Welcome to the Introduction to Augmented Design Webinar

Today's Presenter: Dr. Jennifer Kling

Presentation & Supplemental Files
www.extension.org/pages/60430

Brought to you by: Plant Breeding and Genomics
www.eXtension.org/plant_breeding_genomics

Host: Heather Merk

Sign up for PBG News: http://pbgworks.org

Please fill out the survey evaluation! (You will be contacted via email)

Watch past webinars!

Sign up to learn how to use double haploids for winter wheat improvement!

Thursday November 17, 1pm Eastern Standard Time

www.extension.org/pages/60426

Introduction to Augmented Designs

Applications in Plant Breeding

Jennifer Kling

Oregon State University

Outline – Augmented Designs

- Essential features
  - When are they used in plant breeding?
- Design options
  - Today - one-way control of heterogeneity
- Augmented Block Design - Example
  - Randomization and Field Plan
  - Analysis with SAS
  - Interpretation of Results
- Overview of variations on the basic design
- Software and Further References

Augmented Designs - Essential Features

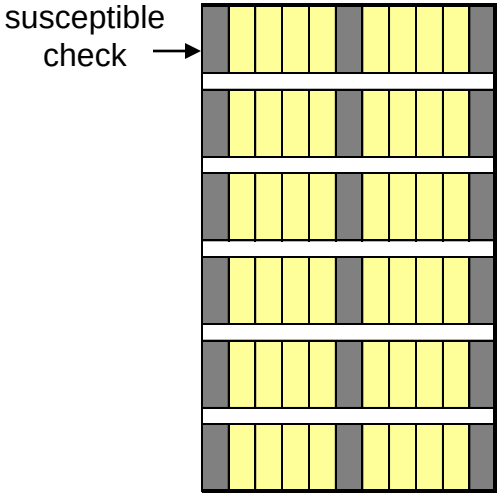
- Introduced by Federer (1956)
- Controls (check varieties) are replicated in a standard experimental design
- New treatments (genotypes) are not replicated, or have fewer replicates than the checks – they augment the standard design

Before
Augmented Designs

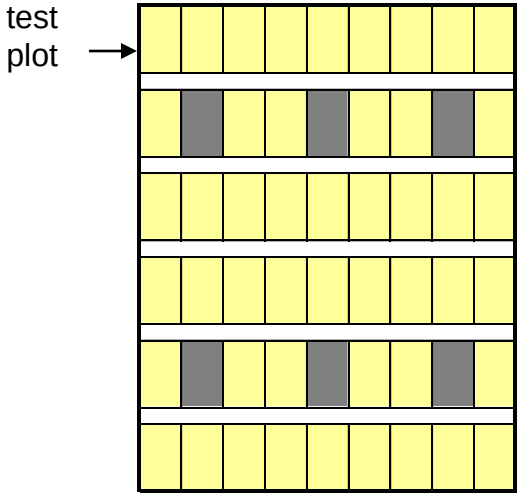
Breeding for Resistance to *Striga* in Sorghum at ICRISAT

Systematic use of checks

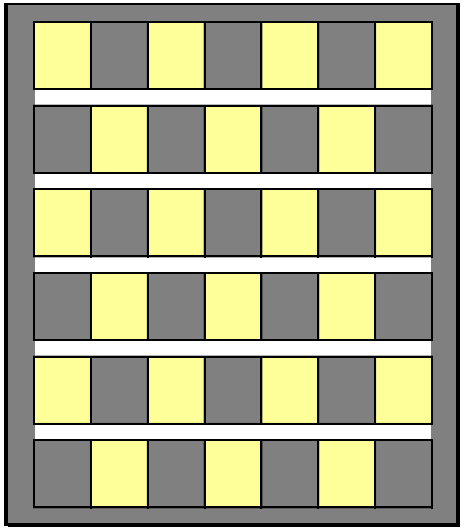
Stage I
Observation Nursery
(unreplicated)



Stage II
Preliminary Screening
(replicated)



Stage III
Advanced Screening
(replicated)



Checkerboard

Augmented Designs - Advantages

- Fewer check plots required than for designs with systematic repetition of a single check
- Provide an estimate of standard error that can be used for comparisons
  - among the new genotypes
  - between new genotypes and check varieties
- Observations on new genotypes can be adjusted for field heterogeneity (blocking)

Some Disadvantages

- Considerable resources are spent on production and processing of control plots
- Relatively few degrees of freedom for experimental error, which reduces the power to detect differences among treatments
- Unreplicated experiments are inherently imprecise, no matter how sophisticated the design

When are they used?

- Early generations
  - seed is limited
  - too many new entries to replicate
  - difficult to maintain homogeneity within blocks with so many entries
- On-farm research
  - growers may prefer to grow a single replication
  - may not be able to accommodate all entries

Design Options

- Choose a design that is appropriate for controlling the heterogeneity in the experimental area
  - One-way blocking
    - Randomized Complete Block Design
    - Incomplete Block Designs (e.g. Lattice Design)
  - Two-way blocking
    - Latin Square (Complete Blocks)
    - Youden Square (Incomplete Blocks)
    - Row-Column Designs

Design Options

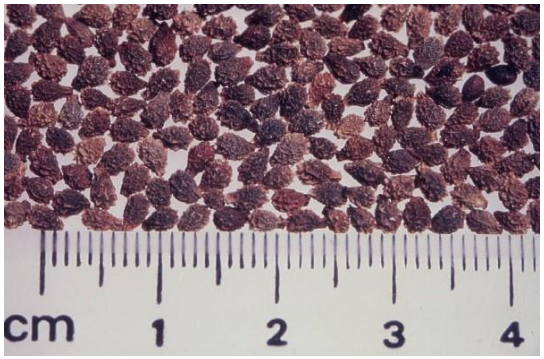
- Underlying design refers to assignment of checks to the experimental units
- All augmented designs are incomplete with respect to the new entries
- Can be replicated in different environments
  - Need to consider relative efficiency compared to other possible designs

Augmented Block Design Example

- Control heterogeneity in one direction
- Simplest case:
  - Checks occur once in every block
  - New entries occur once in the experiment
- This is an ARCBD

Meadowfoam (*Limnanthes alba*)

- Native plant 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

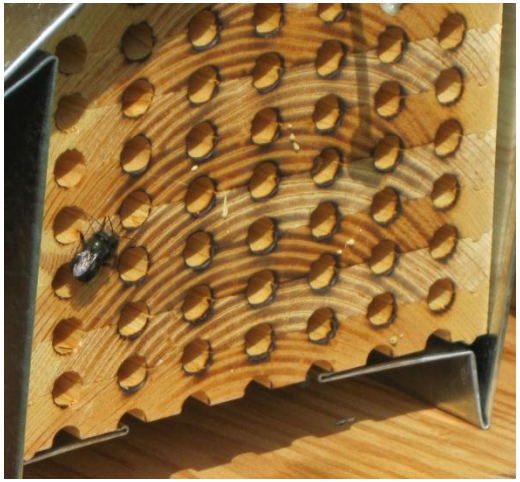


Meadowfoam in Oregon

- Good rotation crop in the Willamette Valley
  - Winter annual
  - Plant and seed meal are high in glucosinolates
  - Same equipment as grass seed

Pollinators

Honeybees



Blue Orchard Bees

Blue Bottle Flies

The Germplasm

- Diverse breeding populations were inherited from a retired breeder (Gary Jolliff)
- Populations were regenerated in the greenhouse and selfed
- S_2 lines were transplanted to the field and allowed to outcross
- Insufficient seed for replicated progeny trials
- Goal - form a broadbased pool for recurrent selection
  - Screen S_2-testcrosses
  - Recombine selected S_2 parents

The Experiment (2007-2008)

- New treatments:
  - 50 S_2-testcross families
- Check varieties:
  - Ross (C1)
    - Cycle 4 of an elite population (OMF58)
    - Widely grown commercial variety
  - OMF183 (C2)
    - Cycle 5 from OMF58
  - Starlight (C3)
    - released variety derived from this germplasm collection

The Design

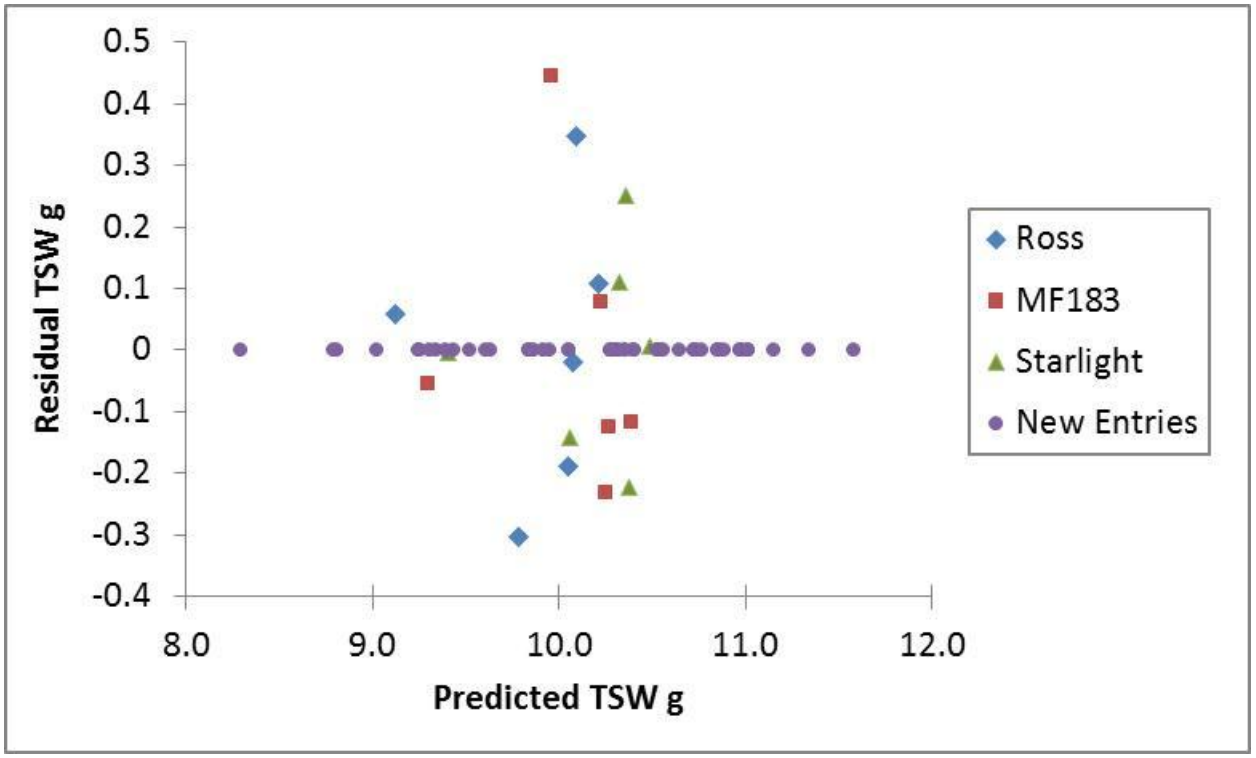
- Block in one direction
- 3 checks (c=3)
- 6 blocks (r=6)
- Error degrees of freedom = (r-1)(c-1) = 10

- 50 new entries (n=50)
- Number of plots = n + r*c = 50 + 6*3 = 68
- Plots per block = 68/6 = 11.3 → 12 (c=3, n=9)
- Last block has only 8 plots (c=3, n=5)
  - Unequal numbers of new entries per block is OK

Statistical Model

$$Y_{ij} = \mu + \beta_i + c_j + \tau_{k(i)} + \varepsilon_{ij}$$

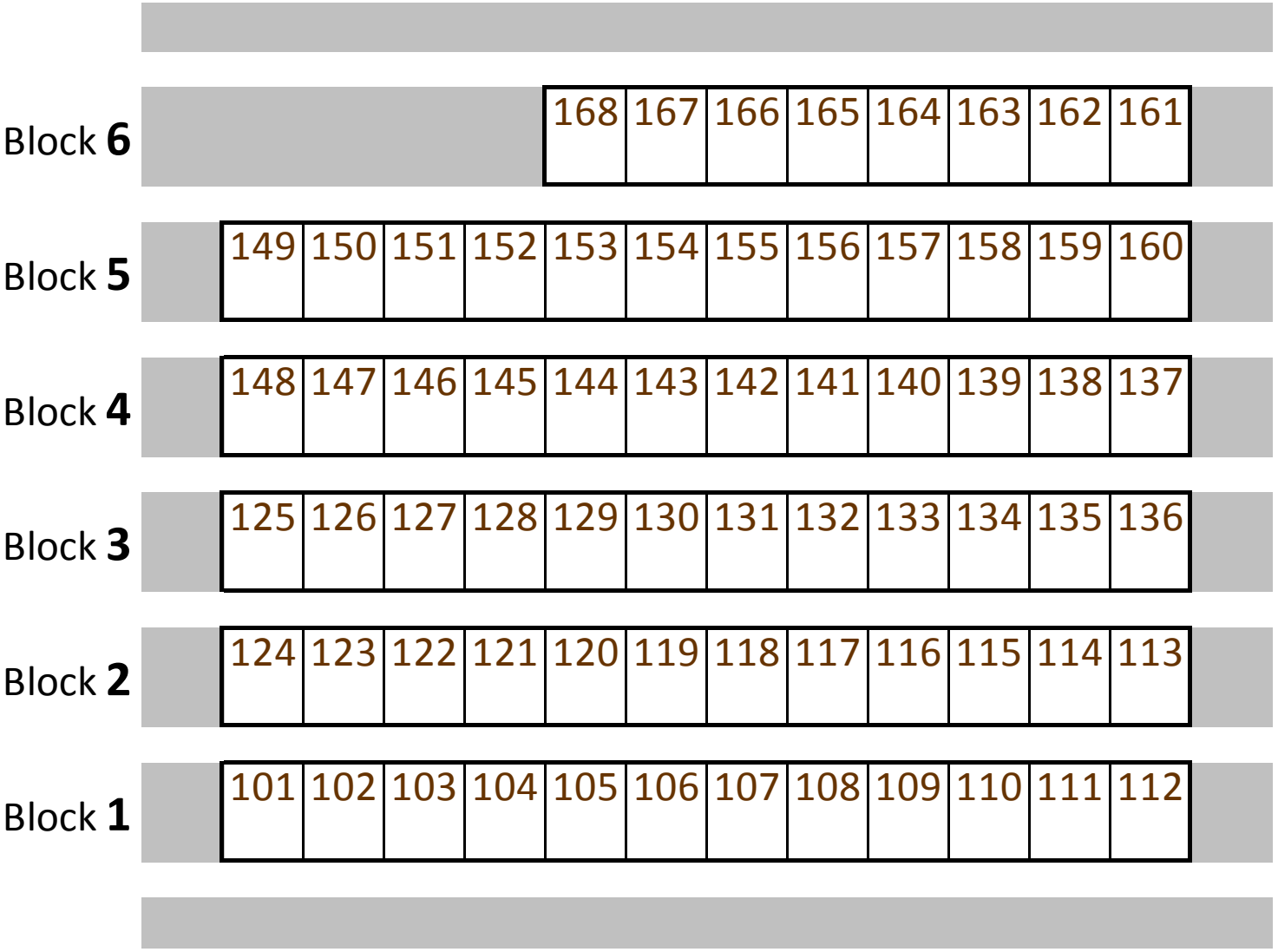
mean + blocks + checks + new entries + error



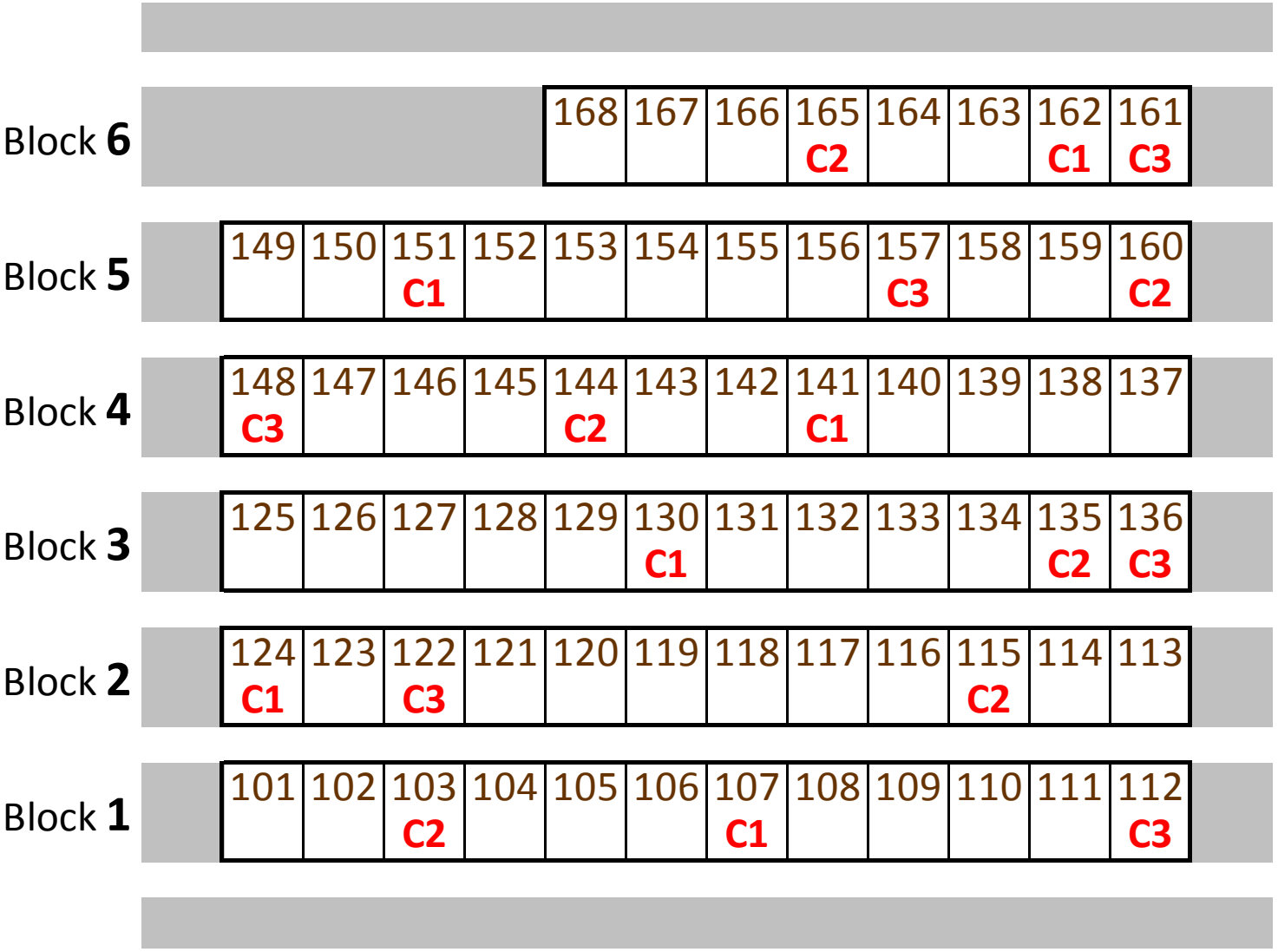
Field Plan

- Include sufficient number of checks and replicates to provide a good estimate of experimental error and adequate power for detecting differences among varieties
- Arrange blocks along field gradient to maximize variation among blocks and minimize variation within blocks
- Assign each of the checks at random to each of the blocks

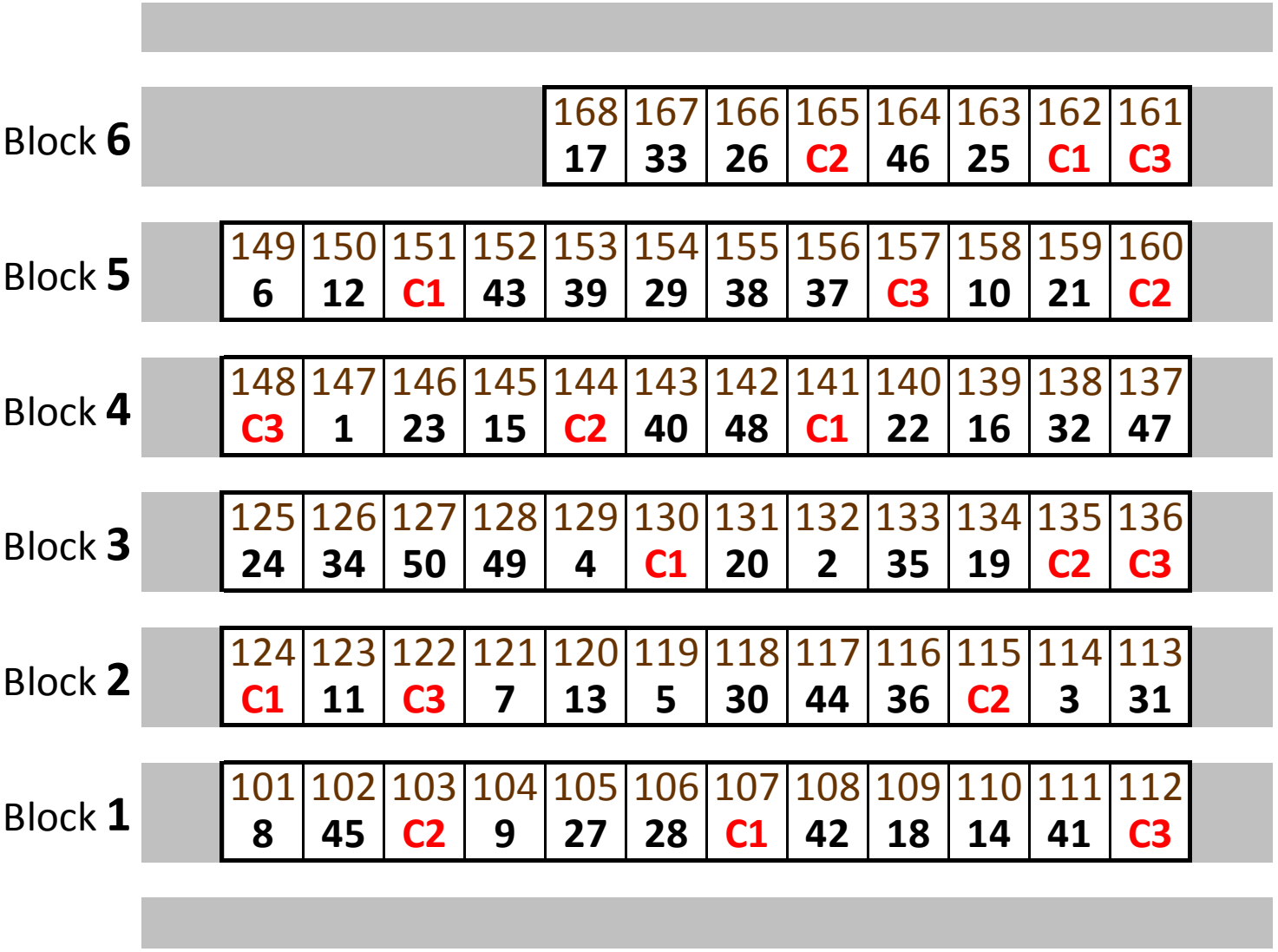
Designation of plots in blocks



Checks assigned to plots in each block



New entries included in final plan



Meadowfoam progeny trials

Data Collection

- Flowering dates, plant height, disease resistance, lodging, seed yield, 1000-seed weight, (oil content)
- For this example: 1000-seed weight (TSW)
  - Relatively high heritability
  - Positively correlated with oil content and oil yield

Fixed or Random?

- Generally assume that blocks are random
  - They represent a larger group of potential blocks
  - We want to make inferences beyond the particular sample of blocks in the experiment
- Genotypes
  - Checks are fixed effects – we are interested in making comparisons with these specific varieties
  - New entries could be fixed or random
    - May be fixed in this case – want to select the winners
    - Most commonly random – particularly in genetic studies

SAS data input – genotypes fixed

```
DATA anyname;
INPUT Plot Entry Name$ Block TSW;
datalines;
   101      8          31        1      9.25
   102     45         126        1      9.43
   103     90       MF183        1     10.30
   104      9          34        1      9.02
   105     27          89        1     10.97
   106     28          90        1     10.52
   107     89        Ross        1      9.86
   108     42         121        1     10.56
   109     18          68        1     10.39
   110     14          57        1     10.97
   111     41         118        1     10.32
   112     91   Starlight        1     10.44
     .      .           .        .         .
     .      .           .        .         .
   161     91   Starlight        6     10.15
   162     89        Ross        6     10.44
   163     25          85        6     10.85
   164     46         132        6      9.52
   165     90       MF183        6     10.14
   166     26          87        6     10.73
   167     33         101        6      9.24
   168     17          66        6      8.79
;
```

Other options for data input

- Import Wizard
- Infile statements
- SAS libraries

Analysis #1 – new entries fixed

```
PROC MIXED;
TITLE 'Augmented Design using PROC MIXED – entries fixed';
CLASS block entry;
MODEL TSW=entry;
RANDOM block;
/*compare all means using Tukey's test*/
LSMEANS entry/pdiff adjust=tukey;
/*test for entries that exceed Starlight*/
LSMEANS entry/pdiff=CONTROLU('91') adjust=dunnett;
ods output lsmeans=TSWadj diffs=TSWpdiff;
RUN;
```

Use export wizard to export

- ***TSWadj***
- ***TSWpdiff***

Results for Analysis #1 (fixed entries)

The Mixed Procedure

Covariance Parameter Estimates

Cov Parm	Estimate
Block	0.1381
Residual	0.06981

Fit Statistics

-2 Res Log Likelihood	17.7
AIC (smaller is better)	21.7
AICC (smaller is better)	22.7
BIC (smaller is better)	21.3

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Entry	52	10	7.51	0.0008

Output from Dunnett Test

TSWpdiff.xls

Effect	Entry	_Entry	Estimate	StdErr	DF	tValue	Probt	Adjustment	Adjp
Entry	1	91	000.3879	000.3131	10	1.24	0.1219	Dunnett-Hsu	0.9351
Entry	2	91	-000.5862	000.3131	10	-1.87	0.9547	Dunnett-Hsu	1.0000
Entry	3	91	000.4366	000.3131	10	1.39	0.0967	Dunnett-Hsu	0.8879
Entry	4	91	-000.1362	000.3131	10	-0.44	0.6636	Dunnett-Hsu	1.0000
Entry	5	91	000.5366	000.3131	10	1.71	0.0587	Dunnett-Hsu	0.7459
Entry	6	91	-000.4782	000.3131	10	-1.53	0.9212	Dunnett-Hsu	1.0000
Entry	7	91	-000.7034	000.3131	10	-2.25	0.9758	Dunnett-Hsu	1.0000
Entry	8	91	-001.0565	000.3131	10	-3.37	0.9965	Dunnett-Hsu	1.0000
Entry	9	91	-001.2865	000.3131	10	-4.11	0.9989	Dunnett-Hsu	1.0000
Entry	10	91	-000.0282	000.3131	10	-0.09	0.5350	Dunnett-Hsu	1.0000
Entry	11	91	001.4666	000.3131	10	4.68	0.0004	Dunnett-Hsu	0.0149
Entry	12	91	000.4518	000.3131	10	1.44	0.0898	Dunnett-Hsu	0.8700
Entry	13	91	000.3266	000.3131	10	1.04	0.1607	Dunnett-Hsu	0.9726
Entry	14	91	000.6635	000.3131	10	2.12	0.0300	Dunnett-Hsu	0.5249
Entry	15	91	-000.4921	000.3131	10	-1.57	0.9265	Dunnett-Hsu	1.0000

Analysis #2 – new entries random

```
DATA arcbd;
set anyname;
if(entry>50) then new=0;
else new=1;
if(new) then entryc=999;
else entryc=entry;
RUN;
```

Plot	Entry	Name	Block	TSW	new	entryc
101	8	31	1	9.25	1	999
102	45	126	1	9.43	1	999
103	90	MF183	1	10.3	0	90
104	9	34	1	9.02	1	999
105	27	89	1	10.97	1	999
106	28	90	1	10.52	1	999
107	89	Ross	1	9.86	0	89
108	42	121	1	10.56	1	999
109	18	68	1	10.39	1	999
110	14	57	1	10.97	1	999
111	41	118	1	10.32	1	999
112	91	Starlight	1	10.44	0	91
113	31	96	2	11.58	1	999
114	3	15	2	9.95	1	999
115	90	MF183	2	9.24	0	90

Wolfinger, R.D., W.T. Federer, and O. Cordero-Brana, 1997

Analysis #2 – new entries random

```
PROC MIXED;
TITLE 'Augmented Design using PROC MIXED – entries random';
CLASS block entry entryc;
MODEL TSW=entryc;
RANDOM block entry*new/solution;
LSMEANS entryc;
ods output solutionr=eblups;
RUN;
```

Use export wizard to export

- ***eblups***

Results for Analysis #2 (random entries)

The Mixed Procedure

Covariance Parameter Estimates

Cov Parm	Estimate
Block	0.01418
new*Entry	0.3659
Residual	0.1748

Fit Statistics

-2 Res Log Likelihood	136.5
AIC (smaller is better)	142.5
AICC (smaller is better)	142.9
BIC (smaller is better)	141.9

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
entryc	3	10	0.89	0.4779

Estimated Best Linear Unbiased Predictors

eblups.xls

Effect	Block	Entry	Estimate	StdErrPred	DF	tValue	Probt
Block	1		000.0189	000.1018	10	0.19	0.8564
Block	2		-000.1264	000.1018	10	-1.24	0.2427
Block	3		000.0188	000.1018	10	0.18	0.8569
Block	4		000.0733	000.1018	10	0.72	0.4882
Block	5		000.0156	000.1018	10	0.15	0.8811
Block	6		-000.0002	000.1044	10	0.00	0.9983
new*Entry		**1**	000.2991	000.3561	10	0.84	0.4206
new*Entry		**2**	-000.2461	000.3561	10	-0.69	0.5052
new*Entry		**3**	-000.0869	000.3561	10	-0.24	0.8122
new*Entry		**4**	000.0585	000.3561	10	0.16	0.8729
new*Entry		**5**	-000.0192	000.3561	10	-0.05	0.9581
new*Entry		**6**	-000.4199	000.3561	10	-1.18	0.2656
new*Entry		**7**	-000.8584	000.3561	10	-2.41	0.0366
new*Entry		**8**	-000.6590	000.3561	10	-1.85	0.0939
new*Entry		**9**	-000.8146	000.3561	10	-2.29	0.0452
new*Entry		**10**	-000.1153	000.3561	10	-0.32	0.7527
new*Entry		**11**	000.6102	000.3561	10	1.71	0.1173
new*Entry		**12**	000.2095	000.3561	10	0.59	0.5693
new*Entry		**13**	-000.1613	000.3561	10	-0.45	0.6602
new*Entry		**14**	000.5051	000.3561	10	1.42	0.1865
new*Entry		**15**	-000.2965	000.3561	10	-0.83	0.4245
new*Entry		**16**	000.0487	000.3561	10	0.14	0.8940

Variations – two-way control of heterogeneity

Design	Features/Applications	Reference
Modified Youden Square	Two-way control of heterogeneity	Federer and Raghavarao, 1975
Augmented Row-column	Entries adjusted for adjacent checks; requires many checks	Federer, Nair and Raghavarao, 1975
Modified Augmented (MAD Type 1)	Systematic placement of controls; Requires square plots	Lin and Poushinsky, 1983
Modified Augmented (MAD Type 2)	Systematic placement of controls; Long, rectangular plots	Lin and Poushinsky, 1985

More Variations

Design	Features/Applications	Reference
Factorial treatments		
Augmented Split Block	Intercropping, stress adaptation	Federer, 2005
Incomplete Blocks		
Augmented Lattice Square	Accommodates larger number of checks	Federer, 2002
α-α Lattice	Accommodates larger number of checks	Williams and John, 2003
Augmented p-rep (partially replicated)	Replicates of entries used to make block adjustments and estimate error	Williams, Piepho, and Whitaker, 2011
Multiple locations		
RCBD, ICBD, row-column, etc.	SAS analysis provided	Federer, Reynolds, and Crossa, 2001

Multiple Locations – Augmented or Lattice Design?

- Lattice Designs
  - Sites are complete replications
  - Information from all plots are used to adjust means for field effects
    - May have greater precision and power to detect differences among entries
- Augmented Designs
  - Flexible arrangement of new entries in field plans
  - Estimate of experimental error obtained from each location
    - Assess site quality
    - Evaluate GXE interactions
    - More flexibility in combining information across sites

Software for Augmented Designs

- AGROBASE GenII by Agronomix, Inc.
  - Randomize and analyze modified augmented type-2 designs
- CIMMYT – sas macro called UNREPLICATE

  http://apps.cimmyt.org/english/wps/biometrics/
  - Developed in 2000 – uses some older SAS syntax
- Statistical Package for Augmented Designs (SPAD) – Indian Agricultural Research Institute, New Delhi

  www.iasri.res.in/iasri**software**/SPAD.ppt

  http://www.iasri.res.in/design/Augmented%20Designs/home.htm

References

- Full list of references provided as supplement

- Federer, W.T. 1961. Augmented designs with one-way elimination of heterogeneity. Biometrics 17(3): 447-473
- Wolfinger, R.D., W.T. Federer, and O. Cordero-Brana. 1997. SAS PROC GLM and PROC MIXED for recovering blocking and variety information in augmented designs. Agronomy Journal 89: 856-859

Acknowledgements

- USDA-CSREES special grant for meadowfoam research
- OMG Meadowfoam Oil Seed Growers Cooperative
- Paul C. Berger Professorship Endowment
- Crop and Soil Science Department, OSU
- Meadowfoam staff
  - Mary Slabaugh
  - Joanna Rosinska
  - Trisha King
  - Alicia Wilson

Questions?

Presentation & Supplemental Files

www.extension.org/pages/60430