

# Optimierung des Feldversuchswesens aus biometrischer Sicht

Versuchstechniker-Tagung der SAATEN-UNION  
am 03. / 04. November 2003  
Hohenlieth



**Andreas Büchse**  
Universität Hohenheim, Fachgebiet Bioinformatik  
[buechse@uni-hohenheim.de](mailto:buechse@uni-hohenheim.de)

# Curriculum vitae

- 1967 in Hannover geboren, verheiratet, eine Tochter
- Studium Landwirtschaft in Göttingen, Promotion im Fach Pflanzenzüchtung
- Zwei Jahre Mitarbeiter am Institut für Zuckerrübenforschung in Göttingen
- Ein Jahr selbständig (Berater für Biostatistik)
- Seit Juli 2002 Assistent im Fachgebiet Bioinformatik (Prof. Piepho), Univ. Hohenheim
- Forschungsschwerpunkt: Optimierung von Versuchsanlagen und Prüfsystemen für die Pflanzenzüchtung

# Inhalt

- 1) Biostatistische Grundlagen
- 2) Versuchsanlagen für Sortenprüfungen (einfaktoriell)
- 3) Mehrfaktorielle Anlagen
- 4) Versuchsserien
- 5) Planung vor Ort und im Detail
- 6) Spezielle Probleme

# Inhalt

- 1) Biostatistische Grundlagen (30 min)
- 2) Versuchsanlagen für Sortenprüfungen (einfaktoriell) (10 min)
- 3) Mehrfaktorielle Anlagen (5 min)
- 4) Versuchsserien (10 min)
- 5) Planung vor Ort und im Detail (15 min)
- 6) Spezielle Probleme (15 min)

# **Biostatistische Grundlagen**

- Varianz, Grenzdifferenz & Co. ...
- Warum Statistik wichtig ist...

# Denk ich an Statistik...

**Nicht den Kopf in den Sand stecken**



# Definitionen

$n$  = Anzahl Werte Stichprobe

$N$  = Anzahl Werte Grundgesamtheit

$s$  = Standardabweichung einer Stichprobe

$s^2$  = Varianz einer Stichprobe

$\sigma^2$  = Varianz der Grundgesamtheit

$x, y \dots$  Einzelwerte

$\mu$  = Mittelwert der Grundgesamtheit

# Arithmetisches Mittel

- Summe aller Werte dividiert durch Anzahl der Werte

$$\textit{arithmetischesMittel} = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$



# Arithmetisches Mittel

Beispiel: Wir haben auf einem Acker zehn Zuckerrüben ausgegraben und wiegen sie.

Wir erhalten die folgenden Zahlen (in Gramm, nach Größe geordnet)

400 400 500 600 700 900 900 1100 1100  
1300

Summe = 7900 g

Mittelwert = 790 g

# Varianz

- beschreibt die Verteilung der Werte um den Mittelwert
  - Anstatt absoluter Abweichungen der Einzelwerte vom Mittelwert, werden Abweichungen quadriert
  - Dadurch erhalten große Abweichungen höheres Gewicht
  - quadrierte Differenzen werden aufsummiert
- = Summe der Abweichungsquadrate (SQ)

# Varianz

- SQ werden abschließend durch Anzahl Werte dividiert (Mean Square = MQ).
- > Varianz ist mittleres Abweichungsquadrat
- Bei Varianz einer Stichprobe Division durch  $n-1$  (Freiheitsgrade)
  - Bei Varianz Grundgesamtheit Division durch  $N$

# Formel für Varianz (Stichprobe)

und aufsummieren

Differenz zum Mittelwert

Quadrieren

$$s^2 = MQ = \frac{SQ}{FG} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 / n}{n-1}$$

Freiheitsgrade

The diagram illustrates the derivation of the variance formula for a sample. It starts with the definition  $s^2 = MQ = \frac{SQ}{FG}$ . The numerator  $SQ$  is expanded as  $\sum_{i=1}^n (x_i - \bar{x})^2$ , and the denominator  $FG$  is  $n-1$ . This is then shown to be equivalent to  $\frac{\sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 / n}{n-1}$ . Red arrows point from the text labels to the corresponding parts of the formula: 'und aufsummieren' points to the summation symbol in the numerator; 'Differenz zum Mittelwert' points to the  $(x_i - \bar{x})$  term; 'Quadrieren' points to the square operation; and 'Freiheitsgrade' points to the  $n-1$  in the denominator.

# Varianz im Rübenbeispiel

Rübe	Gewicht	Durchschnitt	Differenz	Quadrierte Differenz
1	400	790	-390	152100
2	400	790	-390	152100
3	500	790	-290	84100
4	600	790	-190	36100
5	700	790	-90	8100
6	900	790	110	12100
7	900	790	110	12100
8	1100	790	310	96100
9	1100	790	310	96100
10	1300	790	510	260100
Summe	7900	7900	0	909000

Summe der  
Abweichungsquadrate



# Berechnen der Varianz

Summe der Abweichungsquadrate (SQ)  
beträgt 909.000

Anzahl Freiheitsgrade (FG) ist

$$n-1 = 10 - 1 = 9$$

$$\text{Varianz: } SQ / FG = 909.000 / 9 = 101.000$$

# Standardabweichung

- Varianz nur schwer zu interpretieren
- Quadrieren rückgängig machen

$$s = \sqrt{s^2}$$

$$s = \sqrt{101.000} \approx 318$$

# Variationskoeffizient

- Standardabweichung relativiert auf Mittelwert
- Angeben in Prozent

$$318 / 790 = 0,4025 \approx 40\%$$



# **Schließende Statistik**

**Wo liegt der wahre Mittelwert der Grundgesamtheit?**

**Wie schwer ist eine Rübe im Mittel?**

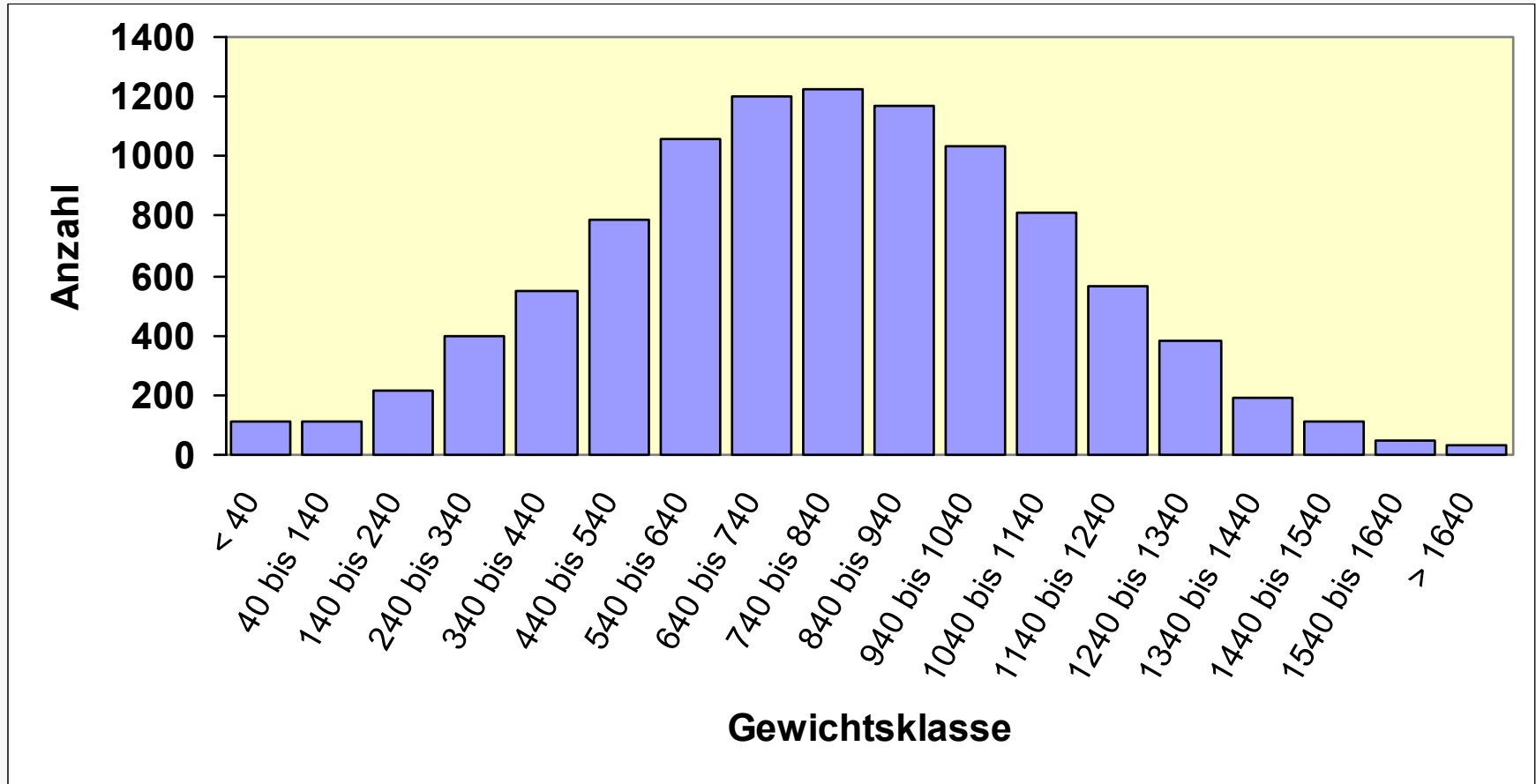
# Stichprobe und Grundgesamtheit

- Annahme: Die Stichprobe von 10 Rüben entstammt einer (unendlich großen) Grundgesamtheit von Rüben mit ...
- Mittelwert 790 g
- Standardabweichung 318 g
- Verteilungsannahme: Rübengewichte sind „normalverteilt“

# Zufallszahlen

- Mit Hilfe von entsprechender Software (z.B. Excel) kann man Zufallszahlen generieren
- z.B. 10.000 normalverteilte Zahlen mit Mittelwert 790 und Standardabweichung 318

# Verteilung von 10.000 Einzelrübengewichten



Mittelwert 790 g

Standardabweichung 318 g

# Stichprobenmittel streuen um den wahren Mittelwert

- Aus den 10.000 Rübengewichten ziehen wir zehn neue Stichproben von je 10 Rüben

Stichprobe 1	921
Stichprobe 2	801
Stichprobe 3	631
Stichprobe 4	703
Stichprobe 5	684
Stichprobe 6	622
Stichprobe 7	699
Stichprobe 8	827
Stichprobe 9	704
Stichprobe 10	853

# Standardfehler des Mittelwerts

- Auch Stichprobenmittelwerte haben eine Varianz bzw. eine Standardabweichung
  - Diese wird als Standardfehler des Mittelwerts  $S_{\bar{x}}$  bezeichnet
  - Mit dem Standardfehler kann berechnet werden in welchem Bereich der „wahre“ aber unbekannte Mittelwert der Grundgesamtheit erwartet werden kann
- > Vertrauensbereich

# Standardfehler des Mittelwerts

- Der Standardfehler kann bereits aus einer einzigen Stichprobe abgeschätzt werden

$$s_{\bar{x}} = \sqrt{\frac{s^2}{n}}$$

Je größer die Varianz der Einzelwerte desto größer wird der Standardfehler!

Je größer die Stichprobe desto kleiner wird der Standardfehler!

$s^2$  = Stichprobenvarianz

$n$  = Anzahl Einzelwerte Stichprobe

# Beispiel

Standardfehler des Mittelwerts in den  
Beispieldaten:

$$s_{\bar{x}} = \sqrt{\frac{s^2}{n}}$$

$$s_{\bar{x}} = \sqrt{\frac{101.000}{10}} = \frac{318}{\sqrt{10}} \approx 100$$



# Vertrauensbereiche

- aus Standardfehler lässt sich Intervall abschätzen in dem wahrer Mittelwert liegt (z.B. 95%-Vertrauensbereich)
- Basis: Stichprobenmittelwerte sind ebenfalls normalverteilt
- Wahrscheinlichkeit, dass wahrer Mittelwert mehr als 1,96 Standardfehler von Stichprobenmittel entfernt ist, beträgt maximal 5%.

# Vertrauensbereiche schätzen

Stichprobenmittel  $\pm 1,96$  Standardfehler  
= 95%-Vertrauensbereich für wahren  
Mittelwert

In Beispiel ist Mittelwert 790g und  
Standardfehler ca. 100g

Vertrauensbereich für den Mittelwert  
erstreckt sich von  
 $790\text{g} - (1,96 * 100\text{g})$  bis  $790\text{g} + (1,96 * 100\text{g})$   
= 594g bis 986g

# Sortenprüfungen

- Wenn Prüfglieder zu vergleichen sind, muss Mittelwert für beide präzise geschätzt werden
- Differenz der Mittelwerte schätzen
- Standardfehler der Differenz  $s_d = \sqrt{\frac{2s^2}{n}}$
- Vertrauensbereich für Differenz berechnen
- Wenn Vertrauensbereich nicht den Wert Null einschließt ist Mittelwertdifferenz signifikant
- > Konzept der Grenzdifferenz (größte gerade noch zufällig mögliche Differenz)

# Grenzdifferenz

Grenzdifferenz nach t-Test bzw. LSD

= Standardfehler der Differenz \* t-Tabellenwert  
(abhängig von  $FG_{\text{Fehler}}$  und Irrtumswahrscheinlichkeit)

$$LSD = t_{Tab} \cdot \sqrt{\frac{2s^2}{n}}$$

Grenzdifferenz ist halbe Breite des  
Vertrauensbereiches für Mittelwertdifferenz

Alle Mittelwertdifferenzen die größer als die  
Grenzdifferenz sind, gelten als signifikant (nicht  
zufällig)

# Varianzanalyse

- Sortenversuche werden meist mit Varianzanalyse (ANOVA) ausgewertet
- Gesamtvarianz wird in Streuungsursachen zerlegt
- Schätzung der Fehlervarianz wird möglich, wenn Wiederholungen angelegt sind
- Diese Fehlervarianz geht in Formeln für Standardfehler und Grenzdifferenzen ein

# Präzision

- Je kleiner die Varianz (bzw. die Standardabweichung) der Einzelwerte
  - Desto geringer ist der Standardfehler des Mittelwertes
  - Desto schmaler werden die Vertrauensbereiche
  - Desto präziser kann ein Mittelwert bzw. Mittelwertdifferenz berechnet werden
- > Je höher Varianz der Einzelwerte, desto größere Stichprobe braucht man, um Mittelwert präzise zu schätzen!

# Relevanz für das Versuchswesen

- Ziel muss Minimierung der Varianz der Einzelwerte (Fehlervarianz) sein
- Sonst brauch ich mehr Wiederholungen
- Das kostet Geld ... ☹️
- Oder ich muss geringe Präzision in Kauf nehmen
- Wahrscheinlichkeit für Fehlentscheidungen steigt an ☹️

**Hierzu eine kleine Simulation ...**  
**... ein hypothetisches Beispiel**





# 25 Sorten: Leistung zwischen 88 und 112

Sorte	Genotypischer Wert
1	112,0
2	111,0
3	110,0
4	109,0
5	108,0
6	107,0
7	106,0
8	105,0
9	104,0
10	103,0
11	102,0
12	101,0

13	100,0
14	99,0
15	98,0
16	97,0
17	96,0
18	95,0
19	94,0
20	93,0
21	92,0
22	91,0
23	90,0
24	89,0
25	88,0

die besten 5 Sorten sollen selektiert werden

# 25 Sorten werden zufällig auf ein Versuchsfeld mit 25 Parzellen verteilt

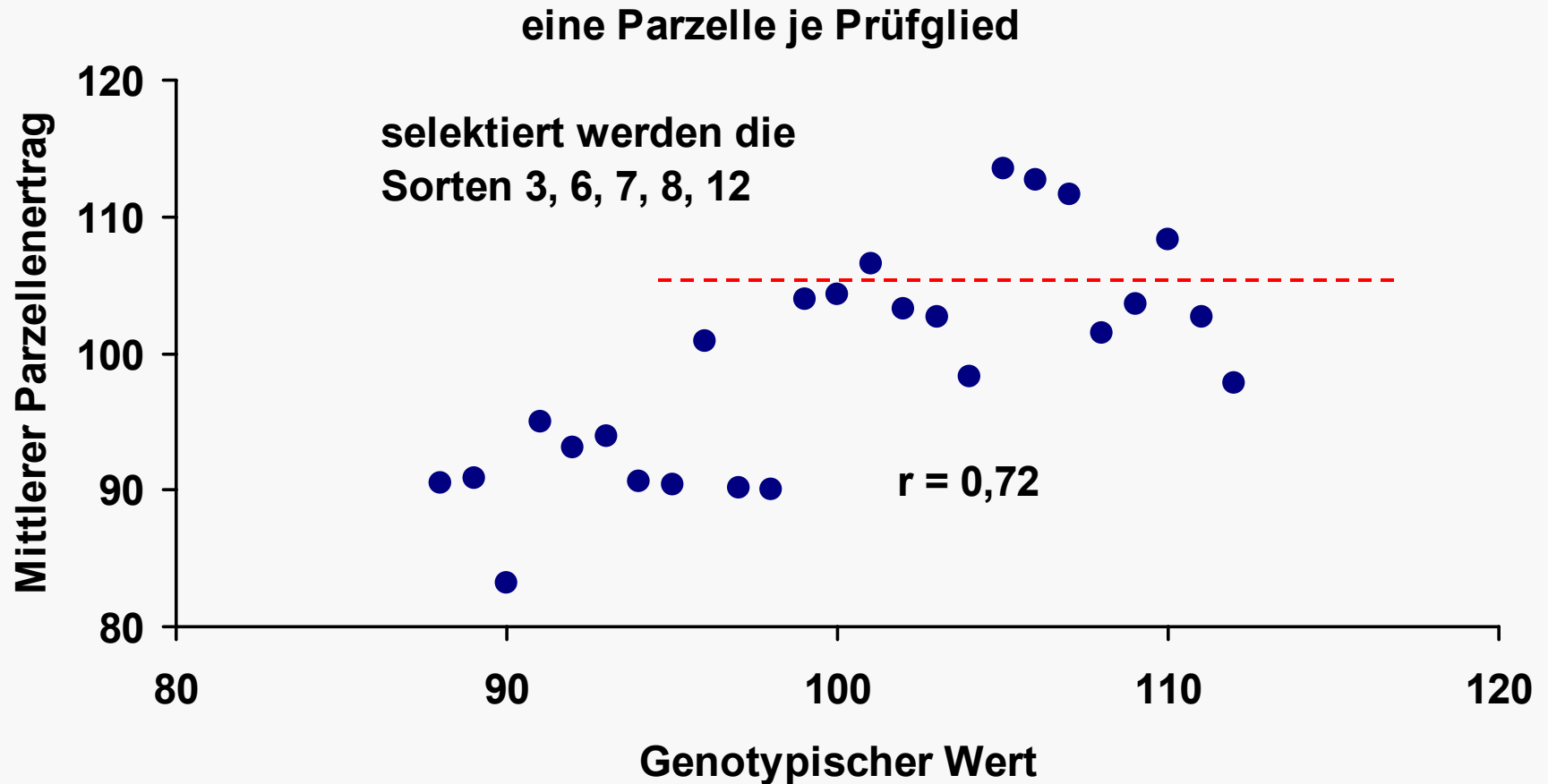
- Jede Parzelle bringt einen Fehler ein

**Karte der Parzelleneffekte**

-14,1	4,6	1,2	-6,8	1,2
-8,3	6,7	5,5	4,9	4,0
-1,7	8,5	4,3	-4,7	-6,8
-5,3	-5,7	5,0	-3,4	1,8
-6,6	-0,3	-7,9	0,9	2,5

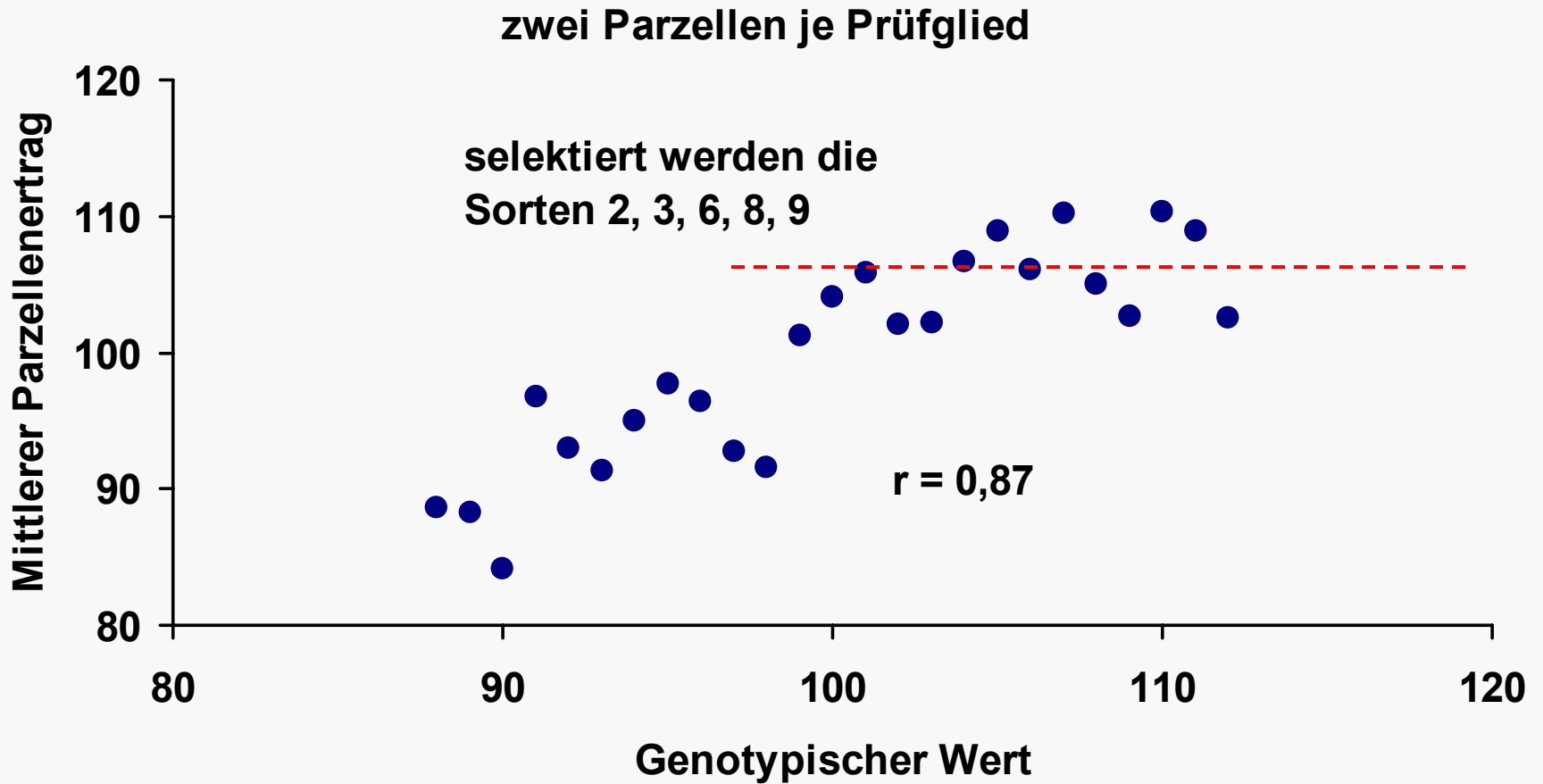
- Standardabweichung der Parzelleneffekte (Versuchsfehler) sei ca. 5,0
- Ertrag ergibt sich als Summe aus Sorten- und Parzelleneffekt

# Selektion auf Basis einer Parzelle ist nicht sonderlich erfolgreich ...

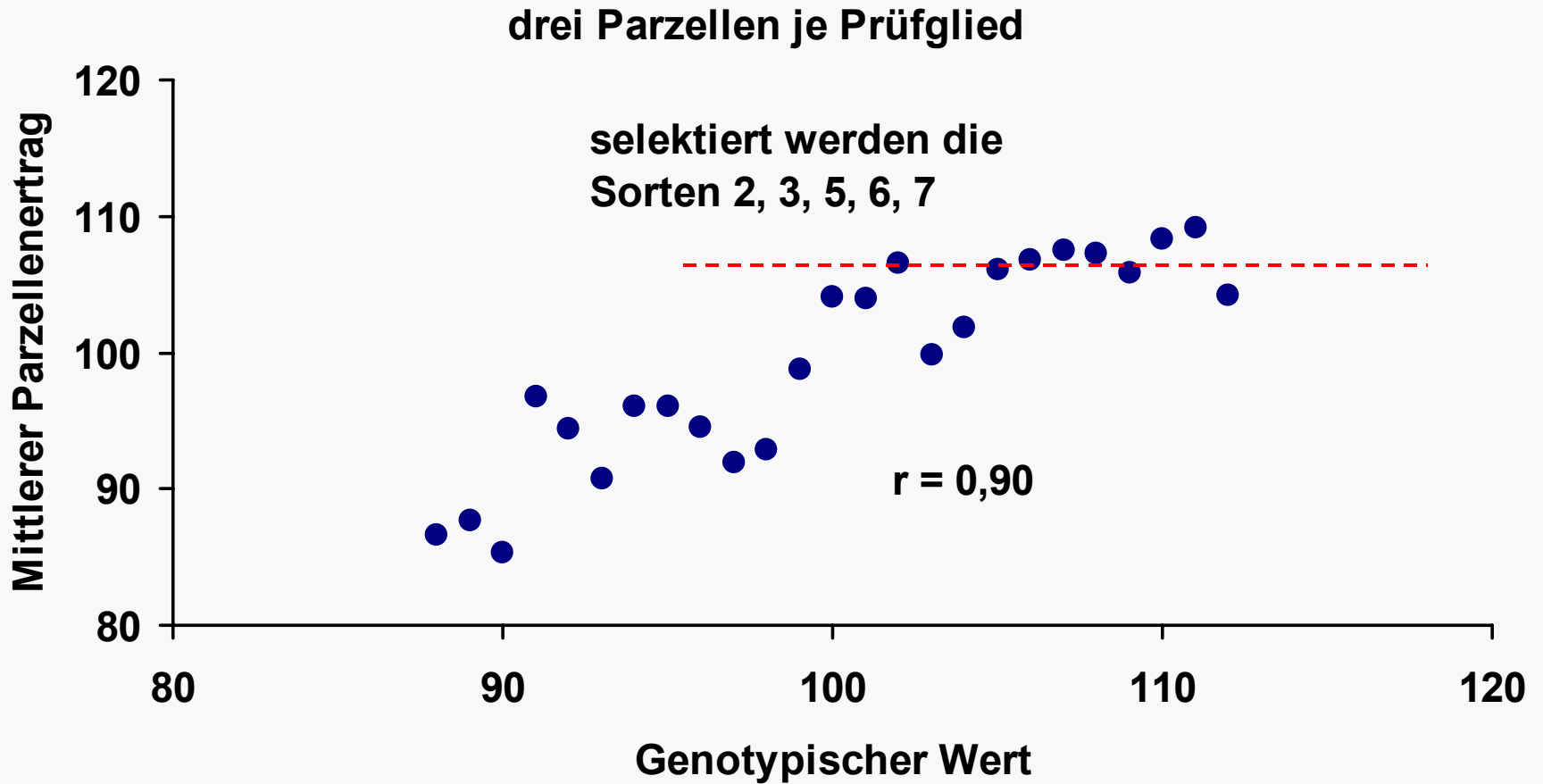




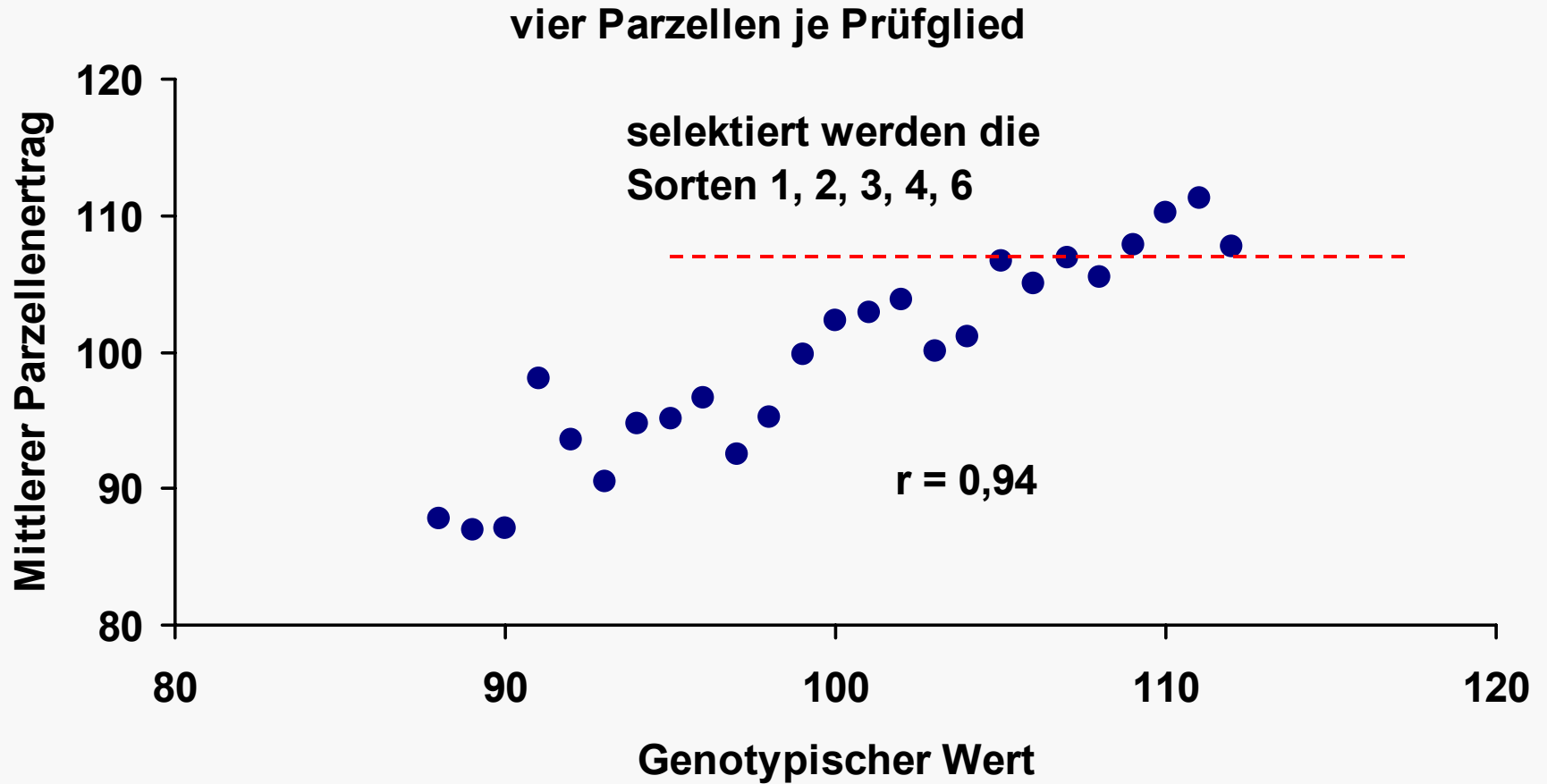
# Zwei Wiederholungen ...



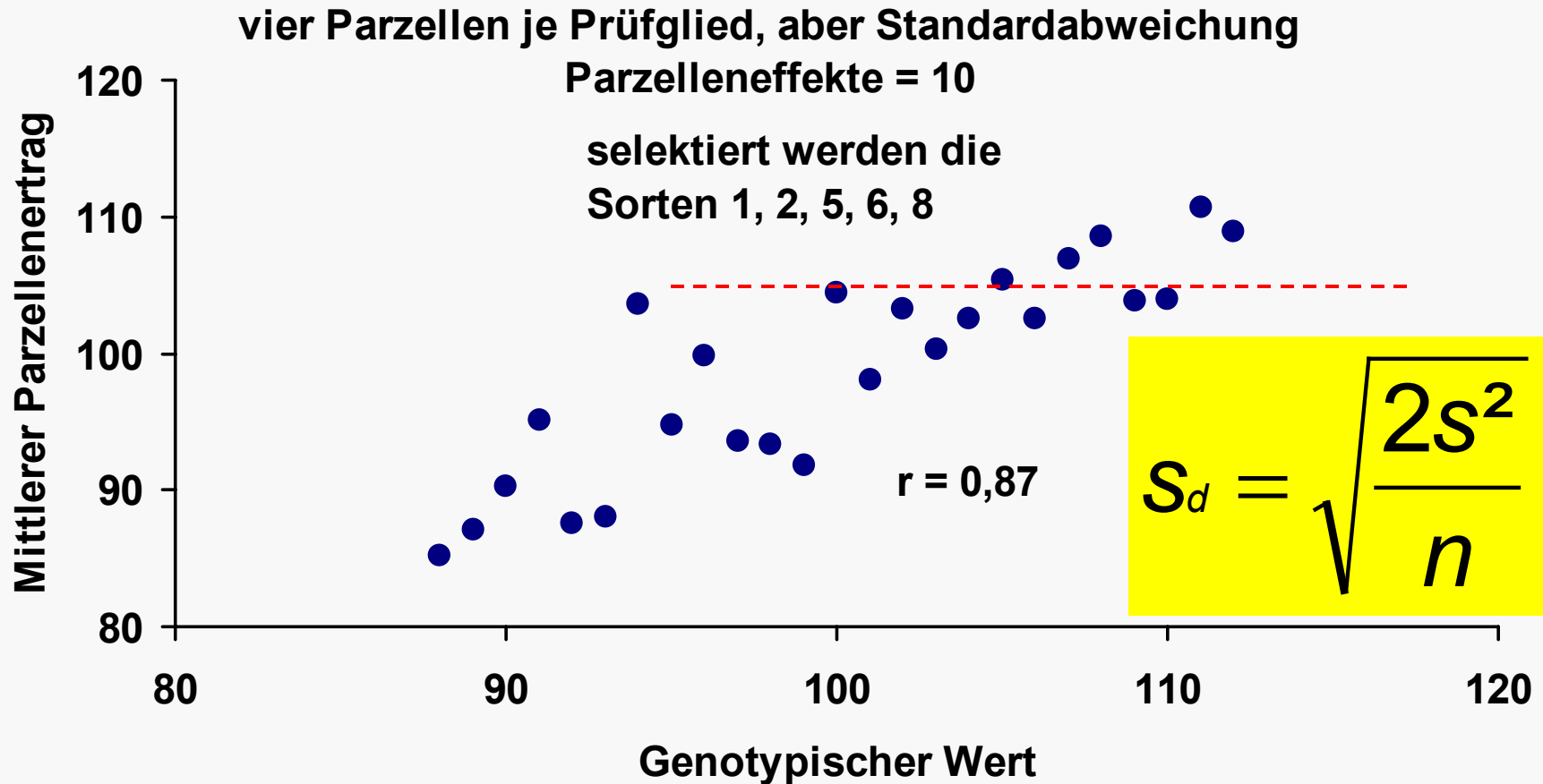
# Drei Wiederholungen ...



# Vier Wiederholungen ...



# Vier Wiederholungen, doppelter Versuchsfehler





# Optimale Anzahl Wiederholungen steigt im Verlauf der Selektion an

- Je geringer Sortenunterschiede sind
- Desto wichtiger wird hohe Trennschärfe bei Selektionsentscheidung
- Desto präziser müssen Mittelwerte und Differenzen geschätzt werden
- Desto mehr Wiederholungen sind nötig

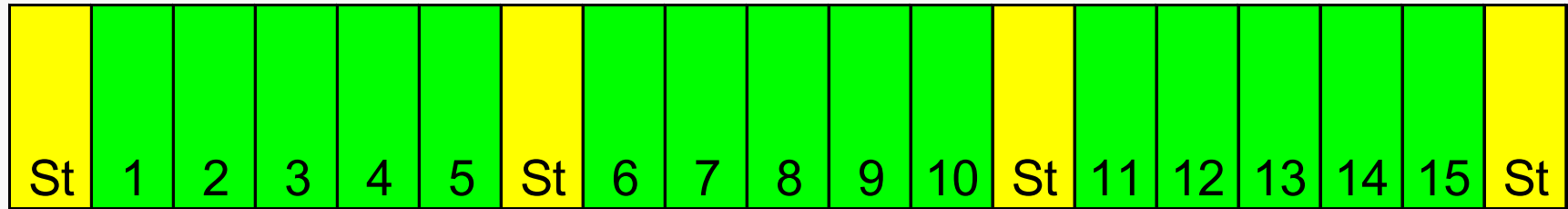


## 2. Versuchsanlagen für Sortenprüfungen



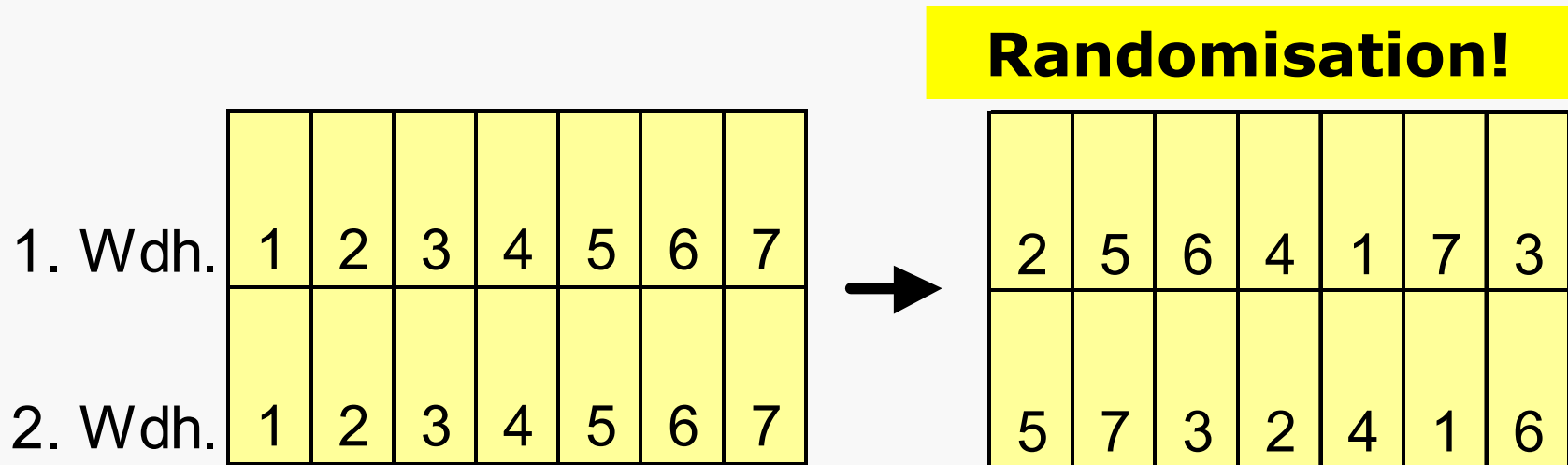


# Standardanlage



- Geringer Saatgutbedarf
- Prüfgliedeffekte werden über Standards korrigiert
- Nur eine Wiederholung > geringe Präzision
- Versuchsfehler kann nur über Standards berechnet werden

# Blockanlage



- Benachbarte Parzellen sind sich ähnlich, deshalb Versuchsfläche in Blöcke aufteilen
- Bodenunterschiede zwischen Blöcken gehen nicht in Versuchsfehler ein!
- Bodenunterschiede innerhalb Blöcken gehen in Versuchsfehler > Blöcke nicht zu groß

# Randomisation

- Randomisation ist wichtig, um zufällige Effekte tatsächlich zufällig auf die Prüfglieder zu verteilen
- Systematische Fehler ausschließen
- Auch erste Wiederholung sollte randomisiert werden!
- Auch in Anlagen ohne Wiederholungen (Standardanlage) sollte randomisiert werden! (lediglich Lage der Standards fixiert)

# Gitteranlage

- Ziel kleinere Blöcke > unvollständige Blöcke
- Nicht mehr alle Prüfglieder in jedem Block
- Prüfgliedmittelwerte werden um Blockeffekte (Bodeneffekte) adjustiert
- Beliebt: Zwei- und Dreisatzgitter (=Gitterquadrate)
- Unvollständige Blöcke in vertikaler und horizontaler Richtung = Row-Column-Design

# Entwickeln von Gitterplänen

## Zweisatzgitter: 9 Varianten

7	8	9
4	5	6
1	2	3

**Wh. I**

3	6	9
2	5	8
1	4	7

**Wh. II**

## Dreisatzgitter: 9 Varianten

**Wiederholung  
I und II wie  
Zweisatzgitter**

7	2	6
4	8	3
1	5	9

**Wh. III**

- Nicht alle Vergleiche mit gleicher Präzision möglich!
- Prüfglieder 1, 6 und 8 niemals im gleichen Block, d.h. nur indirekte Vergleiche möglich

# 7 x 7 Gitter (49 Prüfglieder)

1	24	20	5	44	31	41	14	39	25	9	47	7	17	34	30	26	8	49	40	19	6	}	2. Wdh
2	48	32	37	12	21	1	27	43	18	35	4	11	42	22	23	46	10	3	29	16	38		
3	43	44	45	46	47	48	49	R	R	R	R	R	R	R	36	13	45	28	33	2	15	}	1. Wdh
4	42	41	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22		
5	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		

- Jeweils sieben Parzellen bilden einen Block
- Jeweils sieben Blöcke bilden eine Wiederholung



# 7 x 7 Gitter (49 Prüfglieder)

Lage der  
Blöcke ist  
relativ  
flexibel

1	24	8	18	44	42	7	31	11	23	35	37	20	3	47	2. Wdh
2	39	17	27	29	43	9	2	25	32	13	21	48	38	5	
3	34	49	22	41	12	15	6	19	40	30	10	1	46	26	
4	R	R	R	R	R	R	R	45	14	16	36	4	33	28	1. Wdh
5	R	R	R	R	R	R	R	49	48	47	46	45	44	43	
6	29	30	31	32	33	34	35	36	37	38	39	40	41	42	
7	28	27	26	25	24	23	22	21	20	19	18	17	16	15	
8	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	

# Mögliche Prüfgliedzahlen bei Gittern

Prüfglieder	Parz. je Block	Wdh.	Art
12	3	2,3	Rechteck 3x4
16	4	2,3,4	Quadrat 4x4
20	4	2,3	Rechteck 4x5
25	5	2,3,4	Quadrat 5x5
27	3	3	Kubisch 3x3x3
30	5	2,3	Rechteck 5x6
36	6	2,3	Quadrat 6x6
42	6	2,3	Rechteck 6x7
49	7	2,3,4	Quadrat 7x7
56	7	2,3	Rechteck 7x8
64	4	3	Kubisch 4x4x4
64	8	2,3,4	Quadrat 8x8
72	8	2,3	Rechteck 8x9
81	9	2,3,4	Quadrat 9x9
90	9	2,3	Rechteck 9x10
100	10	2,3	Quadrat 10x10

# Alpha-Gitter (PATTERSON & WILLIAMS 1976)

Größtmögliche Flexibilität hinsichtlich  
Prüfgliedzahl, Blockgröße, Anzahl Wdh.

## Voraussetzungen für Alpha-Gitter

Blockgröße  $k$  muss kleiner oder gleich der Wurzel  
aus der Versuchsgliedzahl  $v$  sein

Zahl der Wiederholungen muss kleiner oder gleich  
Quotient  $v/k$  sein

Versuchsgliedzahl muss bei der Konstruktion des  
Planes ein ganzzahliges Vielfaches von  $k$  sein

# Mögliche Prüfgliedzahlen bei Alpha-Gittern

v	k
12	3,4
14	7
15	3,5
16	4,8
18	3,6,9
20	4,5,10
21	3,7
22	11
24	3,4,6,8,12
25	5
26	13
27	3,9
28	4,7,14
30	3,5,6,10
32	4,8
33	3,11
35	5,7
36	3,4,6,9,12
39	3,13
40	4,5,8,10

v	k
42	3,6,7
44	4,11
45	3,5,9
48	3,4,6,8,12
49	7
50	5,1
51	3
52	4,13
54	3,6,9
55	5,11
56	4,7,8
57	3
60	3,4,5,6,10,12
63	3,7,9
64	4,8
65	5,13
66	3,6,11
68	4
69	3
70	5,7,10

v	k
72	3,4,6,8,12
75	3,5
76	4
77	7,11
78	3,6,13
80	4,5,8,10
81	3,9
84	3,4,6,7,12
85	5
87	3
88	4,8,11
90	3,5,6,9,10
91	7,13
92	4
93	3
95	5
96	3,4,6,8,12
98	7
99	3,9,11
100	4,5,10

# Alpha-Gitter

Rechteck- und Quadratgitter sind lediglich Spezialfälle von Alpha-Gittern

Erstellung der Pläne mit spezieller Software (z.B. CycDesign, Alpha+) oder über sog. „Erzeugende Felder“

Auswertung mit konventioneller Software wie für andere Gitter gewohnt

# Versuche mit sehr vielen Linien

90	89	88	87	86	85	84	S10	83	82
73	74	S9	75	76	77	78	79	80	81
72	71	70	69	68	67	66	S8	65	64
55	56	S7	57	58	59	60	61	62	63
54	53	52	51	50	49	48	S6	47	46
37	38	S5	39	40	41	42	43	44	45
36	35	34	33	32	31	30	S4	29	28
19	20	S3	21	22	23	24	25	26	27
18	17	16	15	14	13	12	S2	11	10
1	2	S1	3	4	5	6	7	8	9

180	179	178	177	176	175	174	S4	173	172
163	164	S10	165	166	167	168	169	170	171
162	161	160	159	158	157	156	S9	155	154
145	146	S2	147	148	149	150	151	152	153
144	143	142	141	140	139	138	S1	137	136
127	128	S3	129	130	131	132	133	134	135
126	125	124	123	122	121	120	S7	119	118
109	110	S6	111	112	113	114	115	116	117
108	107	106	105	104	103	102	S8	101	100
91	92	S5	93	94	95	96	97	98	99

Augmented Design: wiederholte Standards,  
nicht wiederholte Linien

# Nearest-Neighbor-Analyse

## Idee:

- Für Linien jeweils nur eine Parzelle  $\Rightarrow$  beobachteter Wert stark umweltabhängig
- Problem v.a. bei hoher Streuung zwischen Parzellen
- Kann Fruchtbarkeit einer Parzelle aus Erträgen der Nachbarparzellen abschätzen  $\Rightarrow$  Korrektur des beobachteten Wertes einer Parzelle

# Nearest-Neighbor-Analyse

## Umsetzung:

- Viele alte Ansätze mit heuristischen Rechenformeln (Papadakis-Methode, etc.)
- Seit ca. 20 Jahren Beziehung zu gemischten Modellen und Geostatistik ausgenutzt



# Nearest-Neighbor Adjusted BLUP

## Idee:

- Bei Korrektur einer Parzelle sollen direkt benachbarte Parzellen höheren Einfluss haben als entferntere
- Modelliere Umweltstreuung mit Geostatistik
- Baue das ganze in gemischtes Modell ein
- Korrelation zweier Parzellen um so enger je näher beieinander
- Korrelation fungiert als Abstandsmaß; Höhe der Korrelation bestimmt den Einfluss einer Parzelle auf die Korrektur

## **3. Mehrfaktorielle Versuchsanlagen**

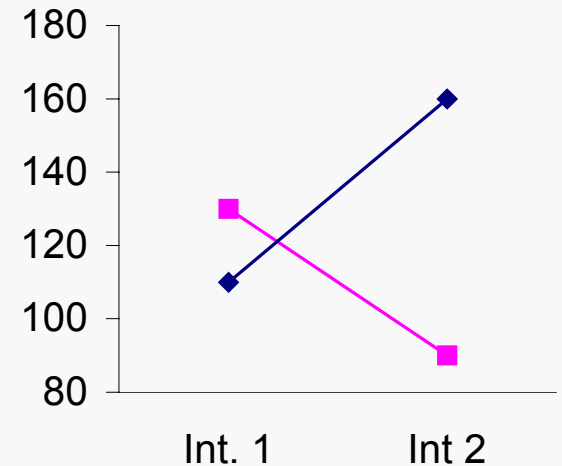
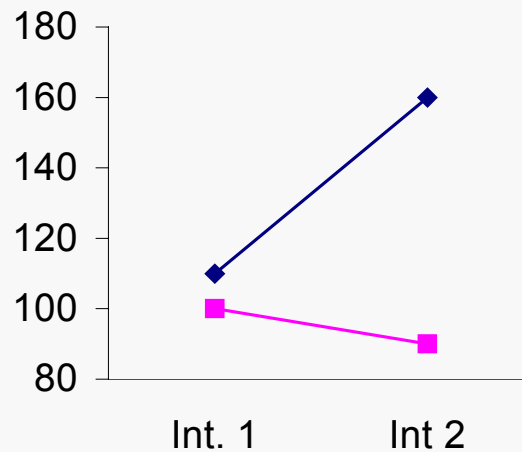
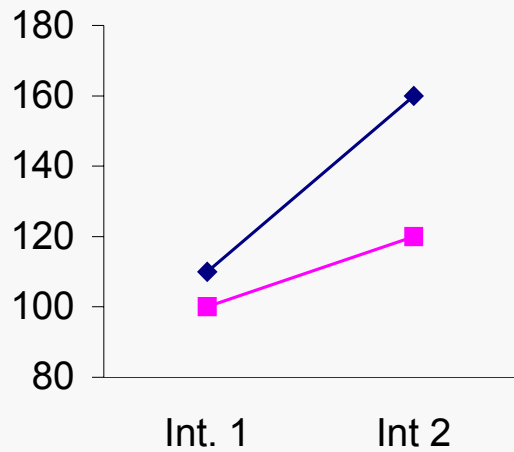
# Mehrfaktorielle Anlage

## Vor- und Nachteile:

- Interaktion Sorte x Produktionsfaktor prüfbar > Aussage wird „breiter“
- Low- / High- Input-Sorten werden sichtbar
- Nur geringere Zahl Wiederholungen möglich
- Wenn Interaktion nicht signifikant, breite Aussage und volle Anzahl Wdh. (innere Wdh.)

# Interaktionen

Ordinal, hybrid, disordinal



# Pläne für mehrfaktorielle Versuche

- Prinzipiell kann jeder mehrfaktorielle Versuch auch als Block- oder Gitteranlage angelegt werden!
- Kombination der Faktorstufen ist dann Versuchsglied
- Normalfall jedoch Spaltanlage

# Spaltanlage

1. Wdh.	3	7	1	8	2	10	5	4	9	6	ohne Fungizid
	6	8	5	2	10	7	1	9	4	3	mit Fungizid
2. Wdh.											mit Fungizid
											ohne Fungizid
3. Wdh.											mit Fungizid
											ohne Fungizid

2 Fungizidstufen, 10 Sorten

# Spaltanlage

## Vorteil:

- Einfache Anlage auf dem Feld

## Nachteil:

- Statistische Analyse wird komplizierter (zweiter Fehlerterm: Großteilstückfehler)
- Weil mehr als ein Fehlerterm:  
„Gemischtes Modell“

# Standardfehler der Differenz für Vergleich der Stufen des Großteilstückfaktors

$$s_d = \sqrt{\frac{2 \cdot MQ_{ra}}{rb}}$$

Anzahl Freiheitsgrade =  $(a-1)(r-1)$

$MQ_{ra}$  *Großteilstückfehler (Behandlung\*Wdh)*

$rb$  *Anzahl Parzellen je  
Großteilstückfaktorstufe (= Anzahl  
Kleinteilstückfaktorstufen x Anzahl  
Wiederholungen)*



# Standardfehler der Differenz für Vergleich der Stufen des Kleinteilstückfaktors

$$s_d = \sqrt{\frac{2 \cdot MQ_e}{ra}}$$

Anzahl Freiheitsgrade =  $a(b-1)(r-1)$

$MQ_e$  Kleinteilstück- bzw. Restfehler

$ra$  Anzahl Parzellen je  
Kleinteilstückfaktorstufe (= Anzahl  
Großteilstückfaktorstufen  $\times$  Anzahl  
Wiederholungen)

# Standardfehler der Differenz für Vergleich der Stufen des Kleinteilstückfaktors auf gleicher Stufe des Großteilstückfaktors

$$s_d = \sqrt{\frac{2 \cdot MQ_e}{r}}$$

Anzahl Freiheitsgrade =  $a(b-1)(r-1)$

$MQ_e$  Kleinteilstück- bzw. Restfehler

$r$  Anzahl Wiederholungen

# Standardfehler der Differenz für Vergleiche von Kleinteilstückfaktorstufen auf unterschiedlichen Großteilstückfaktorstufen

$$S_d = \sqrt{\frac{2 \cdot [MQ_{ra} + (b - 1) \cdot MQ_e]}{rb}}$$

$$FG = \frac{a \cdot (r - 1) \cdot (a - 1) \cdot [(b - 1) \cdot MQ_e + MQ_{ra}]^2}{(a - 1) \cdot (b - 1) \cdot MQ_e^2 + (a) \cdot MQ_{ra}^2}$$

$MQ_{ra}$  Großteilstückfehler

$MQ_e$  Kleinteilstück- bzw. Restfehler

$a$  Anzahl Großteilstückfaktorstufen

$b$  Anzahl Kleinteilstückfaktorstufen

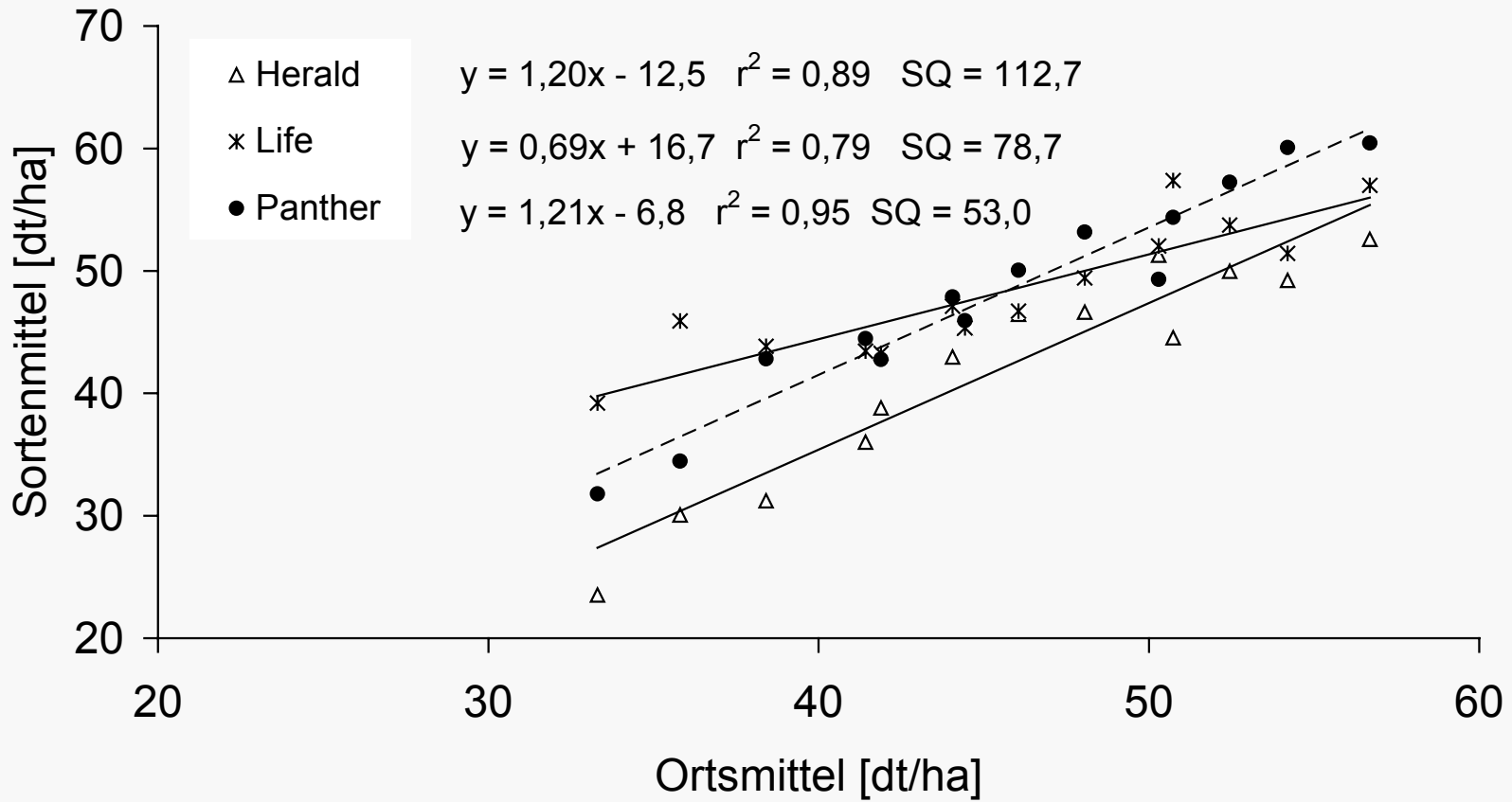
$r$  Anzahl Wiederholungen

## **4. Versuchsserien**

# Warum Versuchserien wichtig sind

- Leistung einer Sorte kann erst nach Prüfung in verschiedenen Umwelten endgültig beurteilt werden
- Nicht nur Mittelwert ist interessant
- Varianz = Ertragsstabilität

# Ertragsstabilität



# Auswertung von Versuchsserien

- Varianz der Genotyp-Umwelt-Interaktion berechnen
- Wenn diese klein > „alles in Butter“
- Wenn groß > evtl. Rangfolgeänderungen der Sorten zwischen den Umwelten
- Sorte x Ort-Interaktion ist für Beratung wichtig: Lokal- bzw. Regionalsorten!?
- Sorte x Jahr ist für den Landwirt wichtig: Da stecken Chancen und Enttäuschungen drin

# Genotyp-Umwelt-Interaktionen

Varianzkomponenten für Ertrag in Wertprüfungen (LAIDIG 1988)

Art	Genotyp	G*Ort	G*Jahr	G*O*J	Fehler
Weizen	9,2	2,2	3,4	12,8	16,4
Gerste	4,6	1,8	1,2	6,4	13,4
Raps	7,9	0,8	1,2	5,1	9,7
Kartoffel	13,8	3,9	6,2	10	9,2
Zuckerr.	16	1,8	2,5	2,6	42,2
Silomais	76,8	14,8	13,6	56,5	105,2



# Standardfehler der Differenz in Serien

$$SED(M_1 - M_2) = \sqrt{2 \left( \frac{\sigma_{\alpha\gamma}^2}{m} + \frac{\sigma_{\alpha\beta}^2}{n} + \frac{\sigma_{\alpha\beta\gamma}^2}{mn} + \frac{\sigma^2}{mnr} \right)}$$

$\sigma_{\alpha\gamma}^2$  = Varianz der Interaktion Sorte\*Jahr  $[(\alpha\gamma)_{ik}]$

$\sigma_{\alpha\beta}^2$  = Varianz der Interaktion Sorte\*Ort  $[(\alpha\beta)_{ij}]$

$\sigma_{\alpha\beta\gamma}^2$  = Varianz der Interaktion Sorte\*Jahr\*Ort

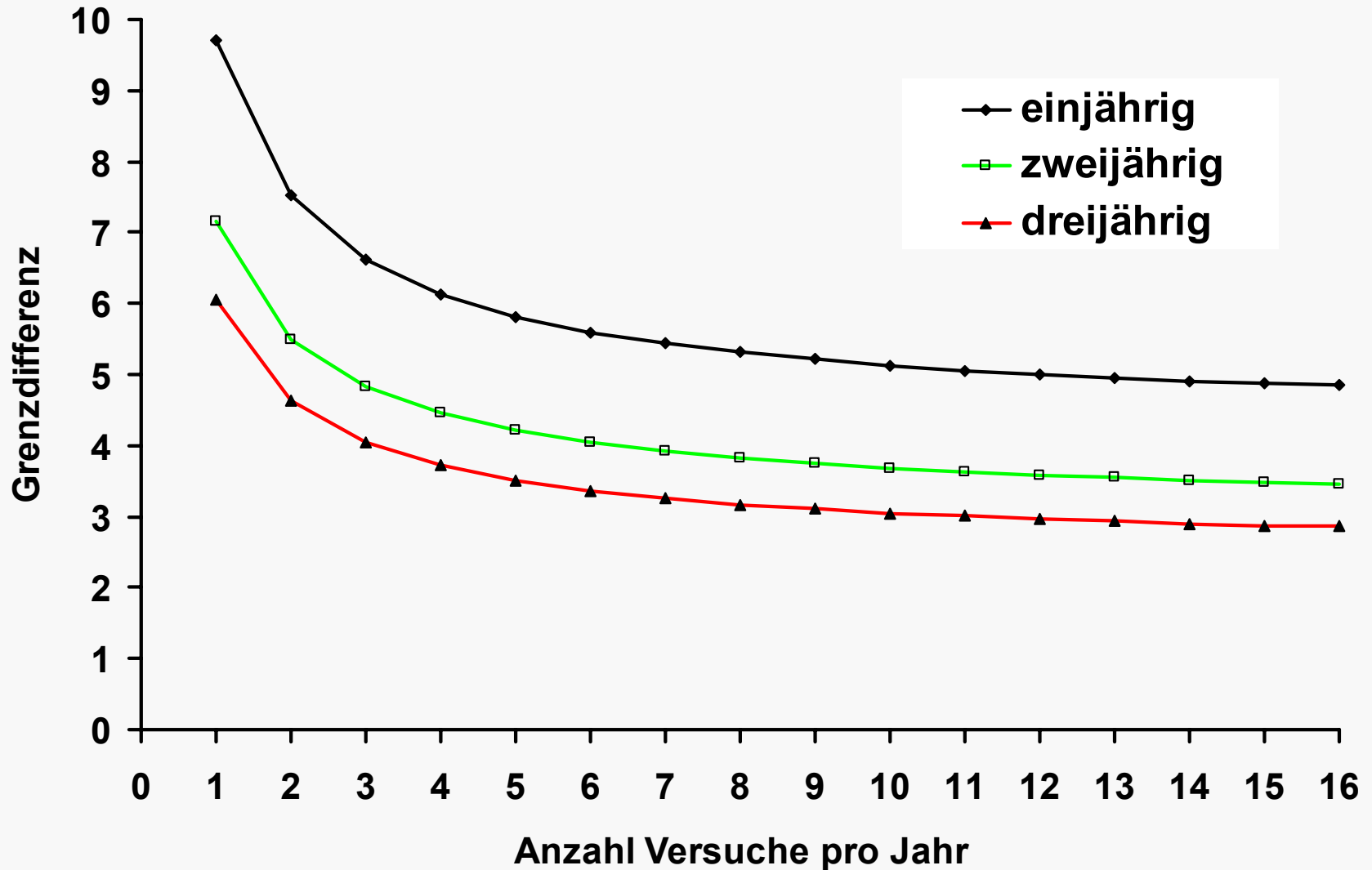
$\sigma^2$  = Fehlervarianz  $[e_{ijk_r}]$

$m$  = Anzahl Jahre

$n$  = Anzahl Orte

$r$  = Anzahl Wiederholungen

# Theoretische Grenzdifferenz in Serien



# Mehrstufigenselektion

- Einjährige Daten sind unsicher
- Selektion auf mehrjähriger Basis besser
- Schrittweise Selektion
  - Im ersten Jahr 50% beste selektieren
  - Im zweiten Jahr wiederum 50% beste (=25% der Linien aus 1. Jahr) auf Basis zweijähriger Daten
  - Im dritten Jahr endgültige Entscheidung auf dreijähriger Datenbasis

# **5. Planung eines Sortenversuches**

# Planung

- Wie viele Standorte?
  - frühe Generationen wenige (evtl. nur 1 Ort)
  - Später mehr (5 Orte bei meisten Arten ausreichend)
- Wie viele Wiederholungen?
  - Je mehr Orte, desto weniger Wdh. sind nötig
  - Mehr Orte sind immer besser als viele Wdh.
  - Für züchterische Versuche 2 Wdh. eigentlich ausreichend

# Planung

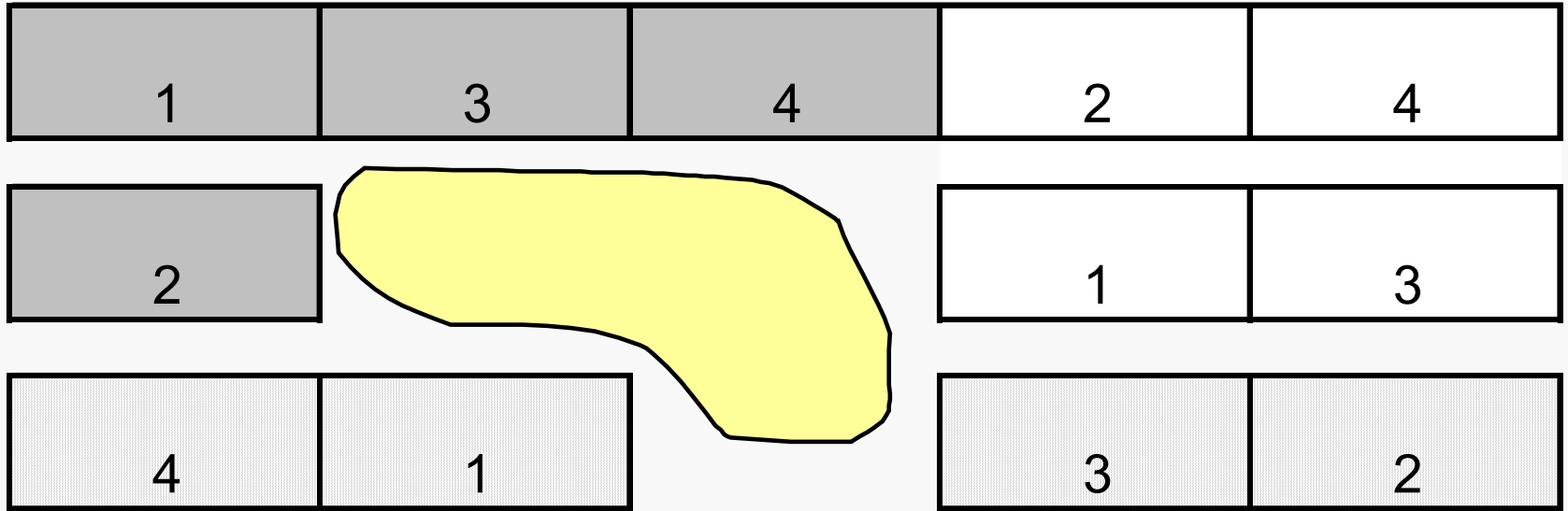
- Welcher Standort?
  - Soll repräsentativ sein für späteres Anbaugebiet
  - Homogene Bodenverhältnisse
  - Für Zuchtziel passendes Klima
  - In der Nähe von Zuchtstation
  - Hohe Korrelation zu WP-Standorten
  - Kompromiss finden

**Bewusste und ausschließliche Auswahl günstiger Standorte schränkt die Übertragbarkeit auf die Grundgesamtheit aller Standorte ein.**

# Planung

- Lage der Versuchsfläche / Design
  - Abstand halten zu Störfaktoren (z.B. Waldrand)
  - Fluss- oder Bachnähe, nasse Stellen, Senken, Kuppen, frühere Wege, Schatten von Bäumen und Gebäuden meiden

# Blöcke um Senke herum legen



3 Wiederholungen

4 unvollständige Blöcke je Wdh.



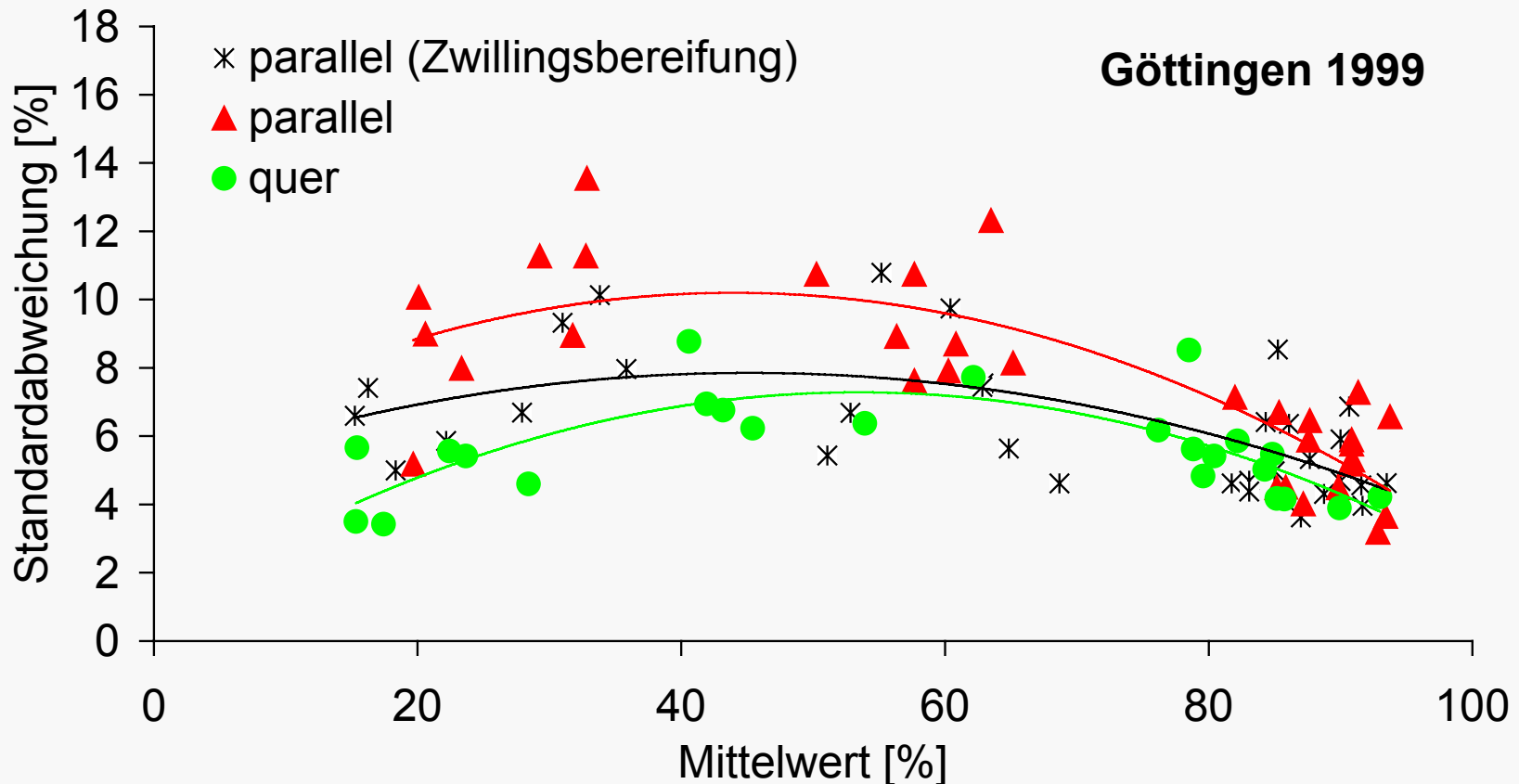
# Planung

- Lage der Versuchsfläche / Design
  - Abstand halten zu Störfaktoren (z.B. Waldrand)
  - Fluss- oder Bachnähe, nasse Stellen, Senken, Kuppen, frühere Wege, Schatten von Bäumen und Gebäuden meiden
  - Homogenen Teil des Schlages wählen (Erfahrungen der Vorjahre berücksichtigen)
  - Richtung der Bodenbearbeitung und Aussaatrichtung beachten



# Richtung der Saatbettbereitung

**Mittelwert und Standardabweichung des Feldaufganges von jeweils 12 benachbarten Zuckerrübenreihen in Abhängigkeit von der Richtung der Saatbettbereitung**



# Wie groß müssen die Teilstücke sein?

- Parzelle muss groß genug sein, um
  - das Ein- und Ausheben der Säaggregate zu ermöglichen
  - Bonituren sinnvoll durchführen zu können
- Parzelle soll klein sein, um
  - Versuchsfläche nicht zu groß werden zu lassen (Bodenunterschiede gering halten!)
  - intensiv betreut werden zu können

# Optimale Teilstückgröße

Fehlervarianz im Parzellenversuch:

$$\sigma^2 = \sigma_p^2 + \frac{\sigma_e^2}{n}$$

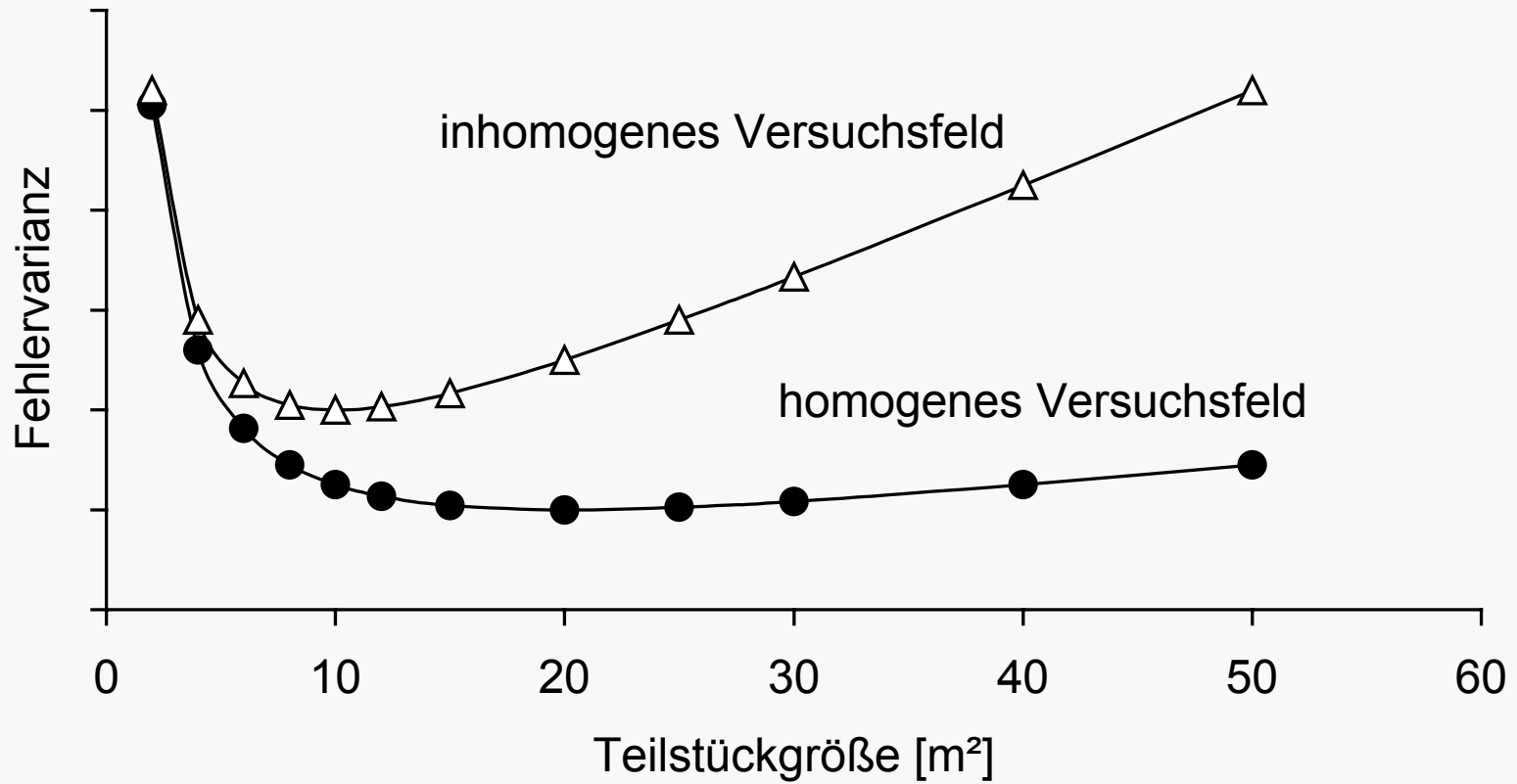
$\sigma^2$  = Varianz der Parzellenwerte  
(Fehlervarianz)

$\sigma_p^2$  = Varianz aufgrund von  
Bodenunterschieden

$\sigma_e^2$  = Varianz der Einzelpflanzen

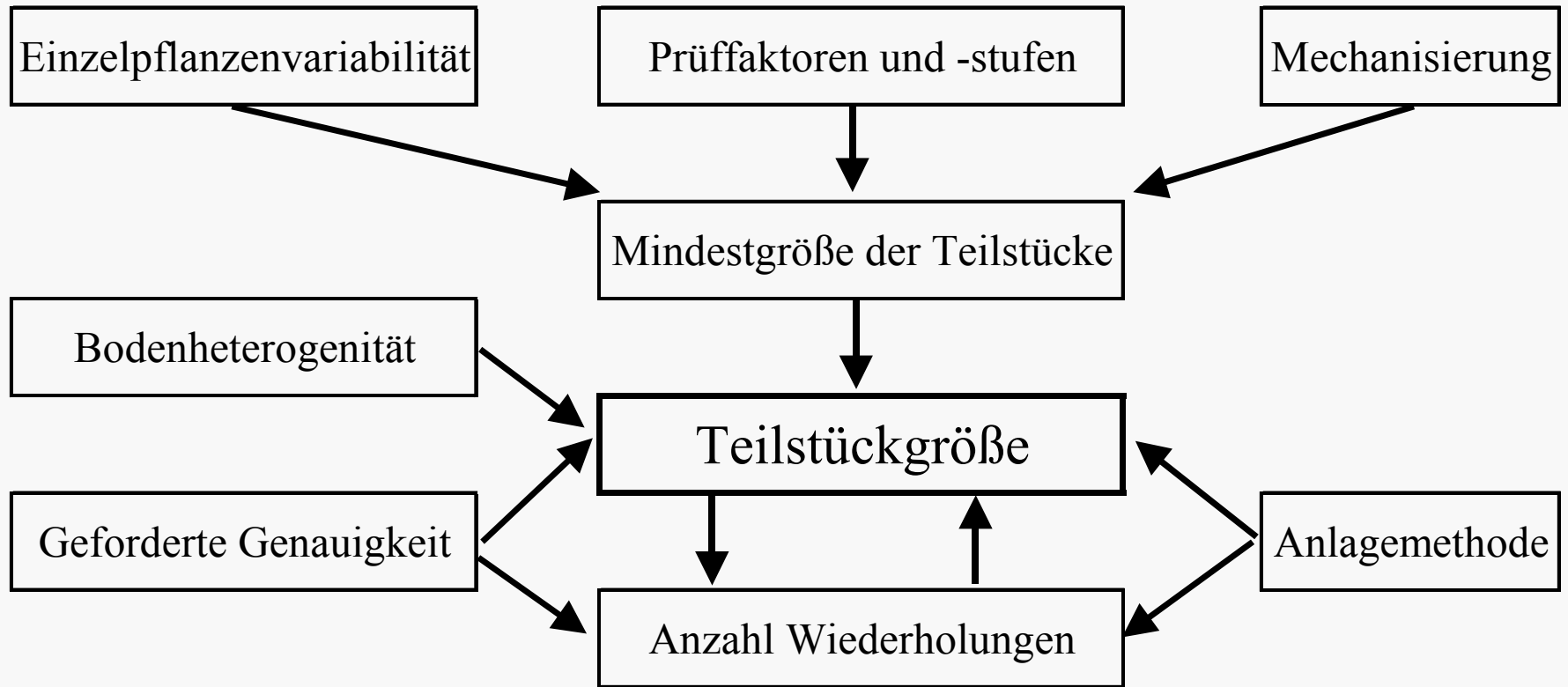
$n$  = Anzahl Pflanzen je Parzelle

# Teilstückgröße



**10-20 m<sup>2</sup> sind meist optimal**

# Was alles zu berücksichtigen ist...





# Randomisation „lenken“

Schritt 1: Grundplan für Dreisatzgitter aufstellen

Wdh 1

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36

Wdh 2

1	7	13	19	25	31
2	8	14	20	26	32
3	9	15	21	27	33
4	10	16	22	28	34
5	11	17	23	29	35
6	12	18	24	30	36

Wdh 3

1	8	15	22	29	36
2	9	16	23	30	31
3	10	17	24	25	32
4	11	18	19	26	33
5	12	13	20	27	34
6	7	14	21	28	35

Block mit 6 Parzellen



# Prüfglieder mehrfach in 1 Spur!

Schritt 2: Mögliche Verteilungen auf dem Feld herstellen

a) 12 Parzellen nebeneinander

	1	2	3	4	5	6	7	8	9	10	11	12
Wdh. 1	13	14	15	16	17	18	19	20	21	22	23	24
	25	26	27	28	29	30	31	32	33	34	35	36
Wdh. 2	1	7	13	19	25	31	2	8	14	20	26	32
	3	9	15	21	27	33	4	10	16	22	28	34
	5	11	17	23	29	35	6	12	18	24	30	36
Wdh. 3	1	8	15	22	29	36	2	9	16	23	30	31
	3	10	17	24	25	32	4	11	18	19	26	33
	5	12	13	20	27	34	6	7	14	21	28	35

# Innerhalb der Blöcke tauschen

Schritt 3: Jedes Prüfglied nur einmal je Spur

a) 12 Parzellen nebeneinander

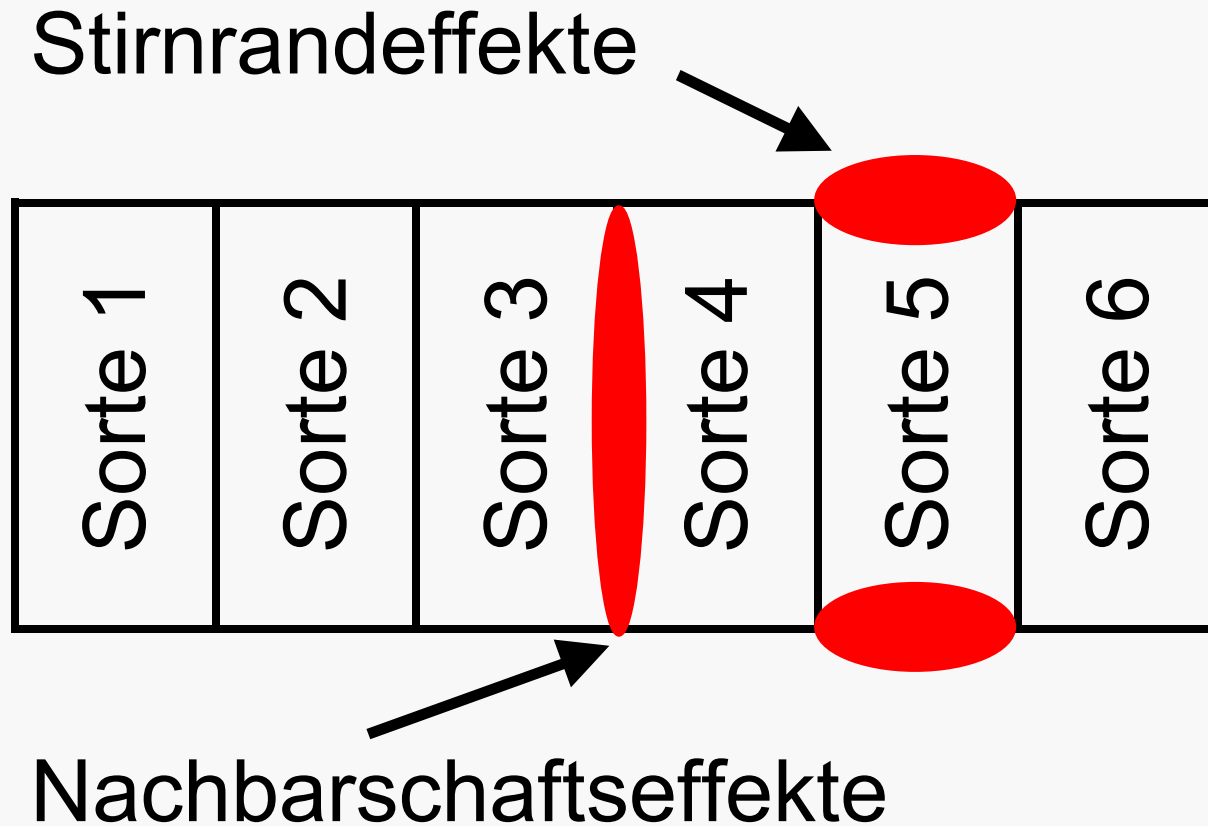
	1	2	3	4	5	6	7	8	9	10	11	12
Wdh. 1	13	14	15	16	17	18	19	20	21	22	23	24
	25	26	27	28	29	30	31	32	33	34	35	36
	31	1	7	13	19	25	2	14	20	26	32	8
Wdh. 2	33	3	9	15	21	27	4	10	16	28	34	22
	35	5	11	17	23	29	6	12	18	24	36	30
	15	8	1	29	22	36	16	9	2	23	30	31
Wdh. 3	3	10	17	24	25	32	18	11	4	19	26	33
	5	12	13	20	27	34	14	7	6	21	28	35

# Abschließend randomisieren

- Der Grundplan auf der vorigen Folie ist abschließend zu randomisieren:
- durch zufälliges Vertauschen kompletter Spalten innerhalb eines Teilblocks (z.B. innerhalb der sechs linken Spalten und über alle drei Wiederholungen hinweg)
- Durch Vertauschen von Zeilen innerhalb Wiederholungen
- Im letzten Schritt Prüfgliednummern zufällig den Nummern für den Plan zuordnen

## **6. Spezielle Probleme**

# Stirrand- und Nachbarwirkungen



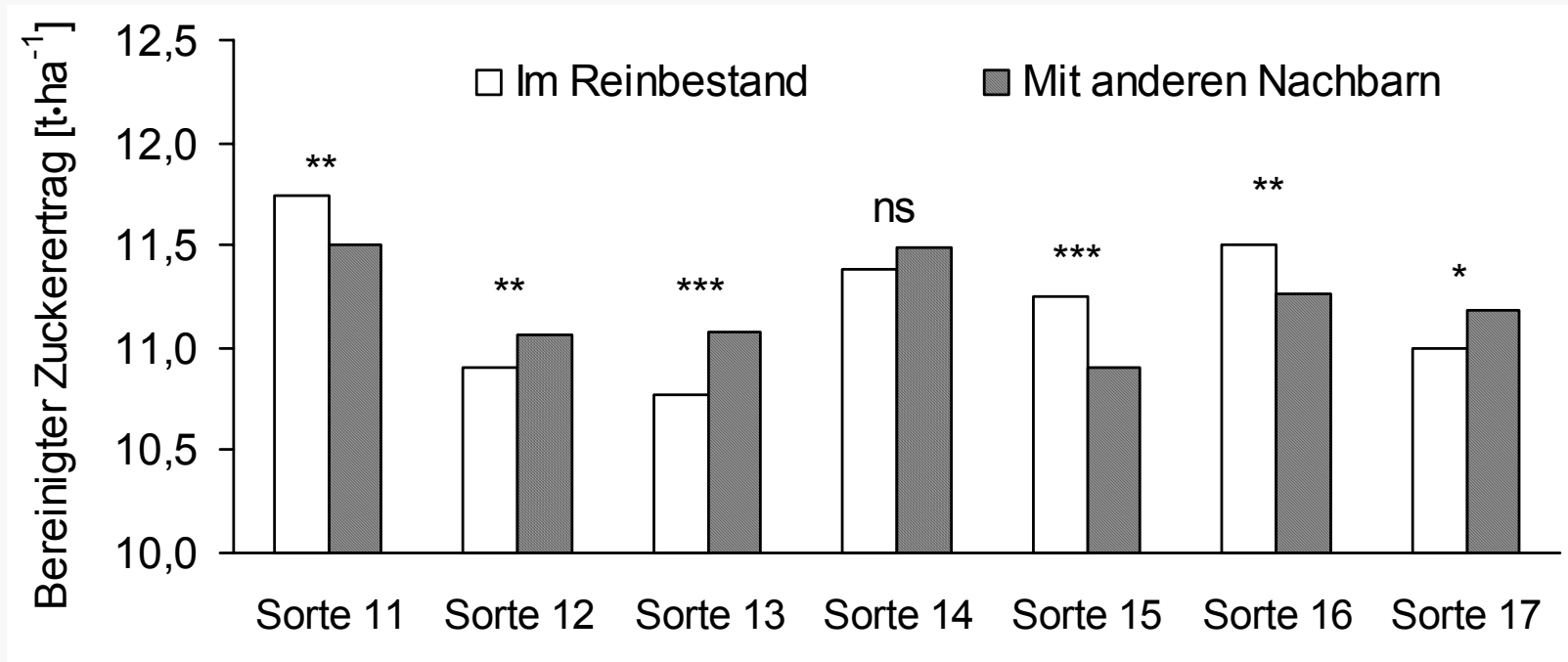






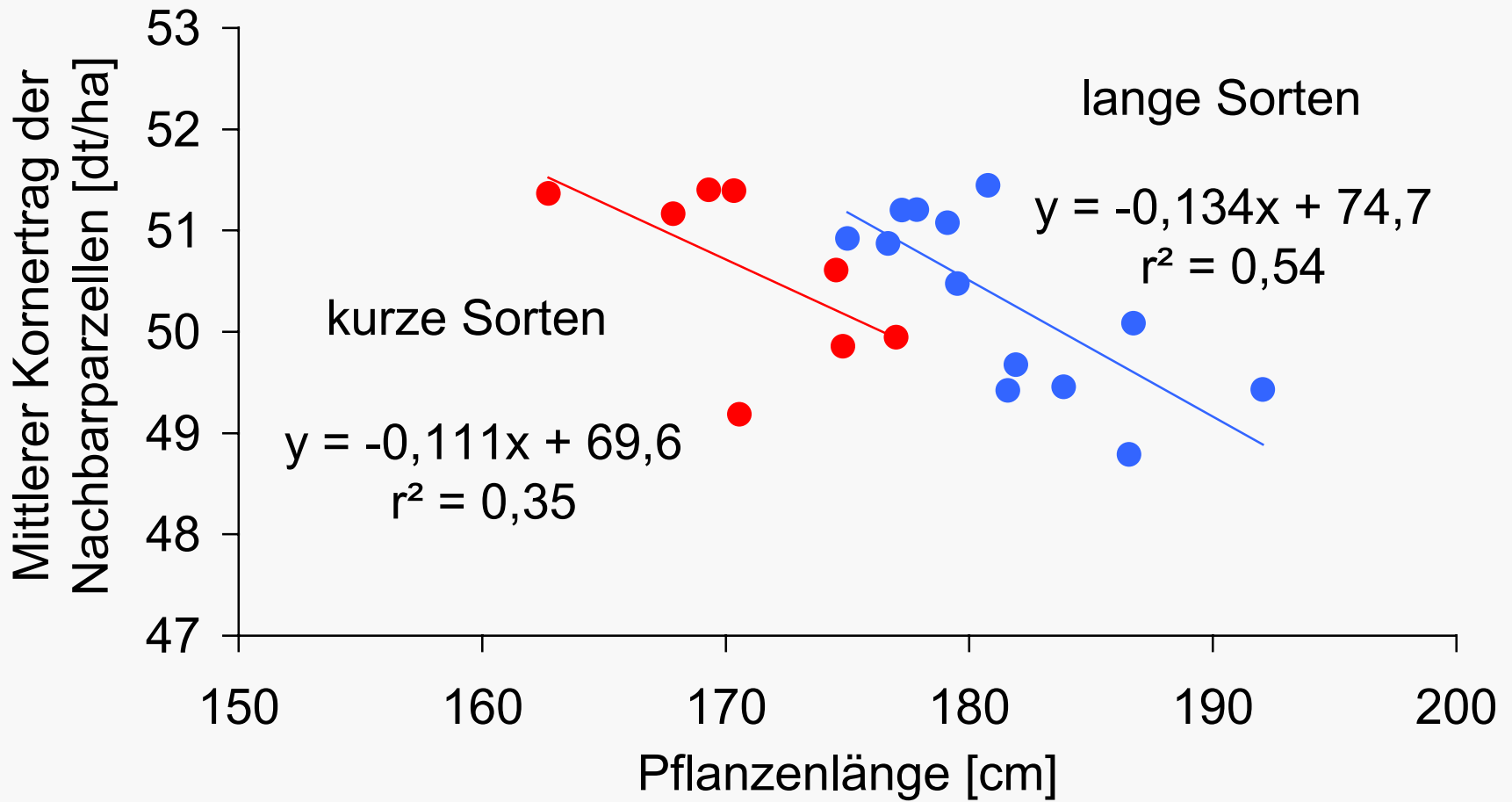


# Nachbarschaftseffekte bei Zuckerrüben





# Nachbarschaftseffekte bei Winterraps



# Ursachen

- Einfluss der Pflanzenlänge (Konkurrenz um Licht)
- Unterschiede im Wurzelwachstum (Konkurrenz um Nährstoffe) können vermutet werden
- Bei großen Unterschieden in der Wüchsigkeit kann auch Konkurrenz um „Raum“ eine Rolle spielen
- Lagerneigung oder Krankheitsanfälligkeit können ebenfalls Bedeutung haben
- **Verzerrungen von bis zu 5% (im Einzelfall auch >10%) im Ertrag sind möglich**

# Lösungen

## **a) Ernte des Parzellenkerns, Verwerfen der Randleihen**

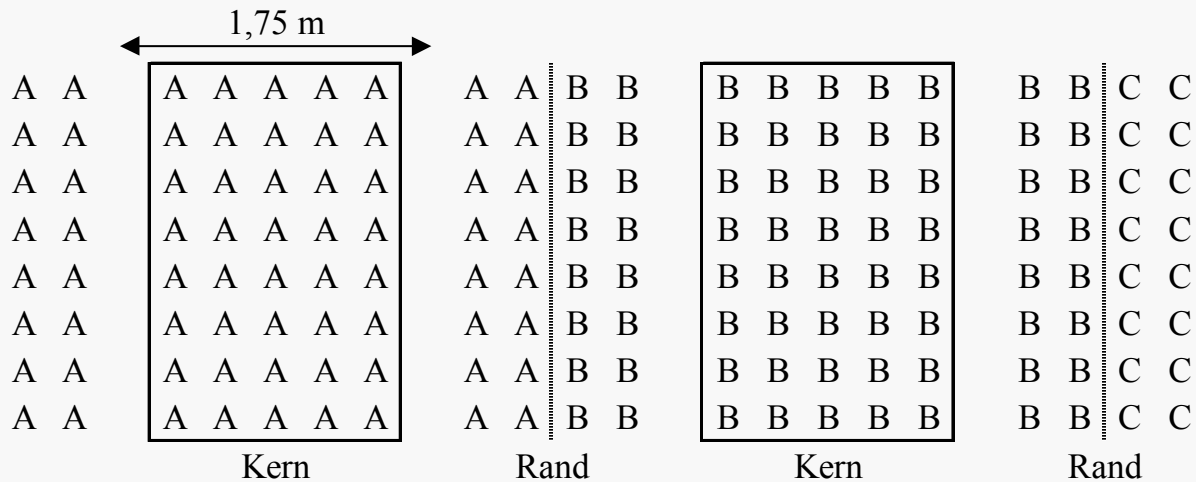
- Vorteil: Starke Reduktion der Effekte
- Nachteil: Hoher Flächenaufwand, größere Bodenunterschiede, je nach Fruchtart technisch aufwendig

## **b) Breitere Parzellen**

- Vorteil: Keine technischen Probleme, geringerer Versuchsfehler

Nachteil: Hoher Flächenaufwand, größere Bodenunterschiede, Effekte werden nur teilweise vermindert

# Anlage spezieller Kerndruschparzellen



# Lösungen

## c) Bildung von Teilsortimenten (lang/kurz)

- Vorteil: Geringer technischer Aufwand
- Nachteil: Kriterien zur Trennung nicht immer eindeutig
- Randparzellen notwendig
- Erhöhter Aufwand bei der Auswertung, weil Spaltanlage...

R	1	2	3	R	R	4	5	6	7	8	R	R	9	10	11	R	R	12	13	14	15	R
R	8	6	4	7	5	R	R	9	11	10	R	R	12	13	14	15	R	R	3	2	1	R
R	12	15	14	15	R	R	4	8	6	7	5	R	R	11	10	9	R	R	1	3	2	R
R	2	3	1	R	R	14	13	12	15	R	R	4	5	8	7	6	R	R	9	10	11	R

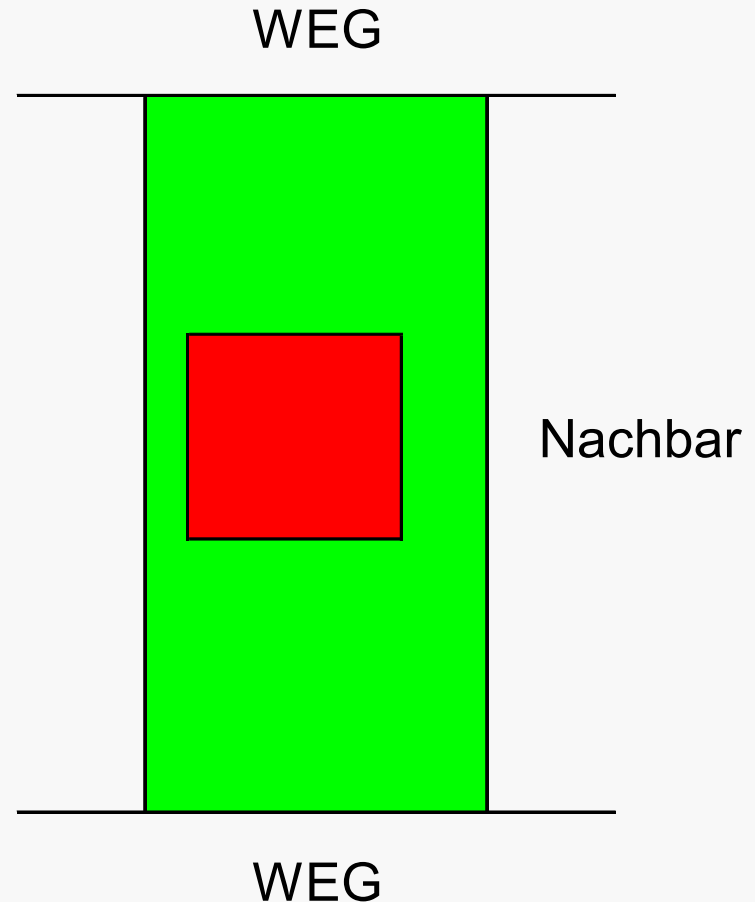
# Vollständige Trennung der Teilsortimente

R	1	2	3	R	R	4	5	6	7	8	R	R	9	10	11	R	R	12	13	14	1
R	3	2	1	R	R	8	6	7	4	5	R	R	9	11	10	R	R	15	14	13	1
R	1	3	2	R	R	4	7	6	8	5	R	R	11	10	9	R	R	12	15	14	1
R	2	3	1	R	R	5	4	8	6	7	R	R	9	10	11	R	R	14	12	15	1

Vergleich zwischen Teilsortimenten nur über Standards möglich, große Grenzdifferenz

# „Einkürzen“ von Fehlstellen

- Geschädigte Flächen (und Randpflanzen!) „ausschneiden“
- Hochrechnen des Ertrages auf die Standardgröße
- Achtung: Stirnrand und Nachbareffekte können bei Einkürzungen stark verzerren!



# Fazit

- Biometrische Prinzipien bereits bei Versuchsplanung berücksichtigen
- Geringer Versuchsfehler spart Geld (weniger Wdh.), erhöht Effizienz (höherer Selektionserfolg)
- Je enger die Sorten beieinander liegen, desto wichtiger ist exakte Versuchsdurchführung und ausreichende Anzahl Wdh.



## Fazit (2)

- Bei Standardanlagen evtl. Effizienzgewinn durch geostatistische Methoden möglich (NNA-BLUP); *Achtung: bei großen Nachbarschaftseffekten wird Korrelation zwischen Nachbarn negativ*
- Auch erste Wiederholung immer randomisieren
- Alpha-Gitter sind flexible Alternative zu Zwei- und Dreisatzgitter

## Fazit (3)

- Mehrfaktorielle Versuche bieten viel Information
- Spaltanlagen und komplexe Designs korrekt auswerten!
- Ertragsstabilität ist wichtige Sorteneigenschaft
- Sorte\*Jahr-Interaktion oft sehr groß
- Erst mehrjährige Daten geben Sicherheit
- 5 Orte pro Jahr meist ausreichend

## Fazit (4)

- Versuchstandorte sollen „gute“ Versuche erlauben aber repräsentativ für Anbaugebiet sein
- Besonderheiten auf dem Versuchsfeld berücksichtigen (nicht jeder Versuch passt an jeden Standort)
- Parzelle so klein wie möglich aber so groß wie nötig
- Prüfglieder möglichst nur 1 mal in Spur
- Rand- und Nachbarwirkungen beachten





**Danke für Ihre Aufmerksamkeit!**

