

## Anvendelse af ufuldstændige blokforsøg

Kristian Kristensen<sup>1</sup>, Jakob Willas<sup>2</sup>, Lise Nistrup Jørgensen<sup>3</sup> og Rene Gislum<sup>4</sup>

<sup>1</sup> *Forskergruppe for Biometri, Afd. for Husdyravl og Genetik, DJF*

<sup>2</sup> *Afd. for Sortsafprøvning, DJF*

<sup>3</sup> *Afd. for Plantebeskyttelse, DJF*

<sup>4</sup> *Afd. for Plantebiologi, DJF*

### Sammendrag

I forsøg med mange forsøgsled, forsøg på uensartede arealer og forsøg, som af praktiske grunde skal opdeles på mindre enheder, vil det ofte være fordelagtigt at anvende blokke, som ikke indeholder alle forsøgsled. Sådanne ufuldstændige blokke kan konstrueres efter flere forskellige principper. I enfaktorielle forsøg anvendes oftest metoder, som sikrer, at alle par af forsøgsled kan sammenlignes med den samme (og bedst mulige) sikkerhed, mens man i flerfaktorielle forsøg oftest anvender metoder, som sikrer, at de mest interessante effekter (hoved- og/eller vekselvirkninger) kan estimeres bedst muligt. Artiklen viser tre eksempler på anvendelse af ufuldstændige blokforsøg: 1) Sortsforsøg i ærter (enfaktoriel), 2) Frøavlsvorsøg med strandsvingel (trefaktoriel) og 3) Forsøg med svampemidler i byg (tofaktorielt split-plot). For hver af disse vises den benyttede plan samt effekten af at benytte planen i stedet et randomiseret blokforsøg med fuldstændige blokke.

**Nøgleord:** Generaliserede lattice forsøg,  $\alpha$ -forsøg, effektivitet, Lsd, enfaktoriel, flerfaktoriel, split-plot.

### Introduktion

En del forsøg har mange forsøgsled. Hvis sådanne forsøg skal anlægges med fuldstændige blokke betyder det at blokkene skal indeholde mange parceller, og blokkene dermed bliver store. Da jordens variabilitet (oftest) stiger med blokstørrelsen, betyder dette, at den tilfældige variation vil blive stor, når der er mange forsøgsled, og det kan så blive besværligt at påvise forskelle mellem forsøgsleddene. Der er flere måder, hvorpå man kan forsøge at løse dette problem.

I nogle tilfælde kan man gøre blokkene mindre ved at dele det nødvendige antal planter på flere parceller (gentagelser). F.eks. kunne man vælge at have 6 parceller á 10 planter for hver behandling i stedet for 3 parceller á 20 planter. Der er dog en grænse for, hvor langt denne proces kan fortsætte i praksis; dels fordi forsøget vil blive dyrere jo flere gentagelser man deler det nødvendige antal planter på, og dels fordi arealet i de enkelte blokke efterhånden

ikke reduceres nævneværdigt, da en stor del af denne nu vil blive brugt til værn og gangarealer.

Man kan forsøge at begrænse forsøgets størrelse ved at dele det i flere mindre forsøg. Herved begrænses størrelsen af de fuldstændige blokke, og derved opnås en sikrere sammenligning af de forsøgsled, som forekommer i samme forsøg. Samtidig har man dog gjort det umuligt at sammenligne forsøgsled, som ikke forekommer i samme forsøg. Derfor er en sådan opdeling af forsøget i flere mindre forsøg (kaldes af nogen serier) kun acceptabel, hvis forsøgsleddene kan deles på disse mindre forsøg, sådan at alle interessante sammenligninger m.m. kan foretages indenfor forsøget. Ved at medtage enkelte forsøgsled i alle forsøg kan man dog muliggøre sammenligning af forsøgsled i forskellige forsøg, men usikkerheden på sådanne indirekte sammenligninger er væsentlig større end sammenligning af forsøgsled i samme forsøg.

Alternativt kan man benytte forsøg med ufuldstændige blokke. Her indeholder blokkene ikke alle forsøgsled og bliver derved mindre. F.eks. kunne et forsøg med 80 forsøgsled med 3 gentagelser lægges ud i 12 ufuldstændige blokke med 20 forsøgsled i hver, eller i 30 ufuldstændige blokke med 8 forsøgsled i hver. Det betyder selvfølgelig, at hver blok kun vil indeholde en del af forsøgsleddene. Hvis forsøgsleddene fordeles tilfældigt på de ufuldstændige blokke, vil man ikke opnå en bedre sikkerhed - tværtimod, forsøget kan i værste tilfælde blive værdiløs. Fordeles forsøgsleddene på de ufuldstændige blokke så de interessante sammenligninger optimeres, kan anvendelsen af ufuldstændige blokke ofte være meget fordelagtigt.

Man skelner normalt mellem to grupper af ufuldstændige blokke:

1. I enfaktorielle forsøg anvendes oftest planer, hvor man sikrer sig, at alle par af forsøgsled kan sammenlignes med samme og (og størst mulig) sikkerhed.
2. I flerfaktorielle forsøg anvendes oftest planer, hvor man sikrer sig, at de interessante effekter (hoved eller vekselvirkninger) kan estimeres og testes med en stor sikkerhed, mens de mindre interessante (eller uinteressante) effekter så bliver estimeret med en lav sikkerhed (eller slet ikke kan estimeres).

## **Enfaktorielle forsøg**

Ufuldstændige blokforsøg har været anvendt i mange år. Metoden går tilbage til Yates (1939), og har gennem årene vist sig at være meget effektiv (se, f.eks. Le Clerg, 1966 eller Patterson and Hunter, 1983). I litteraturen findes mange forskellige typer af ufuldstændige blokforsøg. Her vil kun blive omtalt en type: generaliserede lattice forsøg (også kaldet  $\alpha$ -forsøg), som blev udviklet af Patterson and Williams (1976). Disse forsøg er valgt, dels fordi de er "resolvable" (de ufuldstændige blokke kan samles så de danner hele gentagelser), og dels

fordi de kan konstrueres for alle mulige antal forsøgsled og gentagelser samt for et meget bredt spektrum af blokstørrelser.

### **Eksempel**

I sortsafprøvningen med ærter var der i 2003 34 sorter, som skulle sammenlignes. Ved anvendelse af sædvanlige fuldstændige blokforsøg vurderedes det, at blokkene ville blive for store. Derfor blev det valgt at benytte et generaliseret lattice forsøg ( $\alpha$ -forsøg). Det blev valgt at benytte en plan med 4 gentagelser og 3 ufuldstændige blokke i hver gentagelse. To af de tre ufuldstændige blokke får 11 parceller, mens den sidste får 12 parceller (figur 1). Planen er her vist så de ufuldstændige blokke ligger i hver sin kolonne, men de kan placeres anderledes i marken, f.eks. i forlængelse af hinanden, så der bliver en kolonne for hver gentagelse.

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

**Figur 1. Fordeling af sorter på gentagelser og blokke for ærtforsøget. Hvert tal udgør en parcel og angiver sortsnummeret. Hver kolonne er en blok.**

Af planen fremgår det at alle par af sorter forekommer sammen i samme blok 0, 1 eller 2 gang (der er ingen par af sorter, som forekommer sammen mere end 2 gange). Der optimale er, at alle par af sorter forekommer sammen lige mange gange (f.eks. 1), men det er i praksis sjældent opnåeligt. Det fremgår endvidere, at de ufuldstændige blokke er samlet i hele gentagelser, eksempelvis indeholder de tre blokke længst til venstre alle sorter præcis 1 gang.

Ved placering i marken kan følgende 4 trin anvendes:

1. Vælge et så ensartet areal som muligt.
2. Inddel arealet i 4 lige store områder (4 gentagelser), således at variationen i hver gentagelse er så lille som muligt (= variationen mellem de 4 gentagelser skal være så stor som muligt).
3. Inddel hver af disse 4 gentagelser i 3 ufuldstændige blokke, således at variationen i hver ufuldstændig blok bliver så lille som muligt.
4. Inddel hver ufuldstændig blok i det nødvendige antal parceller, således at variationen mellem parceller bliver så lille som muligt

Analysen af et sådant ufuldstændigt blokforsøg kan foretages med den samme model (model 1 nedenfor), som for et forsøg med fuldstændige blokke (men for computeren er det lidt mere besværligt). Normalt vil man dog opdele effekten af blokke i 2 komponenter nemlig, den som skyldes de hele gentagelser, og den som skyldes de ufuldstændige blokke inden for gentagelser (model 2). I mange tilfælde er det rimeligt at antage, at effekten af de ufuldstændige blokke er tilfældig (model 3), og i så fald vil det ufuldstændige blokforsøg altid være lige så effektivt som et forsøg med fuldstændige blokke.

- (1)  $Y_{vb} = \mu + \alpha_v + \gamma_b + E_{vb}$       Effekt af blokke antages systematisk  
 (2)  $Y_{vrb} = \mu + \alpha_v + \beta_r + \gamma_{rb} + E_{vrb}$       Effekt af blokke indenfor gentagelser antages systematisk  
 (3)  $Y_{vrb} = \mu + \alpha_v + \beta_r + C_{rb} + E_{vrb}$       Effekt af blokke indenfor gentagelser antages tilfældig  
 $E_{vb}$ ,  $E_{vrb}$  og  $C_{rb}$  antages uafhængige og normal fordelte med konstant varians,  $\sigma_E^2$  og  $\sigma_C^2$

Ved anvendelse af ufuldstændige blokke forekommer alle forsøgsled, som nævnt, ikke i alle blokke. Dette bevirker, at en sort, som har været i højt ydende blokke vil have et gennemsnit som er for højt, mens en sort, som har været i lavt ydende blokke vil have et gennemsnit, som er for lavt. Derfor kan simple gennemsnit ikke benyttes til at sammenligne sorterne. I stedet må man benytte de sortsestimater som fås ved udtrykket:  $\hat{\mu} + \hat{\alpha}_v$ . Man kan betragte dette udtryk som simple gennemsnit korrigeret for, om den pågældende sort har været placeret i ufuldstændige blokke, hvor udbyttet i gennemsnit har været lavere end gennemsnittet eller højere end gennemsnittet.

### Effektivitet

Sortsforsøget med ærter blev udført på 4 lokaliteter. I tabel 1 er Lsd-værdierne vist for hver af disse forsøg sammen med den Lsd-værdi man ville have fået, hvis man havde anvendt sædvanlige fuldstændige blokke. Lsd-værdierne blev reduceret noget i to af forsøgene, meget i et forsøg og slet ikke i et forsøg. I tre af forsøgene har man haft en fordel af at benytte  $\alpha$ -design. I forsøgene på lokalitet 3 og 4 kunne man have opnået den samme Lsd-værdi med sædvanlige fuldstændige blokke ved at øge antal gentagelser med 1, mens man på lokalitet 2 skulle have anvendt 4 ekstra gentagelser (d.v.s. i alt 8) for at reducere Lsd-værdien med 29%.

**Tabel 1. Lsd-værdier for sammenligning af sorter på hver af de 4 lokaliteter ved anvendelse af to forsøgstyper samt opnået reduktion ved anvendelse af  $\alpha$ -forsøg**

Lokalitet	Fuldstændige blokke	$\alpha$ -forsøg gens. (min - max)	Relativ reduktion, %
1	4.8	4.8 (4.8-4.8)	0
2	6.9	4.9 (4.7-4.9)	29
3	6.4	6.0 (5.8-6.0)	9
4	3.7	3.4 (3.3-3.5)	8

### ***Fordele og ulemper ved ufuldstændige blokke i enfaktorielle forsøg***

Fordelene er først og fremmest at de gør det muligt at have mange forsøgsled (sorter, pesticider m.m.) i samme forsøg og stadig være i stand til at sammenligne alle forsøgsled med god sikkerhed. Fordelene er især store på marker, hvor det kan være svært at placere tilstrækkelig ensartede fuldstændige blokke. De ufuldstændige blokforsøg kan også benyttes som en slags forsikring mod uventet store variationer i marken, da de altid (med antagelsen om tilfældige blokeffekter) vil være lige så effektive som tilsvarende forsøg med fuldstændige blokke (Yates, 1940). I marken vil de ufuldstændige blokforsøg ikke kunne skelnes fra forsøg med fuldstændig blokke.

Ulemperne er, dels en lidt mere besværlig proces ved konstruktion og planlægning, og dels at alle forsøgsled sjældent vil blive sammenlignet med præcis den samme sikkerhed, idet nogle par af forsøgsled vil forekomme sammen oftere end andre par. Dette vil eksempelvis betyde, at man ved publiceringen ofte ikke blot kan anføre en Lsd-værdi, da denne vil afhænge af, hvilket sortspar man vil sammenligne. Forskellene mellem mindste og største Lsd-værdi er dog oftest kun nogle få procent (se tabel 1). Endelig skal nævnes, at ufuldstændige blokforsøg er lidt mere følsomme overfor manglende værdier samt for ombytning af forsøgsled (ved fejlsåning, forkert behandling m.m.).

### **Flerfaktorielle forsøg**

Forsøg med flere faktorer får hurtigt mange forsøgsled. Eksempelvis vil man i et forsøg med 4 faktorer, som hver har 3 niveauer, få 81 forsøgsled. I princippet kunne her anvendes samme metode som ovenfor, men oftest vil det være mere fordelagtigt at prioritere de enkelte effekter i planen, således at man sikrer en god og sikker estimation af de interessante effekter på bekostning af en mindre sikker estimation af de mindre interessante (eller uinteressante) effekter. En effekt kan være mindre interessant, f.eks. fordi den ikke er vigtig for at besvare det spørgsmål der skal belyses, eller fordi den kun er væsentlig, hvis den giver et meget stort udslag. Også denne type af forsøg har været benyttet i mange år og går helt tilbage til Yates (1935). For forsøg, hvor alle faktorer har  $2^p$  niveauer eller  $3^p$  niveauer er konstruktionen af sådanne planer rimelig simpel (se f.eks. Cochran & Cox, 1957). I andre tilfælde må man enten finde planen i litteraturen eller konstruere den ved en eller anden ad hoc. metode.

### ***Eksempel***

I et forsøg med strandsvingel til frøavl ville man undersøge effekterne af følgende tre faktorer samt deres tovejs vekselvirkninger: 2 dækafgrøder, 4 niveauer af N-tildeling om efteråret og 4 niveauer af N-tildeling om foråret. Forsøget blev udlagt med 4 gentagelser. Hvis forsøget skulle udlægges som et almindeligt randomiseret blokforsøg ville der blive 32 parceller i hver

blok. Med den variation man sædvanligvis har i den type forsøg blev det anset for nødvendigt at reducere antallet ved at anvende ufuldstændige blokforsøg efter følgende principper:

- De hele gentagelser skulle underinddeles i 2 blokke á 16 parceller.
  - Vekselvirkningen mellem de 3 faktorer var mindre interessant og kunne kobles til blokke.
- De 32 behandlinger blev i dette forsøg delt på 2 blokke ved - fra tre-vejs vekselvirkningen - at udvælge 4 ortogonale kontraster med hver 1 frihedsgrad, og derefter koble disse til blokke i hver sin gentagelse. De effekter, der er knyttet til disse 4 kontraster (een i hver gentagelse) bliver således bestemt dårligere end de øvrige, da de nu kun kan estimeres i de 3 gentagelser, hvor de hver især ikke er koblet (figur 2).

Ved gennemgang af figuren finder man, at alle kombinationer af 1 og 2 faktorer forekommer lige mange gange i hver blok. Eksempelvis forekommer hver af de 4 niveauer af N-tildeling om foråret præcis 4 gange i hver blok. Derimod findes kombinationen "*dækafgrøde 1, N-tildeling 1 om efteråret og N-tildeling 1 om foråret*" 1 gang i blokken længst til venstre mens kombinationen "*dækafgrøde 1, N-tildeling 1 om foråret og N-tildeling 2 om efteråret*" ikke forekommer i denne blok.

Gentagelse 1	Gentagelse 2	Gentagelse 3	Gentagelse 4
232	224	143	231
111	114	244	144
144	143	142	243
241	112	113	114
131	121	124	212
214	231	134	242
133	242	121	122
122	213	232	133
223	132	112	111
142	141	131	221
124	134	211	141
221	211	214	234
234	123	222	213
113	244	241	132
212	233	223	123
243	222	233	224

**Figur 2. Markplan for forsøget med strandsvingel til frøavl. Hvert 3-cifret tal udgør en parcel og cifrene angiver den aktuelle behandling i rækkefølgen dækafgrøde, niveau af N-tildeling om efteråret og niveau af N-tildeling om foråret. Hver kolonne er en blok.**

Ved placering i marken bruges de samme principper som ovenfor. Randomiseringen er udført i 3 trin. Først er placeringen af de 4 gentagelser randomiseret, derefter er placeringen af de 2 blokke randomiseret i hver gentagelse, og til sidst er rækkefølgen af de 16 behandlinger i hver blok randomiseret.

Analysen af forsøget foretages også her med en model, som næsten er identisk med den, der anvendes for et forsøg med fuldstændige blokke.

### *Effektivitet*

I forsøget blev der udført telemålinger. Planen forbedrede kun sikkerheden på de ukoblede effekter med ca. 3% for denne variabel (tabel 2). For de ukoblede effekter er effektiviteten, som forventet, uafhængig af om blokeffekterne antages systematiske eller tilfældige. Når blokeffekterne antages systematiske er nogle af sammenligningerne i tre-vejs tabellen, som forventet, mindre sikre ved anvendelse af ufuldstændige blokke end ved anvendelse af fuldstændige blokke. For de koblede parvise sammenligninger i tre-vejs tabellen er effektiviteten bedre når blokeffekterne kan antages at være tilfældige. I eksemplet er der imidlertid kun 3 frihedsgrader til bestemmelse af varianskomponenten for ufuldstændige blokke, hvorfor det kan være betænkeligt at anvende antagelsen om tilfældige blokeffekter ved estimation af de koblede effekter.

**Tabel 2. Lsd-værdier for sammenligning af 3 effekter på telemåling ved anvendelse fuldstændige blokke, ufuldstændige blokke med systematiske blokeffekter og ufuldstændige blokke med tilfældige blokeffekter.**

Effekt	Fuldstændige blokke	Systematiske blokeffekter	Tilfældige blokeffekter	
Dækafgrøde	0.0068	0.0066	0.0066	
N-niveau efterår	0.0096	0.0093	0.0093	
Dækafgrøde × N-niveau efterår	0.0135	0.0132	0.0132	
3-vejs vekselvirknings tabel	0.0270	0.0263	0.0263	48 par
	0.0270	0.0269	0.0260	384 par
	0.0270	0.0274	0.0269	64 par

### *Split-plot*

Split-plot forsøget er et særtilfælde af ovenstående metode. Eksempelvis et forsøg, hvor man skal sammenligne effekten af 6 svampebehandlinger på 4 bygsorter i 3 gentagelser. Her har man 3 effekter at prioritere imellem: 1) svampebekæmpelse 2) sorter og 3) vekselvirkning mellem disse. Hvis formålet med forsøget er at vurdere de 6 svampebehandlinger, vil det ikke være så interessant at sammenligne hovedeffekten af sorter, og det kan vælges at koble denne til blokke. Her er benyttet et forsøg med 4 blokke á 6 parceller i hver gentagelse (figur 3). Kan blokke indenfor gentagelser antages at være tilfældig, hvad vi oftest gør, kan effekten af sorter estimeres og testes, men med en større usikkerhed, dels fordi der kun vil være 6 frihedsgrader til bestemmelse af variationen mellem blokke indenfor gentagelser, og dels fordi denne variation sikkert vil være ret stor.

Gentagelse 1				Gentagelse 2				Gentagelse 3			
4f	3f	2f	1f	2a	1c	3b	4a	3e	4e	1b	2e
4e	3e	2e	1e	2d	1f	3f	4b	3b	4b	1e	2a
4d	3d	2d	1d	2c	1q	3a	4f	3f	4c	1c	2f
4c	3c	2c	1c	2f	1d	3d	4e	3c	4d	1d	2b
4b	3b	2b	1b	2e	1e	3c	4c	3d	4a	1a	2d
4a	3a	2a	1d	2b	1b	3e	4d	3a	4f	1f	2c

**Figur 3. Fordeling af sorter og behandlinger på gentagelser og blokke for bygforsøget. Tallet angiver sortsnummeret, mens bogstavet angiver behandlingen. Hver kolonne er en blok.**

### *Effektivitet*

Planens effektivitet er vurderet ved at sammenligne Lsd-værdierne med dem man ville have fået, hvis man havde anlagt forsøget som et sædvanligt randomiseret blokforsøg (tabel 3). Som det ses af tabellen, halveres Lsd-værdien for sammenligning af svampebehandlinger, hvis det er tilstrækkeligt at betragte hovedvirkningen (d.v.s. ingen vekselvirkning). Er det nødvendigt at sammenligne behandlingseffekterne ved hver enkelt sort, halveres Lsd-værdien næsten. Derimod bliver det ved anvendelse af split-plot planen meget mere usikkert at sammenligne de gennemsnitlige sorteffekter, idet Lsd-værdien for denne sammenligning stiger med næsten 70%.

**Tabel 3. Lsd-værdier for sammenligning af behandlinger og sorter ved anvendelse af et randomiseret blokforsøg og det benyttede split-plot forsøg.**

Effekt	Randomiseret	
	blokforsøg	Split-plot forsøg
Svampebehandling	4.6	2.3
Sorter	3.7	6.2
Vekselvirkning	9.1	4.7 (samme sort)
	9.1	7.5 (forskellige sorter)

### ***Fordele og ulemper ved anvendelse af ufuldstændige blokke i flerfaktorielle forsøg***

Også her er den største fordel, at man kan medtage mange forsøgsled (kombinationer af flere faktorer) og stadig kunne anvende passende små blokke, og dermed undgå at variationen mellem parceller (i store blokke) bliver for stor.

Ulemperne er, at man må prioritere og dermed acceptere at visse effekter enten ikke kan estimeres eller kun kan estimeres med en stor usikkerhed. Desuden skal nævnes, at denne type ufuldstændige blokforsøg er mere følsom overfor antagelsen om at blokeffekterne er additive.



Ikke additive blokeffekter for en effekt (f.eks. hovedeffekten A) kan i visse tilfælde give falske udslag for en anden effekt (måske vekselvirkningen B×C).

## **Konklusion**

De tre eksempler viser, at der i nogle tilfælde - men ikke alle - kan opnås betydelig sikrere estimer ved at anvende ufuldstændige blokforsøg, når der er mange forsøgsled. Hvis blokeffekterne kan antages tilfældig vil forsøg med de viste typer af ufuldstændige blokke aldrig være dårligere end tilsvarende forsøg med fuldstændige blokke. I flerfaktorielle forsøg med systematiske blokeffekter vil de ukoblede effekter aldrig være dårligere. Da der oftest kun er lidt ekstra arbejde i forbindelse med planlægningen, konstruktionen og analysen af sådanne forsøg anbefales det at overveje om man skal bruge forsøg med ufuldstændige blokke, hvis man har mange forsøgsled eller af anden grund vil få store fuldstændige blokke.

## **Referencer**

- Cochran; W.G., and Cox, G.M., 1957. Experimental Designs. Second Edition. John Wiley & Sons, Inc. New York. 611 pp.
- Le Clerg, E.L., 1966. Significance of Experimental Design in Plant Breeding. Plant Breeding. 243-313.
- Patterson, H.D., and Williams, E.R., 1976. A new class of resolvable incomplete block designs, Biometrika 63, 83-92.
- Patterson H.D. and Hunter E.A., 1983. The efficiency of incomplete block designs in National List and Recommended List cereal variety trials. Jour. Agric. Sci. Camb. 101, 422-433.
- Yates, F., 1935. Complex experiments. Jour. Roy. Stat. Soc. Suppl. 2, 181-247.
- Yates, F., 1939. The recovery of inter-block information in variety trials arranged in three-dimensional lattices. Annals of Eugen. 9,136-156.
- Yates, F., 1940. The recovery of inter-block information in balanced incomplete block designs. Annals of Eugenies. 10, 317-325.