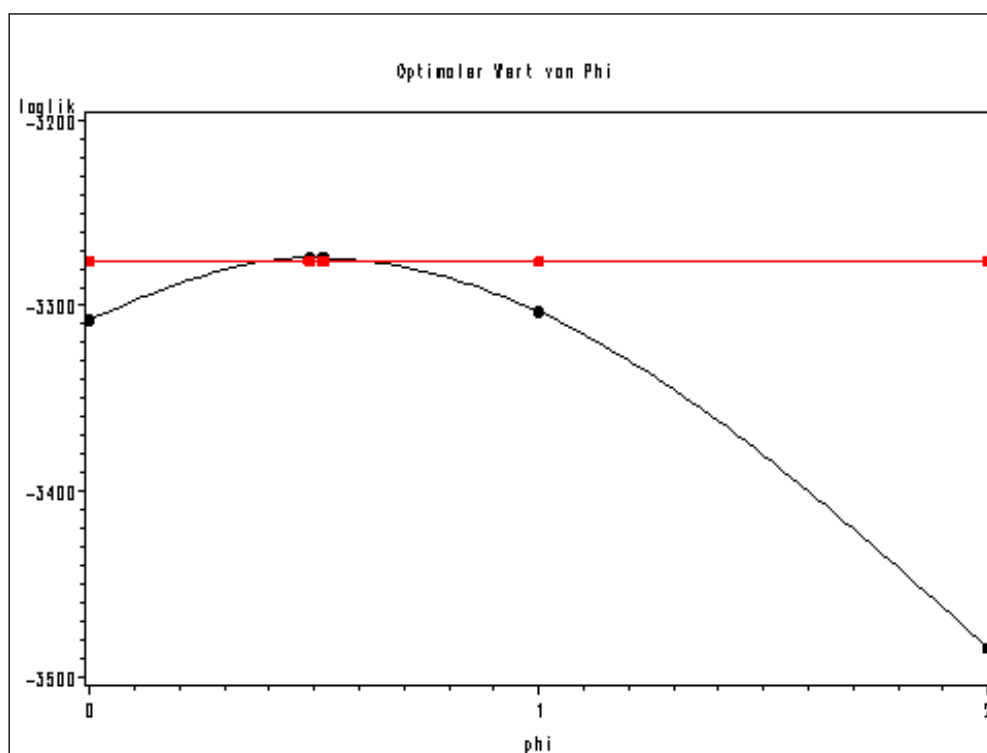


Abb. 1: Schätzung des Transformationsparameter ‚phi‘ (Beispiel)

Der optimale Wert für den Parameter phi sollte dann in den Verfahren ‚VK‘ und ‚MW‘ eingesetzt werden. Hierbei kann sich z. B. ergeben, dass eine logarithmische Transformation optimal ist ($\phi=0$), dass die Daten untransformiert ausgewertet werden können ($\phi=1$) oder dass eine Zwischenform optimal ist.

Warum ist Additivität besonders hinterfragt ? - Diskussion von Hintergründen

Mit der Ausweitung der Auswertungsräume durch das Überlappungsprinzip geht eine zu- nehmende Ertragsdifferenziertheit der Versuchsdurchschnittserträge einher. Dies hatte den Bearbeiter bewogen, a) die Richtigkeit der Additivitätsannahme im Allgemeinen zu hinterfragen und b) ein Werkzeug zu schaffen, mit dem in konkreten Serien-Auswertungen Abweichungen

ggf. berücksichtigt werden können. Bisher wurde die Frage vernachlässigt, ob das Datenverhalten überwiegend additiv oder multiplikativ ist – es wurde ungeprüft nach additivem Modell gerechnet und dazu im Widerspruch ebenso ungeprüft nach multiplikativem Modell dargestellt (Relativwerte). Die Darstellung von Relativzahlen erfolgte nicht wegen der Unterstellung eines multiplikativen Modells, sondern wegen der erleichterten intuitiven Tabelleninterpretation (z.B. bei der Darstellung von Sorten-Erträgen in verschiedenen Umwelten mit unterschiedlichem Durchschnittsertrag) und wegen der Gewöhnung an eine einheitliche Zahlendimension um 100 % unabhängig von der Mengeneinheit (dt, MJ, Anzahl ...). Dieser Widerspruch zwischen Verrechnung (additiv) und Darstellung (multiplikativ) ist nicht sinnvoll bzw. gerechtfertigt. Eine Überprüfung ist in kleinen Datensätzen aus Einzelversuchen bzw. kleinen Serien in der Regel nicht hinreichend möglich, wird aber bei mehrjährigen Auswertungen großer Serien von Sortenversuchen interessant und nützlich. Eine unberücksichtigte Verletzung der Modellvoraussetzungen führt zu verzerrten Schätzwerten für die Grundgesamtheit, und zwar sowohl bei vollständig als auch bei unvollständig balanzierten Daten.

Zum vereinfachten Verständnis der Problematik kann man folgende Fragen stellen:

Sind z.B. die zu erwartenden Sortendifferenzen [dt/ha] in neuen Versuchen

- a) unabhängig vom Ertragsniveau ? → additives Datenverhalten
- b) proportional zum Ertragsniveau ? → multiplikatives Datenverhalten

Für den additiven Auswertungsansatz spricht vermeintlich, dass letztendlich Mengen (dt) und nicht Prozente erzeugt werden sollen. Es muss aber hinterfragt werden, welche Verallgemeinerungsfähigkeit nachfolgende unterschiedlichen Aussagen für die Praxis, also für einen Schlag /eine neue Jahr*Ort-Kombination haben:

- a) Sorte X war im Mittel gegebener Versuche 8 dt/ha überlegen
- b) Sorte X war im Mittel gegebener Versuche 10% überlegen

Wenn hier die Aussage in Richtung b) tendiert, so müssten zumindest relevante multiplikative Anteile vorliegen und es bestünde dann ein Transformationserfordernis bzw. die Notwendigkeit, ein multiplikatives Modell anzupassen. Der Standpunkt, der additive Ansatz wäre Konvention und daher ‚richtig‘ wird der Frage: „Wie erreicht man mit der Stichproben-Auswertung die bestmögliche Verallgemeinerung für die Grundgesamtheit (regionale Empfehlung für ein neues Jahr)?“ nicht gerecht. Die Antwort hängt vom Datenverhalten ab: Ist es additiv ($\phi=1$), so besteht kein Transformationserfordernis. Ist es multiplikativ ($\phi=0$) so ist die log-Transformation optimal. Bei $\{0<\phi<1\}$ ist ein stufenloser Zwischenwert optimal → umsetzbar durch die unten beschriebene und im Verfahren integrierte Box-Cox-Transformation.

Wie kann auf verletzte Additivität reagiert werden ?

Die verwendeten Verfahren arbeiten mit linearen Modellen, setzen also Additivität der Effekte voraus. Ist dieses nicht gegeben, sollten Daten so transformiert werden, dass die transformierten Daten additives Verhalten aufweisen. Mit den transformierten Daten können dann die Verfahren optimal arbeiten. Um im Ergebnis keine abstrakten, dimensionslosen Daten zu erhalten, sondern Daten in der üblichen metrischen Einheit, erfolgt hier zur Ergebnisausgabe im Verfahren ‚MW‘ eine Rücktransformation.

Die varianzstabilisierende Wirkung einer Datentransformation hängt im Wesentlichen von der funktionalen Beziehung zwischen Varianz und Erwartungswert ab (Tab. 2). Die bisherigen Erfahrungen aus der Anwendung der Methode zeigen, dass die Daten in der Regel nahe am additiven Verhalten sind, aber häufig eine leichte Abweichung in Richtung multiplikativen Verhaltens besteht. In fast allen Fällen waren die Modelle für die transformierten Daten signifikant besser als für die untransformierten Daten.

Tab. 2: Datenverhalten und Box-Cox-Transformation

geschätztes phi	verbale Interpretation von Datenverhalten und Transformation
$\varphi \approx 1$	rein additiv; Auswertung identisch zu untransformierten Daten; die Varianz ist konstant und unabhängig vom Erwartungswert; siehe (a)
$0,5 \ll \varphi \ll 1$	näher am additiven als am multiplikativen Verhalten
$\varphi \approx 0,5$	Übergang; identisch zur Wurzeltransformation
$0 \ll \varphi \ll 0,5$	näher am multiplikativen als am additiven Verhalten
$\varphi \approx 0$	rein multiplikativ \rightarrow identisch zur log-Transformation; die Varianz ist proportional zum Erwartungswert; (entspricht etwa einer Auswertung über Relativwerte) siehe (b)
$(\varphi > 1)$	Vorschlag, wenn 1 im Konfidenzintervall liegt: mit $\varphi = 1$ arbeiten und möglichst die Gewichtung nach SE in ‚VK‘ und ‚MW‘ nutzen Additivität könnte z.B. überlagert sein durch einen „Präzisionseffekt“ (je größer die Merkmalsausprägung, desto kleiner die Varianz)

Das Verfahren ‚VK‘ - Bestimmung der Varianzkomponenten

Das Verfahren ‚VK‘ dient der Schätzung der zur Mittelwertsschätzung benötigten Varianzkomponenten incl. der davon abgeleiteten genetischen Korrelationen zwischen einbezogenen Anbaugebieten. Die Schätzung der genetischen Korrelationen ist erforderlich für die optimale Gewichtung von Ziel- und Nachbar-Anbaugebieten.

Wozu Varianzkomponenten schätzen?

Das Leistungspotential einer Sorte kann nicht fehlerfrei ermittelt werden, sondern wird außer von den Einzelversuchsfehlern besonders von Wechselwirkungen mit der Umwelt beeinflusst:

Die Sortenrangfolgen bzw. Sortenrelationen sind von Versuch zu Versuch nicht identisch, ebenso nicht von Jahr zu Jahr, von Ort zu Ort, von Region zu Region. Dieses Verfahren analysiert die Variationsursachen, quantifiziert die Varianzkomponenten, erlaubt den Vergleich ihrer Wirkungsstärke und ermöglicht nachfolgend insbesondere die Berücksichtigung bei der Sortenleistungsschätzung im Verfahren ‚MW‘.

Bei der hier vorgeschlagenen vollen Ausnutzung der verfügbaren Datenbestände ist die Besetzung der einzelnen Sorten in den einzelnen Umwelten z.T. extrem differenziert. Es stellt sich die Frage, wie man vor dem Hintergrund a) dieser differenzierten Besetzung und b) der Sorte*Umwelt-Wechselwirkungen die Daten optimal verdichten soll. Hier führt das gemischte Modell zu einem optimalen Vorgehen. Bezüglich der Berücksichtigung unterschiedlicher Besetzungen in Umwelten [Jahren; Orten etc.] kann vereinfacht gesagt werden, dass ein Ergebnis aus einer zusätzlichen Umwelt nützlicher ist und entsprechend höher gewichtet wird, als aus einer bereits gut besetzten, denn Ergebnisse aus dem gleichem Jahr, Ort etc. sind korreliert und bringen einen geringeren Informationsgewinn als Ergebnisse aus zusätzlichen Jahren, Orten etc.

Im Gegensatz zu herkömmlichen Verrechnungen mit Methoden der Tabellenkalkulation, bei der diverse Rechen- und Entscheidungsregeln zu beachten waren, vereinen diese Modelle eine sehr einfache, elegante Abarbeitung mit Optimalität der Leistungsschätzungen.

Gemischtes Modell und Interpretationen der Varianzkomponenten

Das zugrunde gelegte Modell geht aus den in Tab. 3 ausgewiesenen Varianzkomponenten hervor.

Der Effekt **Ort*Jahr*Typ**, der das Leistungsniveau der Einzelversuche erfasst und berücksichtigt, wird in diesem Modell formal als fixer Effekt behandelt. Die Varianz der Einzelversuche ist erfahrungsgemäß so groß, dass die Nutzung des ‚zufälligen‘ Ansatzes keine Erhöhung der Schätzgenauigkeit für die Sorteneffekte bringt, aber die Rechenzeit erheblich erhöht. Die grundsätzliche Berücksichtigung dieses Effektes ist aber Voraussetzung für die Verrechnung unbalanzierter Daten, also z.B. die gemeinsame Verrechnung von LSV, WP, EU, die Verrechnung langjähriger Daten und die gemeinsame Auswertung mehrerer Regionen.

Tab. 3: Ausgabe der Varianzkomponenten durch PIAFStat
(Beispiel Winterweizen Anbaugebiet D-Nord 1997-2006)

Obs	CovParm	Subject	Estimate	Alpha	Lower	Upper
1	Intercept	S	2.4479	0.2	1.8866	3.3803
2	JAHR	S	1.7807	0.2	1.4813	2.2036
3	ort	S	0.7020	0.2	0.5326	0.9933
4	JAHR*r	S	0.8983	0.2	0.6797	1.2764
5	JAHR*ort	S	3.5613	0.2	3.1065	4.1458
6	typ(JAHR*ort)	S	1.2629	0.2	0.9272	1.8856
7	UN(1,1)	S	0.9773	0.2	0.5673	2.4387
8	UN(2,1)	S	0	.	.	.
9	UN(2,2)	S	0.5621	0.2	0.3020	1.7921
10	UN(3,1)	S	0	.	.	.
11	UN(3,2)	S	0	.	.	.
12	UN(3,3)	S	0	.	.	.
13	Residual		1.0000	.	.	.
	mittlerer SE ² in Versuchen		7.4235			

Tab. 4: Ausgabe der genetischen Korrelationen durch PIAFStat

1	Korrelation zw. AG1 und AG2	0.76238
2	Korrelation zw. AG1 und AG3	0.84539
3	Korrelation zw. AG2 und AG3	0.90181

ad (Obs 1) Sorten-Varianz innerhalb der Sorten-Gruppen

Zur Optimierung der Gewichtung von Regionen ist Sorte hier formal als zufällig betrachtet. Bei der Mittelwertsschätzung wird Sorte aber als fix, also als Individuum angesprochen (lsmeans).

ad (Obs 2) Sorte * Jahr Wechselwirkung im gesamten Großraum

Leistungs-Relationen der Sorten sind in den Jahren über alle Gebiete hinweg unterschiedlich. Die unterschiedliche Besetzung der Sorten in Jahren im Großraum wird optimal berücksichtigt.

ad (Obs 3) Sorte * Ort Wechselwirkung in Anbaugebieten

Leistungs-Relationen der Sorten sind an den Standorten innerhalb der Anbaugebiete unterschiedlich. Wenn diese VK rel. klein ist, bestätigt dies die vorgenommene Regionalisierung dahingehend, dass die Sorten an den Standorten innerhalb jedes Anbaugebietes im lang-jährigen Mittel vergleichsweise ähnlich reagieren. Die unterschiedliche Besetzung der Sorten an Orten (Züchter, BSA, LDS...) im Großraum wird optimal berücksichtigt.

ad (Obs 4) Sorte * Jahr Wechselwirkung in Anbaugebieten

Leistungs-Relationen und Besetzung der Sorten sind in den Jahren innerhalb der Anbaugebiete zusätzlich zu (Obs 2) unterschiedlich. Wenn diese Komponente kleiner als (Obs 2) ist, deutet dies auf relativ ähnliche klimatische Bedingungen im gesamten einbezogenen Großraum hin. Die unterschiedliche Besetzung der Sorten in Jahren in Anbaugebieten wird optimal berücksichtigt.

ad (Obs 5) Sorte * Jahr * Ort Wechselwirkung in Anbaugebieten

Die Sorten reagieren in Einzelversuchen spezifisch auf Ereignisse, die weder am Standort reproduzierbar noch für das Jahr typisch sind. Diese Komponente ist oft groß, was die Interpretation der Sorten-Reaktionen auf Umwelteinflüsse, z.B. die Verallgemeinerungsfähigkeit von (häufig nur scheinbaren) Jahres- oder Orts- Effekten beschränkt. Die unterschiedliche Besetzung der Sorten wird optimal berücksichtigt.

ad (Obs 6) Sorten Wechselwirkung gleicher Ort und gleiches Jahr

In zeitgleichen Versuchen an einem Standort können bereits real abweichende Sortenrelationen eintreten, die nicht nur im Rest-Fehler begründet sind (z.B. Verrechnungsorten in WP, LSV, EU). Bei der Interpretation im Einzelfall wäre wichtig zu wissen, wie/ob sich die Anbaubedingungen dieser Versuche unterschieden haben (Schlag, Saatzeit, Betreuer, Intensität ...).

Diese Varianzkomponente kann nur geschätzt werden, wenn mit Wichtung nach SE (Standardfehler) gearbeitet wird. Die z.T. mehrfache Besetzung einzelner Sorten in Versuchen am gleichen Ort und Jahr wird optimal berücksichtigt.

ad (Obs 7; 9; 12) Sorte * Anbaugesamt Wechselwirkung im „Großraum“

Diese VK werden je Anbaugesamt ausgewiesen. Sie sind in ihrer Relation zu **(Obs 1)** die Basis für die Gewichtung von Nachbargebieten und damit einer der entscheidendsten Aspekte der Hohenheim-Gülzower-Serienauswertung.

ad (Obs 13) Sorten Wechselwirkung im Einzelversuch

Dies ist die Varianz (SE^2) der MW je Sorte in Einzelversuchen, wenn die Option ‚Ohne Gewichtung nach SE‘ gewählt wurde. Bei Verrechnung mit Gewichtung nach SE wird hier formal ‚1‘ als Konstantfaktor für die hinterlegten MW-spezifischen SE^2 ausgewiesen. Unterschiedliche Schätzgenauigkeiten der Mittelwerte (durch unterschiedliche Wiederholungsanzahl oder s%) werden nur bei ‚Gewichtung nach SE‘ optimal berücksichtigt. Das Verfahren ‚VK‘ weist bei Gewichtung nach SE zur Information den mittleren SE^2 zusätzlich aus, um auch für (Obs 13) eine informative Größenordnung zu erhalten.

Zur Gewichtung der Anbaugesamte

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

- a) genetische Korrelation zum Zielgebiet (Tab. 4)
je größer, desto stärker gewichtet, für alle Sorten gleich
- b) Prüfumfang im Anbaugesamt
je größer, desto stärker gewichtet, für die Sorten/ Jahrgänge differenziert!

Für das jeweilige Zielgebiet haben sich die Neuen Bundesländer auf eine Mindestzielbesetzung von 5 Standorten verständigt. Eine ausgewogene Besetzung aller Anbaugesamte mit jeweils möglichst > 3 Orten auch in einbezogenen Nachbargebieten ist sehr wichtig. Der Ansatz, z.B. von einem Nachbargebiet nur den nächstliegenden Ort zu verwenden, lässt eine hinreichende Schätzung der genetischen Korrelationen nicht zu und würde zu einer z.T. sehr hohen Gewichtung dieses Ortes führen, weil an ihn die Anforderung gestellt ist, ein ganzes Anbaugesamt zu repräsentieren. Das erscheint dann im Ergebnis z.T. nicht ganz nachvollziehbar, liegt aber an extremen Vorgaben des Anwenders.

Da in jedem LSV-Jahrgang die Wichtungsanteile anders sind (je länger geprüft, desto höhere Eigengewichtung des Zielgebietes ist optimal), ist die häufig zu beobachtende Erwartung, dass für alle Sorten die Wichtung proportional ist, nicht zutreffend. Es erfolgt für jede Sorte in Abhängigkeit von ihren konkreten Prüfungsumfängen in Regionen/Jahren etc. eine optimale Wichtungsverteilung.

Das Verfahren ‚MW‘ - Bestimmung der Mittelwerte

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

Für ein vorzuziehendes Zielgebiet (Abb. 2) werden (mehrjährige) Sorten-Mittelwerte in folgender Weise bestimmt: die MW sind auf eine einheitliche Ebene /Vergleichsbasis /mittleres absolutes Leistungsniveau so projiziert, dass sie unmittelbar und optimal miteinander vergleichbar sind (Adjustierung). Eine Relativierung kann bei Bedarf außerhalb dieses Verfahrens erfolgen.

Als Ergebnis-Listen werden die Sorten-Mittelwerte sowie Versuchs-Mittelwerte, Anzahl Versuche je Sorten in Anbaugesamten und Gewichte der Anbaugesamte je Sorte in optionalen Sortenvarianten angeboten. Alle Tabellen können optional in Excel abgespeichert und dort weiterverarbeitet werden.

Ergebnislisten – Grundoutput (Tab. 5 und 6)

Die Abkürzungen in den Tabellenköpfen bedeuten:

- MW - adjustierter Mittelwert der Sorten für das 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

Tab. 5: Ausgabe der Sortenmittelwerte durch PIAFStat

adjustierte Mittelwerte der Sorten im Zielgebiet 4

für das Merkmal ERTR86DT Intensitätsstufe: 2

Auswertungszeitraum: 2001 bis 2006

Mindestanzahl Versuche je Sorte im Zielgebiet: 1

Wichtung nach Versuchspräzision: ja gewählter Transformationsparameter: Phi = 0,8

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.

Tab. 6: Ausgabe der Versuchsmittelwerte durch PIAFStat

adjustierte Versuchs- Mittelwerte

für das Merkmal ERTR86DT Intensitätsstufe: 2

Anbaugebiet Ort	Jahr	VNR	Versuchsmittel
5 Vipperow	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
Biestow	2003	109	98.6

usw.

Die adjustierten Mittelwerte und Vertrauensintervalle für den paarweisen Vergleich werden in der LFA u.a. entsprechend Abb. 3 grafisch dargestellt. Sorten, deren Schenkel sich nicht überlappen, unterscheiden sich im langjährigen, regionalen Erwartungswert signifikant voneinander. Diese Darstellungsform hat sich in der Praxis sehr gut bewährt, da hier Mittelwert und Schätz-

genauigkeit der Sortenunterschiede ohne Spezialkenntnisse intuitiv richtig und sinnvoll interpretiert werden können.

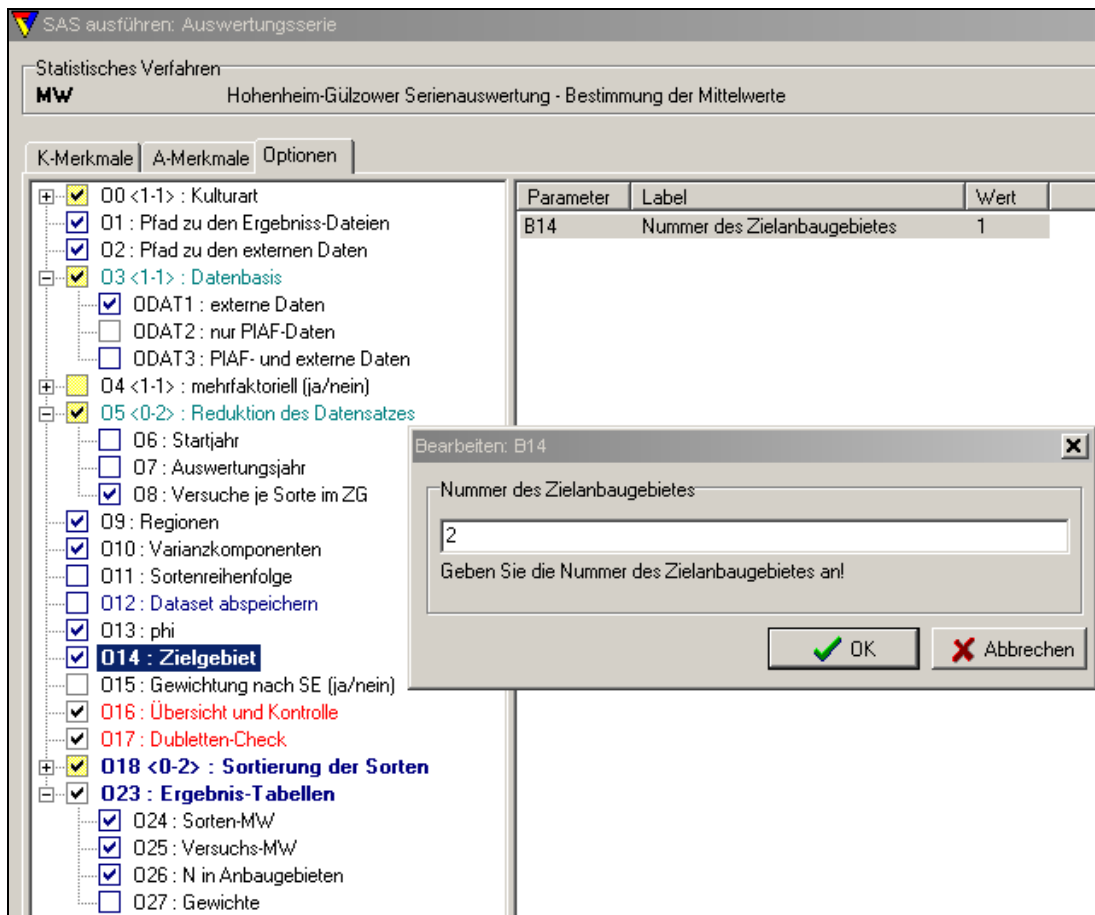


Abb. 2: Optionen des PIAFStat-Verfahrens ‚MW‘ (Hardcopy)

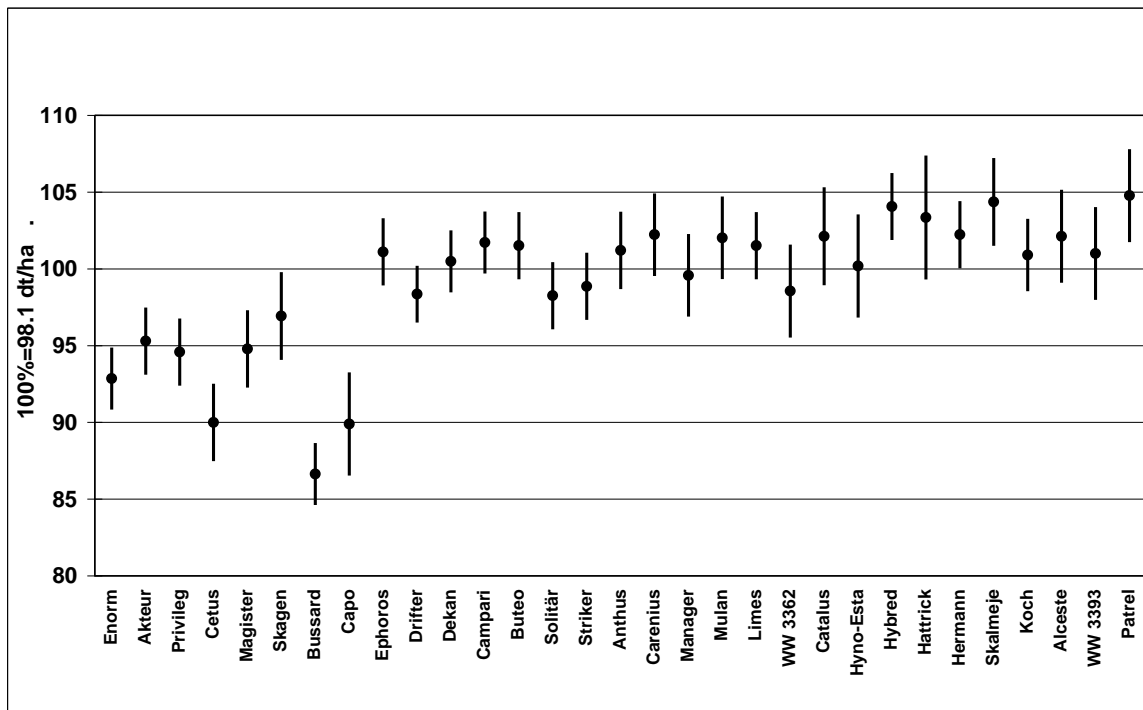


Abb. 3: Schätzwerte für Kornertrag und 90%-Vertrauensintervalle für paarweise Vergleiche nach Hohenheim-Gülzower Serienauswertung am Beispiel: Winterweizen (B und E) 2001-2006 im Anbaugebiet D-Nord

Verwendete 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
- Piepho, H.-P. (2000): Zur Durchführung multipler Vergleiche in Nicht-Standardsituationen. Zeitschrift für Agrarinformatik 8(1), 16-20
- 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
- Möhring, J., A. Büchse, H.-P. Piepho, V. Michel, J. Rath, F. Laidig. 2004. Gesundsparen ohne Nachteile. DLG-Mitteilungen 6(2004)
- Möhring, J., A. Büchse, H.-P. Piepho. 2005. Auswertung von landwirtschaftlichen Sortenversuchen mit PROC MIXED – Spagat zwischen Theorie und Praxis. SAS: Verbindung von Theorie und Praxis, Proceedings der 9. Konferenz der SAS-Anwender in Forschung und Entwicklung. 279-288
- Press W.H. 1989. Numerical recipes in Pascal. Cambridge University Press
- Roßberg, D., V. Michel, R. Graf, R. Neukampf. 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
- Weber, E. 1986. Grundriss der Biologischen Statistik. VEB Gustav Fischer Verlag. Jena
- Zenk, A., J. Möhring, V. Michel. 2005. Einbindung neuer Methoden zur Routineauswertung von landwirtschaftlichen Versuchen mit Hilfe von SAS-Macros. SAS: Verbindung von Theorie und Praxis, Proceedings der 9. Konferenz der SAS-Anwender in Forschung und Entwicklung. 407-417