

Anmerkungen zu der Versuchsanlage 2001/2002 von Tandem 1

1. Jedes der drei Prüfglieder :
 - Kontrolle I Mais im Direktsaatverfahren, Vorfrucht Wintererbsen
 - Kontrolle II Mais im Direktsaatverfahren, Vorfrucht Wintererbsen-Roggengemenge
 - Kontrolle Umbruch Mais in Hauptfruchtstellung, Vorfrucht Wintererbsen-Roggengemenge, im April untergepflügt
 kommt **nur auf einer** Parzelle (**einer** Versuchseinheit/Randomisationseinheit) vor.

Es gibt keine Wiederholungen. Dadurch kann keine Schätzung des Versuchsfehlers (Variabilität zwischen gleich behandelten Versuchseinheiten) erfolgen; eine Bewertung der Prüfglied-Unterschiede ist nicht möglich.

2. Auf jeder Parzelle wurden an sieben Stellen Bodenproben entnommen. Diese sind (aus Gründen der Praktikabilität) nur entlang von drei Seiten der Parzellen angeordnet. Damit stellt sich die Frage der Repräsentativität der Proben für die Gesamtfläche der Parzellen.

Um repräsentativ zu sein, müsste jeder Stichprobenpunkt auf der Gesamtfläche dieselbe Wahrscheinlichkeit haben, gewählt zu werden. Durch die a-priori-Festlegung auf die Ränder wird diese Wahrscheinlichkeit für Probenpunkte im Innern der Parzelle auf Null gesetzt. Dadurch können Verzerrungen entstehen, z.B. wenn es „Randeffekte“ gibt.

Wie sicher ist es, dass die Variabilität auf einem beliebigen Transsekt gleich ist wie entlang der drei Ränder ?

3. An jedem der $3 \times 7 = 21$ Probenpunkte wurde an neun Terminen eine Probe entnommen und an dieser zwei Merkmalswerte bestimmt : N_{\min} und N_{mic} (in μg pro g Boden ? und hochgerechnet auf kg/ha Bodenfläche ?).

Die zwei Stickstoffwerte bilden eine zweidimensionale Zufallsvariable; die neun zeitlich gestaffelten Ergebnisse eine „Zeitreihe“ oder „repeated measurement“, „longitudinale Daten“. Da diese alle aus demselben Probenstandort stammen (und dazwischen keine Randomisation oder eine Zufallsauswahl stattfand), muss die Kovarianz-Struktur zwischen den aufeinanderfolgenden Messwerten bei der Auswertung berücksichtigt werden.

Stochastisch unabhängig sind die 21 zweidimensionalen „Zeitreihen“. (Dabei bleibt die Einschränkung unter Pt. 2, dass die Probenpunkte entlang der Seitenränder relativ systematisch angeordnet waren, unberücksichtigt.)

4. Die Daten weisen eine deutliche Heteroskedastizität auf (bes. deutlich ausgeprägt bei N_{\min}) : je grösser der mittlere Merkmalswert, desto grösser die Varianz (Standardabweichung).

Dieser Tatsache ist bei der statistischen Auswertung Rechnung zu tragen. Folgende Lösungswege bieten sich an :

a. Durch eine geeignete Datentransformation (z.B. Übergang zu Logarithmen) kann die Heteroskedastizität oftmals behoben werden.

b. Mit modernen Rechenverfahren können auch heteroskedastische Daten analysiert werden (weighted least squares).

c. Rangverfahren bieten eine Alternative, da die Vergabe von Rängen invariant unter monotoner Transformation ist.

5. Daten-Auswertung.

5.1 Problemlos sind Mittelwert-Vergleiche für ein Merkmal zwischen den Parzellen zu einem festgelegten Zeitpunkt (Querschnitt-Studie). Hierzu liegen 21 stochastisch unabhängige Beobachtungen vor. Unter entsprechenden Verteilungsannahmen (z. B. nach einer geeigneten Datentransformation) wäre eine Varianzanalyse durchaus möglich; bei gleichzeitiger Betrachtung beider Merkmale eine zweidimensionale MANOVA.

*Zu beachten ist dabei, dass Vergleiche zwischen Parzellen **NICHT** mit Vergleichen zwischen den Prüfgliedern gleichgesetzt werden können, da hierzu eine Schätzung der Versuchsstreuung fehlt (vgl. Pt.1) !*

Vorsicht ist geboten, wenn Querschnitt-Vergleiche für mehrere Zeitpunkte durchgeführt werden. Die statistischen Tests sind nicht voneinander unabhängig, und Korrekturen bei den Irrtumswahrscheinlichkeiten für den Fehler erster Art müssen angebracht werden.

5.2 Für Mittelwert-Vergleiche zwischen verschiedenen Zeitpunkten innerhalb einer Parzelle müssen Methoden für „repeated measurements“ angewandt werden: entweder multivariate Ansätze oder Methoden für „Profile“ (Analyse der Unterschiede zwischen den Zeitpunkten).

5.3 Vergleiche der Zeitverläufe zwischen verschiedenen Parzellen müssen ebenso mit Methoden für „repeated measurements“ („longitudinale Daten“) vorgenommen werden.

Verteilungs-unabhängige (Nicht-Parametrische) Analyse : vgl. z. B.

*Edgar BRUNNER und Frank LANGER : **Nichtparametrische Analyse longitudinaler Daten**. R. Oldenbourg Verlag München Wien 1999.*

6. Ausblick

Unabdingbare Voraussetzung für einen Vergleich zwischen den drei Prüfgliedern KI, KII und KU wäre eine Versuchsanlage mit mehreren Wiederholungen der Behandlungen, d.h. es müssten (im balancierten Fall) **3n Parzellen** zur Verfügung gestellt werden, denen durch einen geeigneten Randomisationsplan (z.B. CRD oder RCBD) jeweils eines der drei Prüfglieder zugeteilt wird.

Der Stichprobenplan für die Gewinnung der Bodenproben könnte beibehalten werden. Ggf. könnte man die Zahl der Proben pro Parzelle reduzieren, so dass die Gesamtzahl der Proben konstant bleibt. Entscheidend dabei ist, dass die Variabilität zwischen den gleich behandelten Parzellen eine Schätzung des Versuchsfehlers („experimental error“) für Vergleiche zwischen den unterschiedlich behandelten Parzellen ermöglicht.

Oscar KEMPTHORNE : *The Design and Analysis of Experiments*.

John Wiley and Sons, Inc., New York-London-Sydney, 1952, 5th printing 1966.

p. 163 :

„Throughout this book we shall use the term „*experimental unit*“ to *denote the unit that is allocated a treatment independently of the other units*. The experimental unit can contain *several observational units*; [...], *the variation between the observational units within an experimental unit is usually of little value in assessing the errors of estimates of treatment effects*.“

p. 211 :

“[...] the *experimental unit* is defined as the *unit for which a treatment is chosen at random*“.