

MOLL, Eckard, PIEPHO, Hans-Peter und RICHTER, Christel

Blocks fix oder zufällig?

1 Auswahl der Versuchsfelder

Die Versuchsstandorte für Feldversuche sind von unterschiedlicher Größe. Die Bodengüte ist keinesfalls einheitlich. Im Allgemeinen werden auf einem Versuchsstandort mehrere Versuche angelegt, die selten in Beziehung zueinander stehen. In der Regel werden die Blocks nicht auf der gesamten Fläche eines Versuchsstandorts verteilt. Der Versuch wird im Allgemeinen als einheitliches Ganzes aufgefasst, für das es gilt, seinen Platzbedarf unter Berücksichtigung der geforderten Rahmenbedingungen (Konstantfaktoren) wie bestimmte Vorfrucht, keine Durchführung von speziellen Behandlungen im Vorjahr u. dgl. in das Konzept der jährlich durchzuführenden Versuche einzupassen. Die Festlegung der eigentlichen Versuchsfelder erfolgt meist nicht randomisiert, sondern in Abhängigkeit vom Flächenbedarf, unter Berücksichtigung der Konstantfaktoren sowie der Durchführbarkeit der im Versuch durchzuführenden Behandlungen und der damit verbundenen Wege um den Versuch für den Einsatz der Technik. Die Entscheidung, wo der Versuch realisiert wird, trifft der Versuchsfelderleiter im Allgemeinen unter Aspekten, die von der eigentlichen Versuchsfrage völlig unabhängig sind. Die Festlegung der Konstantfaktoren eines Versuches bestimmt seinen Aussagebereich. Dazu gehört auch die Bodenwertzahl der ausgewählten Fläche auf der die Lage der konkreten Blocks als mehr oder minder zufällig betrachtet werden kann.

2 Warum Blocks im Versuchswesen?

Die Grundlage eines jeden Versuchs ist die Randomisation seiner Prüfglieder, d.h. die zufällige Zuordnung der Prüfglieder zu den Versuchseinheiten. Diese vollständige Randomisation der Prüfglieder wird im landwirtschaftlichen Versuchswesen meist durch die Bildung von Blocks eingeschränkt. Dabei ist zwischen vollständigen Blocks, die alle Prüfglieder enthalten, und unvollständigen Blocks, die jeweils eine Auswahl von Prüfgliedern umfassen, zu unterscheiden. Ziel der durch die Blockbildung eingeschränkten Randomisation ist es, die Differenzen zwischen den Prüfgliedern möglichst ohne störende Einflüsse zu schätzen. Das wird vor allem dann erreicht, wenn die störenden Einflüsse innerhalb der Blocks gering sind. Neben dieser wichtigsten Zielstellung der Blocks als Plan- und Störfaktor sind die Blocks auch Organisationsprinzip für die Arbeitsorganisation, d.h. alle Bodenbearbeitungs- und Pflegemaßnahmen, die nicht Prüffaktoren sind, alle Beobachtungen und Bonituren einschließlich der Ernte werden blockbezogen zeitgleich, mit gleicher Technik und von der gleichen Person durchgeführt. Die Blockbildung als Organisationsprinzip erfasst nicht nur die lokalen Effekte, sondern auch Personal-, zeitliche und technische Störeffekte.

Die Blocks können unter der Maßgabe, dass die Störfaktoren innerhalb der Blocks einen geringen Einfluss haben, auf der gesamten Fläche des Versuchsstandorts, für die die gleichen Konstantfaktoren vorliegen, verteilt werden. Wird der (einortige und einjährige) Feldversuch als eine flächenmäßige Einheit betrachtet, dann sind die Blocks Gruppierungen von Prüfgliedern innerhalb eines weitestgehend geschlossenen Teilfläche des Versuchsstandorts.

3 Statistische Versuchsfrage eines einortigen und einjährigen Feldversuchs und Anzahl der Wiederholungen

Bei der Formulierung der statistischen Versuchsfrage gilt es, die unterschiedlichen Zielstellungen der Feldversuche zu berücksichtigen. Damit soll nicht zwischen den

Sortenversuchen, die den größten Umfang der Feldversuche darstellen, Düngungsversuchen, Pflanzenschutzversuchen und anderen unterschieden werden. Die *statistische* Versuchsfrage wird dadurch bestimmt, ob das Schätzen oder das Testen der Erwartungswerte als statistisches Analyseverfahren als der wichtigere Aspekt angesehen werden soll. Im praktischen Feldversuchswesen zeigt es sich, dass beide Problemstellungen von Interesse sind. Sollte beispielsweise ein Auftraggeber „nur“ an den Mittelwerten interessiert sein, so ist es für die Praxisberatung auch interessant, Mittelwerte zu vergleichen. Im Allgemeinen treffen also beide statistische Versuchsfragen aufeinander.

Im praktischen Versuchswesen hat man sich aber unabhängig von der statistischen (Haupt-) Frage auf Feldversuche mit 4 Wiederholungen „geeignet“. Versuche mit mehr als 4 Wiederholungen sind eher eine Ausnahme. Weitaus stärker ist der Trend, aus ökonomischen Zwängen bei der Versuchsdurchführung, jedoch in der Regel nicht statistisch begründet, Versuche mit geringerer Zahl von Wiederholungen durchzuführen.

4 Modellwahl und Zielstellung einer statistischen Analyse

4.1 Statistisches Schließen

Die Präzisierung der Versuchsfrage umfasst die Art des statistischen Schließens und die Festlegung des Aussagebereichs des Versuchs. Nach obigem kann der statistische Schluss von der Form sein, dass Parameter zu schätzen sind, statistische Tests oder aber Auswahlverfahren durchzuführen sind. Dabei bestimmt die Modellwahl die biometrische Auswertung. Sie ist nur dann ohne Belang, wenn kein statistischer Schluss vom Versuch auf eine definierte Grundgesamtheit erfolgen soll, also nur eine Beschreibung der Ergebnisse des Versuchs oder eine gewisse Datenverdichtung vorgenommen werden soll, ohne das Schätzen von Parametern oder das Testen von Hypothesen über Parameter.

Für den Blockfaktor gibt es prinzipiell die Möglichkeit, ihn als fixen oder zufälligen Faktor im statistischen Modell zu berücksichtigen. Wird er als fix angenommen, so können Blockeffekte geschätzt und auch miteinander verglichen werden. Das ist aber im Allgemeinen nicht das Ziel eines Feldversuchs. Dem Blockfaktor kommt mit seiner Rolle, die vollständige Randomisation der Prüfglieder gruppiert einzuschränken, eine Sonderstellung unter den Faktoren des Modells zu. Es geht i.d.R. nur darum, zwecks Nutzung der Inter-Block-Information bei unvollständigen Blocks die Blockvarianz zu schätzen. Es gibt im Allgemeinen für die Blocks keine Fragestellung für einen Test.

Das Hauptaugenmerk liegt folglich im Schätzen bzw. im Testen von Prüfgliedeffekten. Da bei der varianzanalytischen Auswertung das Schätz- und das Testproblem aufeinander treffen, bedarf es genauerer Betrachtungen.

4.2 Zum Schätzen der Erwartungswerte der Prüfglieder

Zur Schätzung von Prüfglied-Mittelwerten für eine Region, das trifft auch für den Versuchsstandort zu, müssen die Versuchsstandorte bzw. Flächen am Standort zufällig aus dem Gebiet bzw. Standort (der Grundgesamtheit) entnommen sein. Nur so wird gewährleistet, dass die Schätzfunktionen für Parameter die ihnen unterstellten Eigenschaften haben. Wird nicht so verfahren, dann können die Ergebnisse nicht als repräsentativ für die Grundgesamtheit (Region oder Versuchsstandort) betrachtet werden. Da die Blocks im Allgemeinen als zusammenhängende Flächen nicht zufällig aus der Menge der möglichen Blocks des Versuchsstandorts ausgewählt werden, kann die Schätzung der Erwartungswerte und Varianzen der Prüfglieder nicht repräsentativ für die Grundgesamtheit Versuchsstandort, sondern nur für

die ausgewählte Versuchsfläche sein, obwohl die Prüfglieder innerhalb der Blocks randomisiert wurden. Das trifft so auch auf die geschätzten Konfidenzintervalle der Erwartungswerte der Prüfglieder und auf ihre geschätzten Standardfehler zu.

4.3 Betrachtung der Prüfgliedifferenzen, d.h. Testen von Hypothesen über die Erwartungswerte der Prüfglieder

Wie bereits festgestellt, soll häufig der Versuchsstandort die Grundgesamtheit sein. Die Versuchsfläche wird meist nicht randomisiert, aber von der Versuchsfrage unabhängig ausgewählt. Solange Additivität von Block- und Behandlungseffekten auch in benachbarten Flächen/Regionen gilt, können die bezüglich der Differenzen gewonnenen Ergebnisse auf den gesamten Standort übertragen werden. Wird diese Annahme a priori unterstellt, dann ist eine Zufallsstichprobe von Blocks aus der Menge aller Blocks des Standortes nicht notwendig. In dem Fall sollte den Standardfehlern der Mittelwerte prinzipiell keine Beachtung geschenkt werden (PIEPHO 2006), den Standardfehlern der Differenzen und denen anderer Kontraste immer.

Die Blockbildung dient vor allem dazu, die Differenzen zwischen den Prüfglied-Mittelwerten unverzerrt und mit möglichst hoher Präzision zu schätzen. Für die Differenzen der Mittelwerte können die Standardfehler und Konfidenzintervalle geschätzt und die Tests von Hypothesen über die Erwartungswerte der Prüfglieder durchgeführt werden. Bei vollständigen Blocks ist es egal, ob Blocks als fest oder zufällig betrachtet werden (solange die Blockvarianzschätzung positiv bleibt), da die statistischen Schlüsse für Mittelwertdifferenzen hier identisch sind. Bei unvollständigen Blocks nimmt man Blockeffekte als zufällig, wenn man die Inter-Block-Information nutzen will, sonst fest (Intra-Block-Analyse).

4.4 Blocks als zufällige Effekte im Modell eines einortigen und einjährigen Feldversuchs

Betrachten wir die Auswirkungen der Modellwahl, d.h. Blocks fix oder zufällig, und gehen von der z.B. in SAS gebräuchlichen REML-Schätzung aus, dann beeinflusst die Modellwahl sowohl die Schätzwerte der LsMeans (least squares means, adjustierte Mittelwerte) als auch die Schätzung von den dazugehörigen Standardfehlern, wenn die Versuchsanlage unbalanciert ist. Ist für die Blocks „zufällig“ anzunehmen, darf es keinesfalls bei nichtpositiven Blockvarianzen zu einer automatischen Reduktion des Modells kommen. Das würde den Randomisationsschritten, die zur Annahme führen, dass die Blocks zufällig sind, widersprechen. Eine mögliche Modellreduktion kann in der SAS-Prozedur MIXED durch das Setzen der Option nobound verhindert werden. Bei zu geringer Wiederholungsanzahl für einige Prüfglieder und/oder unvollständigen Blocks kann es hierdurch aber zu numerischen Problemen kommen, wenn die Varianz-Kovarianz-Matrix nicht-positiv-definit wird. In diesen Fällen sollten die Blockeffekte wie fixe Parameter im Modell behandelt werden.

PIEPHO (2006) und RICHTER (2006) beschäftigen sich mit den Beziehungen zwischen dem linearen Modell und dem Randomisationsmodell. Beide plädieren dafür, Blocks in die Randomisationprozedur für einen Versuch einzubeziehen und bei der Analyse von zufälligen Blocks auszugehen. Während PIEPHO (2006) die Relation bei endlicher Grundgesamtheit näher betrachtet, diskutiert RICHTER (2006) die Beziehung im Hinblick auf eine unendliche Grundgesamtheit. Ihre Schlussfolgerungen sind in die Empfehlungen aufgenommen worden.

4.5 Konsequenzen der statistischen Analyse einer Versuchsserie durch die Wahl der Blocks als fix oder zufällig

Bei der Versuchsserie sind in das lineare Modell zwei die Umwelt beschreibende Faktoren und ihre Wechselwirkungen aufzunehmen. Das sind der Faktor Orte für die Versuchsstandorte und Jahre für die Versuchsjahre. Natürlich gilt auch hier wie für die Blocks, dass, wenn der Faktor Orte zufällig sein soll, eine Zufallsstichprobe aus den Orten einer Region zu ziehen ist, gleiches für den Faktor Jahre. Da diese Problematik aber nicht Gegenstand der hier auf die Blocks gerichteten Betrachtungen ist, soll auch nicht näher darauf eingegangen werden.

Betrachten wir zunächst den einfachsten Fall, dass die zu einer Versuchsserie zusammenzufassenden Einzelversuche orthogonal in Orte, Jahre und Prüfglieder sind. An allen Versuchsstandorten werden in allen Versuchsjahren dieselben Prüfglieder angelegt. Zudem gehen wir noch vom gleichen, für jeden Einzelversuch neu randomisierten Modell der Versuchsanlage aus. Sind die Einzelversuche zudem noch balanciert oder höchstens gering unbalanciert, lässt sich eine die unterschiedlichen Fehlervarianzen der Einzelversuche berücksichtigende gewichtete Analyse der Versuchsserie vornehmen. Sehen wir uns dabei nur die Blocks an, so stehen sie in hierarchischer Beziehung zu den Umwelten. Ist einer der Umweltfaktoren zufällig, so ist der die Blocks im Modell betreffende Term Blocks(Orte*Jahre) ebenfalls zufällig. Sind die Blocks zufällig und es ist keiner der Umweltfaktoren zufällig, dann ist in einer Auswertung dieser Term natürlich zufällig.

Unterscheiden sich die Modelle der Versuchsanlagen der zu einer Versuchsserie zusammenzufassenden Einzelversuche nicht stark – eine der Versuchsanlagen ist statt einer zweifaktoriellen Spaltenanlage eine zweifaktorielle Streifenanlage oder zweifaktorielle Blockanlage –, dann bietet sich die Möglichkeit einer Analyse der Versuchsserie mit Hilfe gewichteter Prüfgliedmittelwerte an. Der Vorteil ist, dass die Wichtung für jedes Prüfglied und modellbezogen erfolgt. PIEPHO (1999) schlägt als Gewicht das reziproke Quadrat der Standardfehler der Mittelwerte vor. Dieses Vorgehen ist dann geeignet, wenn der Einzelversuch mit festen Blocks ausgewertet wird. Bei zufälligen Blocks sind die Kovarianzen zwischen Prüfgliedmittelwerten oft nicht zu vernachlässigen, so dass Gewichte anders zu berechnen sind. Ein entsprechendes Verfahren basierend auf der Inversen der Varianz-Kovarianz-Matrix der Prüfgliedmittelwerte eines Versuches schlagen SMITH et al. (2001) vor. Dieser Weg wird auch beschrieben, wenn die Versuchsserie nicht orthogonal ist. Für den großen Bereich der Sortenversuche muss i.d.R. von nicht orthogonalen Versuchsserien ausgegangen werden. Die gewichtete Analyse ist geeignet und praktikabel.

5 Empfehlung

Es wird empfohlen, grundsätzlich von zufälligen Blocks auszugehen. Zudem ist das gewählte Modell mit seiner Randomisationsstruktur die Grundlage der Analyse, d.h. dass in der SAS-Prozedur MIXED die Option `nobound` zu setzen ist.

In Abweichung von dieser Grundregel sollte in folgenden zwei Ausnahmen für das Testen der Mittelwerte von fixen Blocks ausgegangen werden. Für die Formulierung der ersten Abweichung von genanntem Grundsatz wird zugrunde gelegt, dass ab einer Wiederholungszahl von 4 bei der Schätzung eines Erwartungswertes erste Stabilitätserscheinungen eintreten. Für das praktische Versuchswesen bedeutet das, dass jedes Prüfglied mindestens viermal vorhanden sein sollte. Für eine relativ zuverlässige Schätzung der Blockvarianz braucht man allerdings mehr als 10 Blocks.

Fixe Blocks sollten in folgenden Ausnahmen gewählt werden,

Ausnahme 1:

Wenn es vorkommt, dass mehrere Prüfglied-Mittelwerte aus weniger als 4 Wiederholungen berechnet werden (bei vollständigen und unvollständigen Blocks).

Ausnahme 2:

wenn die REML-Schätzung bei zufälligen unvollständigen Blocks Null oder negativ wird, die Varianz-Kovarianz-Matrix nicht-positiv-definit wird. Bei vollständigen Blocks können und sollten negative Varianzschätzwerte zugelassen werden, weil die Varianz-Kovarianz-Matrix immer positiv-definit bleibt (NELDER 1954).

6 Literatur

NELDER, J. A. (1954): The interpretation of negative components of variance
Biometrika **41**, 544-548

PIEPHO, H.-P. (2006): Anmerkungen zur Frage „Blocks fix oder zufällig“
schriftliche Ausarbeitung, März 2006

PIEPHO, H.-P. (1999): Stability analysis using the SAS system
Agronomy Journal **91**, S. 154-160

RICHTER, Ch. (2006): Gedanken zur Problematik „Blocks im Feldversuchswesen - fix oder zufällig“
Diskussionsbeitrag zum Thema „Blocks zufällig oder fix“ der AG Versuchswesen

SMITH, A., CULLIS, B., and GILMOUR, A. (2001): The analysis of crop variety evaluation data in Australia
Australian and New Zealand Journal of Statistics **43**, 129-145