Lattice Designs Applications in Plant Breeding

Jennifer Kling Oregon State University



Presentation Outline

- Why use incomplete block designs?
- Distinguishing features
- Lattice designs basic plans and field layout
- Statistical analysis
- Cyclic designs and α-Lattices
- Meadowfoam breeding example
- Randomization and field plan
- SAS analysis and interpretation

Blocking

Purpose

- Reduce experimental error, increase precision
- Compare treatments (genotypes) under more uniform conditions

The problem

- In breeding programs, the number of treatments may be large
- As blocks get larger, the conditions within blocks become more heterogeneous
- Other factors may limit the number of experimental units in a block
 Size of a growth chamber, greenhouse, or available field area
 Number of samples that can be processed at one time

Incomplete Block Designs

 Plots are grouped into blocks that are not large enough to contain all treatments (genotypes)

Distinguishing features

- Number of blocking criteria
- Balanced or partially balanced
- Resolvable (or not)
- Shape (square, rectangular)
- Process for generating design



Types of Incomplete Block Designs

Single blocking criterion

- Randomized incomplete blocks
- Two blocking criteria based on Latin Squares
 - Latin Square is a complete block design that requires N=t². May be impractical for large numbers of treatments.
 - Row-Column Designs either rows or columns or both are incomplete blocks
 - Youden Squares two or more rows omitted from the Latin Square

Balanced Incomplete Block Designs

- Each treatment occurs together in the same block with every other treatment an equal number of times
 - t = number of treatments
 - k = number of units per block (block size)
 - b = total number of blocks in the experiment
 - r = number of replicates of each treatment
 - λ = number of times that treatments occur together in the same block
- All pairs of treatments are compared with the same level of precision even though differences between blocks may be large



Balanced Incomplete Block Designs

For balance, $r = \lambda(t-1)/(k-1)$

- t = number of treatments
- k = number of units per block (block size)
- b = total number of blocks in the experiment
- r = number of replicates of each treatment
- λ = number of times that treatments occur together in the same block
- N = total number of experimental units
- λ must be an integer; N = b*k = r*t
- Example: t = 10 treatments with block size k = 4

 \Rightarrow N = b*k = r*t = 6*10 = 60

In plant breeding, the minimum number of replications required for balance is often too large to be practical

Partially Balanced Incomplete Block Designs

- Different treatment pairs occur in the same blocks an unequal number of times or some treatment pairs never occur together in the same block
 - Mean comparisons have differing levels of precision

Greater precision when treatments occur in the same block

- Statistical analysis more complex
- Common for plant breeding trials due to large number of entries



Resolvable Incomplete Block Designs

- Blocks are grouped so that each group of blocks constitute one complete replication of the treatment
 - "block" = incomplete block = "sub-block"
 - replication = "superblock"
- Trials can be managed in the field on a rep-by-rep basis
- Field operations can be conducted in stages (planting, weeding, data collection, harvest)
- Complete replicates can be lost without losing the whole experiment
- If you have two or more complete replications, you can analyze as an RCBD if the blocking turns out to be ineffective

Lattice Designs are Resolvable

 Lattice designs are a well-known type of resolvable incomplete block design



s = number of blocks in each complete replication

3

Lattice Designs

Square lattice designs

- Number of treatments must be a perfect square (t = k²)
- Blocks per replicate (s) and plots per block (k) are equal (s = k) and are the square root of the number of treatments (t)
- For complete balance, number of replicates (r) = k+1

Rectangular lattice designs

- t = s*(s-1) and k = s-1
- Example: 4 x 5 lattice has 4 plots per block, 5 blocks per replicate, and 20 treatments

Alpha lattices

- t = s*k
- more flexibility in choice of s and k

The Basic Plan for a Square Lattice

Block	Rep I	Rep II	Rep III	Rep IV
1	123	147	159	168
2	456	258	267	249
3	789	369	348	357

- Balance each treatment occurs together in the same block with every other treatment an equal number of times
 - Once in this case, so $\lambda = 1$

Basic plans can be found in Cochran and Cox, 1957

Randomization

Field Arrangement

- Blocks composed of plots that are as homogeneous as possible
- Randomization of a Basic Plan
 - Randomize order of blocks within replications
 - Randomize the order of treatments within blocks



Example of Randomization of a 3 x 3 Balanced Lattice (t = 9)

1 Assign r random numbers

Random	Sequence	Rank
372	1	2 ←
217	2	1
963	3	4
404	4	3

2 From basic plan Block Rep III **Rep IV** Rep I Rep II 123 147 168 1 159 456 258 267 249 3 789 369 348 357

3 Randomize order of replications

4 Randomize blocks within reps

Block	Rep I	Rep II	Rep III	Rep IV	1	Rep	I	П		IV	
1	147_	123	168	159			3	2	3	1	
2	258	456	249	267			2	1	1	3	
3	369	789	357	348			1	3	2	2	

5 Resulting new plan										
Block	Rep l	Rep II	Rep III	Rep IV						
1	369	456	357	159						
2	258	123	168	348						
3	147	789	249	267						

Partially Balanced Lattices

Simple Lattices

- Two replications use first two from basic plan
- 3x3 and 4x4 are no more precise than RCBD because error df is too small

Triple Lattices

- Three replications use first three from basic plan
- Possible for all squares from 3x3 to 13x13

Quadruple Lattices

- Four replications use first four from basic plan
- Do not exist for 6x6 and 10x10
 a can repeat simple lattice, but analysis is different

Linear Model for Lattice Design

Example is for a single blocking factor



ANOVA

- Form of the analysis is the same for simple, triple and quadruple lattices
- Two error terms are computed
 - E_b Error for block = SSB/r(k-1)
 - E_e Experimental error = SSE/((k-1)(rk-k-1))

Source	df	SS	MS
Total	rk²-1	SSTot	
Replications	r-1	SSR	
Treatments (unadj)	k ² -1	SST	
Block in rep (adj)	r(k-1)	SSB	E _b
Intrablock error	(k-1)(rk-k-1)	SSE	E _e

Computing Sums of Squares

- SSTOT = $\sum Y_{ijl}^2 (G^2 / rk^2)$
 - G is the grand sum of all plots in the experiment
- SSR = $(1/k^2)\sum_{j} R_j^2 (G^2/rk^2)$
 - R_i is the sum of all plots in the jth replication
- $SSB = (1/kr(r-1))\sum_{ji}C_{ji}^2 (1/k^2r(r-1))\sum_{ji}C_{ji}^2$ (adjusted)
 - C_{jl} = sum over all replications of all treatments in the lth block of the jth replication minus rB_{jl}
 - B_{il} = sum of the k plots in the lth block of the jth replication
 - C_j = sum of all C_{jl} in the jth replication
- SST = $(1/r) \sum T_i^2 (G^2 / rk^2)$ (unadjusted)
 - T_i = sum of the ith treatment across replications
- SSE = SSTot SSR SSB SST

Adjustment factor

- Compare E_b with E_e : If $E_b \le E_e$
 - Adjustment of treatment means will have no effect
 - Analyze as if it were an RCBD using replications as blocks
- If E_b > E_e then compute an adjustment (weighting) factor A

•
$$A = (E_b - E_e)/(k(r - 1)E_b)$$

used to compute adjusted treatment means

$$\overline{Y}_{i(adj)} = \left(T_i + \sum AC_{jl}\right) / r$$

For all blocks in which the ith treatment occurs

Testing Treatment Differences

 To test significance among adjusted treatment means, compute an adjusted mean square

•
$$SSB_{u} = (1/k) \sum B_{jl}^{2} - (G^{2}/rk^{2}) - SSR$$

• $SST_{adj} = SST-Ak(r-1)[((rSSB_u)/(r-1)(1+kA)) - SSB]$

• MST_{adj} = SST_{adj} /(
$$k^2$$
-1)

 Finally, compute the F statistic for testing the differences among the adjusted treatment means

Standard Errors

Compute the effective error mean square

•
$$E'_{e} = \left[1 + \left((rkA) / (k+1)\right)\right]E_{e}$$

- SE of adjusted treatment mean
 - $=\sqrt{\mathsf{E}_{e}^{'}/\mathsf{r}}$
- SE of difference between adjusted means in same block

•
$$= \sqrt{(2E_{e} / r)(1 + (r - 1)A)}$$

SE of difference between adjusted means in different blocks

• =
$$\sqrt{(2E_e/r)(1+rA)}$$

For larger lattices (k > 4) it is sufficient to use

•
$$=\sqrt{2E'_e/r}$$

Relative Efficiency

Estimate the error mean square of an RCBD

•
$$E_{RCBD} = (SSB+SSE)/((k^2-1)(r-1))$$

Then the relative efficiency of the lattice is

• RE = E_{RCBD}/E_{e}'



Numerical Example - Simple Lattice

Rep I

Block		Barley	Yield kg	B _{jl}	C _{jl}	Adj		
1	(19)	(16)	(18)	(17)	(20)			
	18.2	13.0	9.5	6.7	10.1	57.5	17.6	1.54
2	(12)	(13)	(15)	(14)	(11)			
	13.3	11.4	14.2	11.9	13.4	64.2	4.2	0.37
3	(1)	(2)	(3)	(4)	(5)			
	15.0	12.4	17.3	20.5	13.0	78.2	5.3	0.46
4	(22)	(24)	(21)	(25)	(23)			
	7.0	5.9	14.1	19.2	7.8	54.0	-0.5	-0.04
5	(9)	(7)	(10)	(8)	(6)			
	11.9	15.2	17.2	16.3	16.0	76.6	3.9	0.34
Selection number in ()					Sum	330.5	30.5	2.66

Example from Petersen, R.G. 1994. Agricultural Field Experiments

Numerical Example - Simple Lattice

Rep 2

Block		Barley	B _{jl}	C _{jl}	Adj			
1	(23)	(18)	(3)	(8)	(13)			
	7.7	15.2	19.1	15.5	14.7	72.2	-9.9	-0.86
2	(5)	(20)	(10)	(15)	(25)			
	15.8	18.0	18.8	14.4	20.0	87.0	-13.3	-1.17
3	(22)	(12)	(2)	(17)	(7)			
	10.2	11.5	17.0	11.0	15.3	65.0	-10.4	-0.91
4	(14)	(24)	(9)	(4)	(19)			
	10.9	4.7	10.9	16.6	9.8	52.9	15.5	1.35
5	(6)	(16)	(11)	(21)	(1)			
	20.0	21.1	16.9	10.9	15.0	83.9	-12.4	-1.08
Select	Selection number in () Sum						-30.5	-2.67
				Gran	nd Sum	691.5	0.0	0.0

Initial ANOVA

Source	df	SS	MS
Total	49	805.42	
Replication	1	18.60	
Selection (unadj)	24	621.82	
Block in rep (adj)	8	77.59	9.70=E _b
Intrablock error	16	87.41	5.46=E _e

 $E_b > E_e$ so we compute the adjustment factor, A

 $A = (E_{b} - E_{e})/(k(r-1)E_{b}) = (9.70 - 5.46)/((5)(1)(9.70)) = 0.0874$

Adjustment of mean for Entry 1

 $Y_{1(adj)} = \left(T_1 + \sum AC_{jl}\right) / r = \left[30 + (0.0874 * 5.3) + (0.0874 * 1.35)\right] / 2 = 14.69$

Intrablock ANOVA of Adjusted Means

Source	df	SS	MS	F
Total	49	805.42		
Replication	1	18.60		
Selection (adj)	24	502.39	20.93	3.83**
Block in rep (unadj)	8	252.18		
Intrablock error	16	87.41	5.46=E _e	

It may be better to use the effective error in the denominator of the F test

• $E_e' = (1+(rkA)/(k+1))E_e = (1+(2*5*0.0874)/6)*5.46 = 6.26$

• F = 20.93/6.26 = 3.34**

See the supplemental Excel spreadsheet for more details

Relative Efficiency

- How does the precision of the Lattice compare to that of a randomized complete block design?
 - First compute MSE for the RCBD as:
 E_{RCBD} = (SSB+SSE)/(k² 1)(r -1) =
 - (77.59 + 87.41)/(24)(1) = 6.88
- Then % relative efficiency =
 - $(E_{RCBD} / E_{e'})100 = (6.88/6.26)*100 = 110.0\%$
 - There is a 10% gain in efficiency from using the lattice

Cyclic Designs

Incomplete Block Designs discussed so far

- Require extensive tables of design plans
- Need to avoid mistakes when assigning treatments to experimental units and during field operations

Cyclic designs are a type of incomplete block design

Relatively easy to construct and implement

 Generated from an initial block 	Block	Treatment Label
	1	0, 1, 3
• Example: 6 treatments, block size = 3	2	1, 2, 4
 Add one to each treatment label 	3	2, 3, 5
for each additional block	4	3, 4, 0
 Modulo t=6 	5	4. 5. 1

6

5, 0, 2

Good reference: Kuehl, 2000, Chapt. 10

Alpha designs

- Patterson and Williams (1976) described a new way to construct cyclic, resolvable incomplete block designs
- α-designs are available for many (*r*,*k*,*s*) combinations
 - *r* is the number of complete replicates
 - *k* is the block size
 - s is the number of blocks per replicate
 - Number of treatments t = sk
- Efficient α-designs exist for some combinations for which conventional lattices do not exist
- Can accommodate unequal block sizes
- Two types: α(0, 1) and α(0, 1, 2)
 - Indicates values of λ that occur in the trial (2 or 3 associate classes)

Alpha Designs - Software

- Gendex <u>http://designcomputing.net/gendex/alpha/</u>
 - Can generate optimal or near-optimal α -designs
 - Up to 10,000 entries
 - Evaluation/academic copy is free and can be downloaded
 - Cost for commercial perpetual license is \$299
- CycDesigN
- Agrobase
- R agricolae package
 - design.alpha, design.lattice, design.cyclic, design.bib
- SAS PROC PLAN
 - some code required
 - http://www.stat.ncsu.edu/people/dickey/courses/st711/Demos/

Gend	0		
Alpha	RCD	MIGA	RAT
IBD	RRCD	SOD	NOA
CIBD	FEADO	СUТ	LHD

Efficiency Factors for Lattice Designs

Balanced lattice
$$E = \frac{(k+1)(r-1)}{r(k+2)-(k+1)}$$

$$\mathsf{E} = \frac{\mathsf{k} + \mathsf{1}}{\mathsf{k} + \mathsf{3}}$$

Triple lattice

$$\mathsf{E} = \frac{2\mathsf{k} + 2}{2\mathsf{k} + 5}$$

Use as large a block size as possible while maintaining homogeneity of plots within blocks

 Alpha lattice (upper bound)

$$\mathsf{E} = \frac{(t-1)(r-1)}{(t-1)(r-1) + r(s-1)}$$

t = # treatments, k = block size, s = # blocks/rep, r = # complete reps

Analysis of Lattice Experiments

SAS

- PROC MIXED, PROC LATTICE
- PROC VARCOMP (random genotypes)
- PROC GLIMMIX (non-normal)
- ASREML
- GENSTAT
- R
 - stats, Ime4 packages
- Agrobase
 - www.agronomix.com/



Meadowfoam (Limnanthes alba)

- Native to vernal pools in the PNW
- First produced as a crop in 1980
- Seed oil with novel long-chain fatty acids
 - light-colored and odor free
 - exceptional oxidative stability
- Used in personal care products
- Potential uses
 - fuel additive
 - vehicle lubricants
 - pharmaceutical products





GJ Pool Progeny Trial, 2012 (α -Lattice)



Plot size 4 ft x 12 ft (planted 2000 seeds/plot)

Average seed yield 1612 kg/ha

Autofertile Progeny Trial, 2013

- 132 entries
- 127 TC + 5 checks
- 11 x 12 α -lattice
- 2 replications
- 1 location
- 4 ft x 10 ft plots
- 1450 seeds/plot



Average seed yield 879 kg/ha

Meadowfoam Reproduction

- Winter annual, diploid
- Factors that promote outcrossing:
 - Protandrous
 - Heterostylous
 - Many native pollinators
 - Inbreeding depression in L. alba ssp alba
- Potential for selfing
 - Perfect flowers
 - Self-compatible
 - Flowers close at night







Figure 2.—Sequential reproductive development of a Mermaid meadowfoam flower: A. Flower opening with initial pollen availability (dehiscence) and unreceptive stigmas B. Maximum pollen shed with unreceptive stigmas C. Maximum stigmatal receptivity with reduced pollen availability

Autofertile Pool

Derived from crosses between

- Outcrossing L. alba ssp. alba populations
- Self-pollinating lines from L. alba ssp. versicolor
- Inbreeding depression??
- Development of progeny for evaluation
 - Selfed plants in the greenhouse for two generations, without pollinators
 - Selected for autofertility (seed number)
 - Planted S₂ families in an isolated nursery with honeybees to produce testcross progeny

S₂ Families in Isolated Nursery

- Planted 5
 seeds per
 family in
 short rows
- 2 replicates
- Randomized blocks
- Honeybee pollinators



ALPHA 7.0: Construct alpha designs of size (r,k,s) (c) 2013 Design Computing (designcomputing.net/)

Best alpha design for v=132, r=2, k=11, and s=12.

Efficiency 0.8542

Plan (blo	ocks are ro	ows):								
82	83	97	117	48	54	25	2	99	38	107
100	74	86	65	85	94	71	18	126	28	31
84	1	53	120	101	72	104	122	114	66	68
10	112	7	15	33	121	127	59	14	67	130
131	124	118	102	3	39	93	43	109	55	106
103	44	37	32	62	63	108	23	80	70	60
89	27	79	24	4	64	81	29	21	61	128
69	47	41	49	19	98	110	76	75	30	36
105	22	40	9	90	5	91	16	132	95	111
26	113	35	45	12	20	73	11	6	119	34
125	116	13	123	92	51	52	56	88	46	129
57	42	17	78	50	58	8	77	115	96	87
131	59	86	36	114	25	16	119	103	27	88
90	2	23	42	116	65	34	128	67	43	68
93	58	15	64	70	101	82	92	111	100	19
18	5	102	72	73	61	10	49	87	62	99
115	26	1	83	31	30	3	52	40	37	112
85	6	97	121	125	78	66	24	9	47	106
17	109	107	56	60	130	45	84	126	41	29
81	118	132	46	108	28	98	96	38	35	7
51	12	4	76	33	104	44	105	94	39	57
123	54	8	55	32	110	22	71	20	120	21
122	79	69	48	77	129	95	124	11	80	14
75	50	91	113	63	13	117	89	74	53	127

Field Map for the Autofertile Trial (Rep 1)

N

BORDER $1131\ 1130\ 1129\ 1128\ 1127\ 1126\ 1125\ 1124\ 1123\ 1122\ 1121\ 1120\ 1119\ 1118\ 1117\ 1116\ 1115\ 1114\ 1113\ 1112\ 1111$ 10 ft 57 129 92 123 13 116 125 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 16 | 132 | 95 | 111 | 26 | 113 | 35 1088 1087 1086 1085 1084 1083 1082 1081 1080 1079 1078 1077 1076 1075 1074 1073 1072 1071 1070 1069 1068 1067 **128 61** 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 131 124 118 102 106 103 63 108 1044 1043 1042 1041 1040 1039 1038 1037 1036 1035 1034 1033 1032 1031 1030 1029 1028 1027 1026 1025 1024 1023 10 68 66 114 122 104 72 101 120 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022

BORDER

107 100

18 126

Block 1

Data Set – Variables Included

- Plant height cm
- 1000-seed weight (TSW)
- Seed oil content (%)
- Seed yield kg/ha



Other options for data input Import Wizard Infile statements SAS libraries

SAS Data Input

ods html close; ods html;

data AF2013;

Input PLOT REP BLOCK ENTRY NAME\$ Height TSW Oil Yield; Oilyld=yield*oil/100;

datalines;

1001	1	1	82	179-53-2	22.5	9.48	26.16	<mark>782</mark>
1002	1	1	83	179-56-1	22.5	9.38	27.42	<mark>832</mark>
1003	1	1	97	179-83-1	23.5	8.56	24.25	<mark>944</mark>
1004	1	1	117	188-14-1	19.5	7.90	28.11	721
								•
2129	2	24	89	179-63-1	24.0	10.10	26.68	<mark>858</mark>
2130	2	24	74	179-38-1	22.5	9.35	28.40	<mark>631</mark>
2131	2	24	53	179-4-2	20.5	9.51	25.76	1127
2132	2	24	127	188-133-2	22.5	9.62	25.37	721

;

Mixed Model Analysis of Yield

PROC MIXED;

TITLE 'Lattice analysis of yield: PROC MIXED, entries fixed';

CLASS REP BLOCK ENTRY;

MODEL Yield = ENTRY;

RANDOM REP BLOCK (REP) / solution;

ods output solutionr=syield;

```
/*test for entries that differ from Ross*/
```

```
LSMEANS entry/pdiff=CONTROL('128');
```

```
ods output lsmeans=Yieldadj diffs=Ylddiff;
```

RUN;

Use export wizard to export syield Yieldadj Ylddiff

Analysis of Yield (fixed entries)

The Mixed Procedure Covariance Parameter Estimates

Cov Parm	Estimate
REP	2435.58
BLOCK (REP)	9900.50
Residual	12657

Fit Statistics

-2 Res Log Likelihood	1753.2
AIC (smaller is better)	1759.2
AICC (smaller is better)	1759.3
BIC (smaller is better)	1755.2

Type 3 Tests of Fixed Effects

	Num	Den		
Effect	DF	DF	F Value	Pr > F
Entry	131	109	6.17	<.0001

Analysis of Yield – Random Effects

Cov Parm	Height	TSW	Oil	Yield
REP	0.191	0.0369	0.4252	2435.6
BLOCK(REP)	1.568	0	0.1696	9900.5
Residual	2.238	0.1972	1.0360	12657.0



Conclusion

- α-Lattice Designs?
- Yes!
- You have nothing to lose
- They are most helpful when you need them the most



Lattice Design References

- Cochran, W.G., and G.M. Cox. 1957. Experimental Designs, 2nd edition. Wiley, New York.
- Hinkelman, K., and O. Kempthorne. 2006. Design and Analysis of Experiments.
 Volume 2. Wiley, New York.
- John, J.A., and E.R. Williams. 1995, Cyclic and Computer Generated Designs, 2nd edition. Chapman and Hall, London, UK.
- Kuehl, R.O. 2000. Chapt. 10 in Design of Experiments: Statistical Principles of Research Design and Analysis, 2nd edition. Duxbury Press.
- Patterson, H.D., and E.R. Williams. 1976. A new class of resolvable incomplete block designs. Biometrika 63: 83–92.
- Piepho, H.P., A. Büchse, and B. Truberg. 2006. On the use of multiple lattice designs and α-designs in plant breeding trials. Plant Breeding 125: 523–528.
- Yau, S.K. 1997. Efficiency of alpha-lattice designs in international variety yield trials of barley and wheat. Journal of Agricultural Science, Cambridge 128: 5–9.

Acknowledgements

Funding

- OMG Meadowfoam Oil Seed Growers Cooperative
- Paul C. Berger Professorship Endowment
- Crop and Soil Science Department, OSU

Meadowfoam Staff

- Gary Sandstrom
- Ann Corey
- Student workers
- Lattice Design Examples
 - Matthais Frisch
 - Roger Petersen
 - Nan Scott

