Experiment Design Using ARM Software

Steven R. Gylling, Ph.D.
Gylling Data Management, Inc.

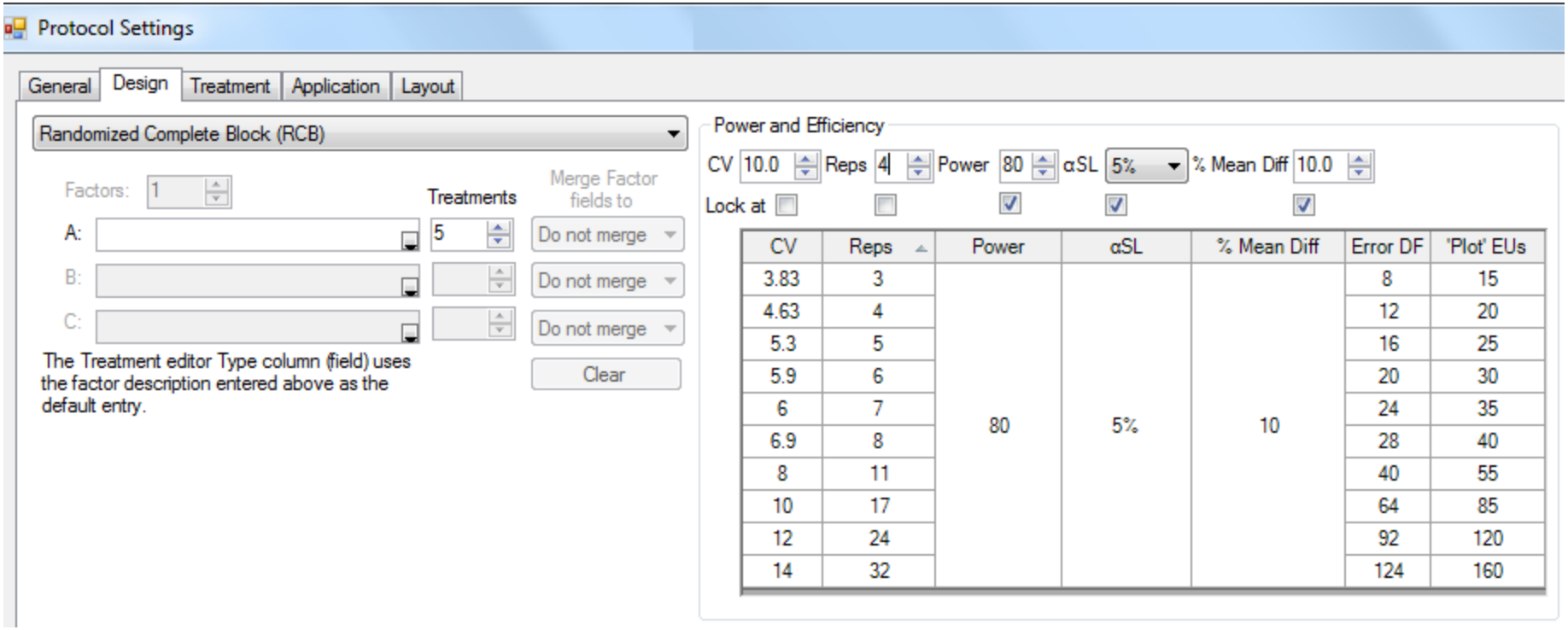
Plan Experiments to Have:

- A reasonable chance of distinguishing anticipated treatment differences
- The optimum number of replicates required to meet objectives
- An efficient experimental design and randomization for desired precision
- Cost-effective utilization of the available experimental area

Why is Planning Critical?

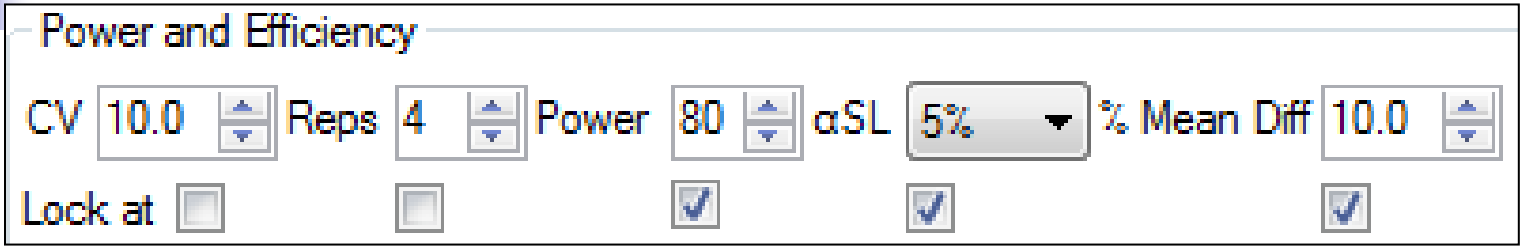
- Can reduce costs by selecting optimum number of replicates and samples
- Expected treatment differences are typically < 10%, and frequently < 5%, so small precision gains can help to:
  - Distinguish an actual treatment difference (reject null hypothesis H_0)
  - Strengthen evidence of no treatment diff.) (do not reject null hypothesis H_0)

ARM 2015
Power and Efficiency Planner



CV	Reps	Power	αSL	% Mean Diff	Error DF	'Plot' EUs
3.83	3	80	5%	10	8	15
4.63	4				12	20
5.3	5				16	25
5.9	6				20	30
6	7				24	35
6.9	8				28	40
8	11				40	55
10	17				64	85
12	24				92	120
14	32				124	160

Power and Efficiency Planner



- "Lock at" to keep 3-4 columns constant
- Calculates table of possible values for "unlocked" columns (e.g. Rep or CV)
- Values entered by protocol writer are carried into trials created from protocol, conveying protocol expectations to trialist

Power and Efficiency Planner

Plan replicates to achieve required precision
5 treatments with CV=5, 10% mean diff.

Power and Efficiency

CV 5.0 Reps 4 Power 80 αSL 5% % Mean Diff 10.0

Lock at ☑ ☐ ☑ ☐ ☑

CV	Reps	Power	αSL	% Mean Diff	Error DF
5	7	80	1%	10	24
	5		5%		16
	4		10%		12
	4		15%		
	3		20%		8
	3		25%		
	3		30%		
	2		40%		4
	2		50%		

Summary		
Reps	αSL=Significance	
7	1%	0.01
5	5%	0.05
4	10-15%	0.1-0.15
3	20-30%	0.2-0.3
2	40-50%	0.4-0.5

Power and Efficiency Planner

CV effect on minimum detectable % mean difference at 5% αSL for 10 treatments, 4 reps

Power and Efficiency

CV 5.0 Reps 4 Power 80 αSL 5% % Mean Diff 10.0

Lock at ☐ ☑ ☑ ☑ ☐

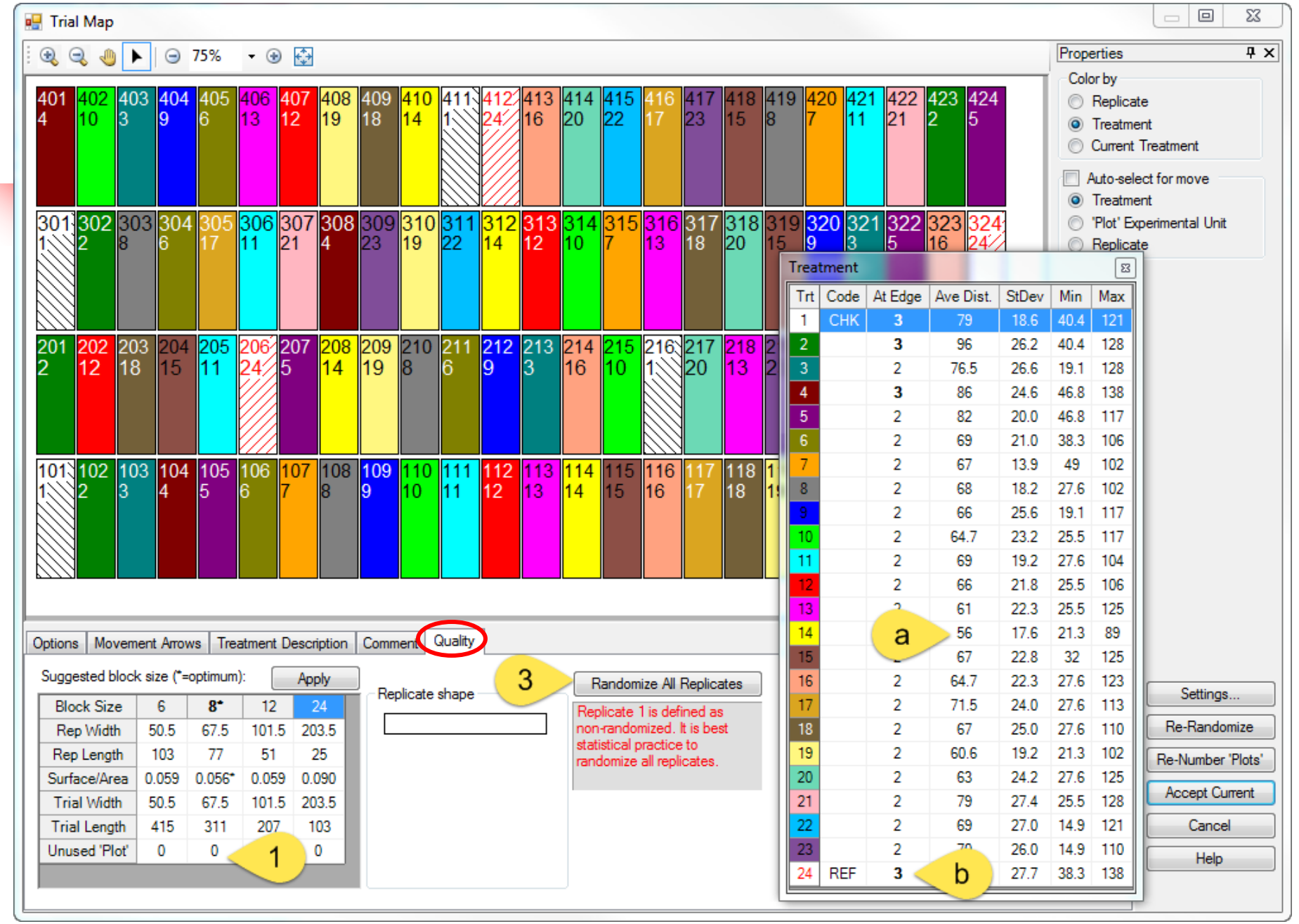
CV	Reps	Power	αSL	% Mean Diff	Error DF
2.92				6	
3				6.17	
3.9				8	
4				8.2	
4.86				10	
5	4	80	5%	10.3	27
5.84				12	
6				12.3	
6.8				14	
7				14.4	

Detectable Difference between Trt. Means	
CV	% Mean Diff.
3	6.17% difference
4	8.2% difference
5	10.3% difference
6	12.3% difference
7	14.4% difference

Randomization Quality Review

Goal is to improve experiment precision:

1. Arrange replicates as squares, not strips
2. Equalize treatment distribution
   a. Balance average distance from all other treatments
   b. Balance "Edge effect" across treatments
3. Randomize all replicates


Trial Map
75%
Properties
Color by
Replicate
Treatment
Current Treatment
Auto-select for move
Treatment
'Plot' Experimental Unit
Replicate
Treatment
Trt Code At Edge Ave Dist. StDev Min Max
1 CHK 3 79 18.6 40.4 121
2 3 96 26.2 40.4 128
3 2 76.5 26.6 19.1 128
4 3 86 24.6 46.8 138
5 2 82 20.0 46.8 117
6 2 69 21.0 38.3 106
7 2 67 13.9 49 102
8 2 68 18.2 27.6 102
9 2 66 25.6 19.1 117
10 2 64.7 23.2 25.5 117
11 2 69 19.2 27.6 104
12 2 66 21.8 25.5 106
13 61 22.3 25.5 125
14 56 17.6 21.3 89
15 67 22.8 32 125
16 2 64.7 22.3 27.6 123
17 2 71.5 24.0 27.6 113
18 2 67 25.0 27.6 110
19 2 60.6 19.2 21.3 102
20 2 63 24.2 27.6 125
21 2 79 27.4 25.5 128
22 2 69 27.0 14.9 121
23 2 26.0 14.9 110
24 REF 3 27.7 38.3 138
a
b
Options
Movement Arrows
Treatment Description
Comment
Quality
Suggested block size (*=optimum):
Apply
Block Size 6 8* 12 24
Rep Width 50.5 67.5 101.5 203.5
Rep Length 103 77 51 25
Surface/Area 0.059 0.056* 0.059 0.090
Trial Width 50.5 67.5 101.5 203.5
Trial Length 415 311 207 103
Unused 'Plot' 0 0 0
1
Replicate shape
3
Randomize All Replicates
Replicate 1 is defined as non-randomized. It is best statistical practice to randomize all replicates.
Settings...
Re-Randomize
Re-Number 'Plots'
Accept Current
Cancel
Help

Arrange Replicates as Squares not Strips

"Optimum" is smallest surface-to-area ratio

Options | Movement Arrows | Treatment Description | Comment | Quality

Suggested block size (*=optimum): Apply

Replicate shape

Block Size	6	8*	12	24
Rep Width	50.5	67.5	101.5	203.5
Rep Length	103	77	51	25
Surface/Area	0.059	0.056*	0.059	0.090
Trial Width	50.5	67.5	101.5	203.5
Trial Length	415	311	207	103

Options | Movement Arrows | Treatment Description | Comment | Quality

Suggested block size (*=optimum): Apply

Replicate shape

Block Size	6	8*	12	24
Rep Width	50.5	67.5	101.5	203.5
Rep Length	103	77	51	25
Surface/Area	0.059	0.056*	0.059	0.090
Trial Width	50.5	67.5	101.5	203.5
Trial Length	415	311	207	103

Options | Movement Arrows | Treatment Description | Comment | Quality

Suggested block size (*=optimum): Apply

Replicate shape

Block Size	6	8*	12	24
Rep Width	50.5	67.5	101.5	203.5
Rep Length	103	77	51	25
Surface/Area	0.059	0.056*	0.059	0.090
Trial Width	50.5	67.5	101.5	203.5
Trial Length	415	311	207	103

Options | Movement Arrows | Treatment Description | Comment | Quality

Suggested block size (*=optimum): Apply

Replicate shape

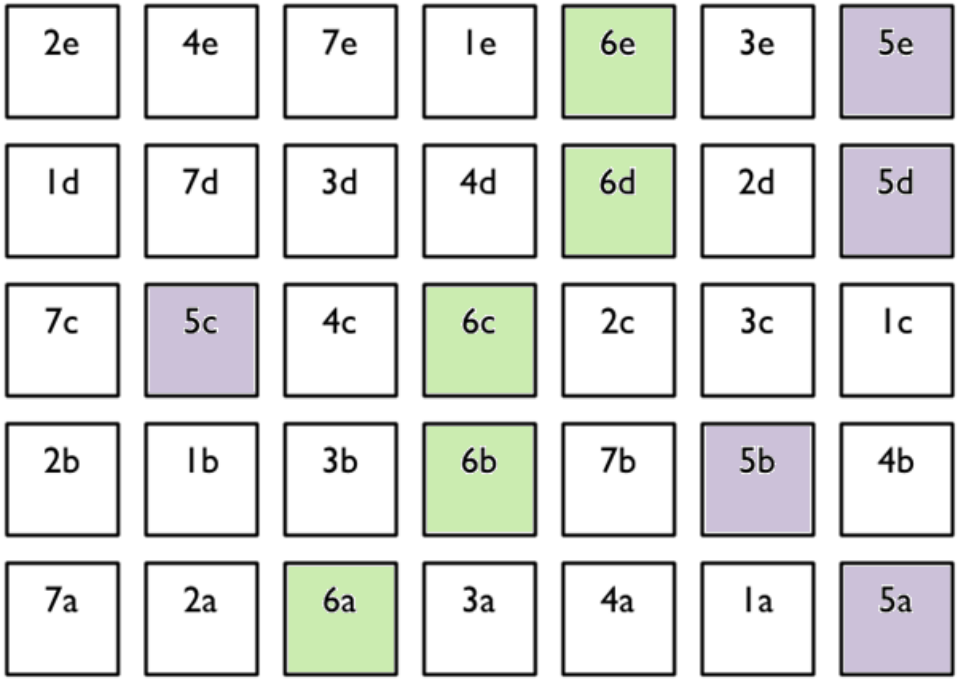
Block Size	6	8*	12	24
Rep Width	50.5	67.5	101.5	203.5
Rep Length	103	77	51	25
Surface/Area	0.059	0.056*	0.059	0.090
Trial Width	50.5	67.5	101.5	203.5
Trial Length	415	311	207	103

Substantially reduces avg. dist. between trt.

Equalize Treatment Distribution

"Undesirable" layout of 7 treatments and 5 replicates in Randomized Complete Block:

- Trt. 6 in middle 3 columns of all reps
- Trt. 5 in right 2 cols for all but one plot



Uses "Average Distance of Treatment" Comparison (ADTC)

- van Es and van Es, "Spatial Nature of Randomization and Its Effect on the Outcome of Field Experiments", Agron J, 85:420-428 (1993).
- Goal is to create spatially-balanced designs.
- Comparison between treatments 1 and 2 is taken from 5 plots for each treatment.
- Measure the plot-to-plot distance for each plot containing treatment 1 to the paired plot within replicate containing treatment 2, for a total of 5 distances.
- ADTC for treatment pair 1-2 is average of the 5 distances.
- Repeat this comparison for all treatment pairs.

Unequal Treatment Distribution

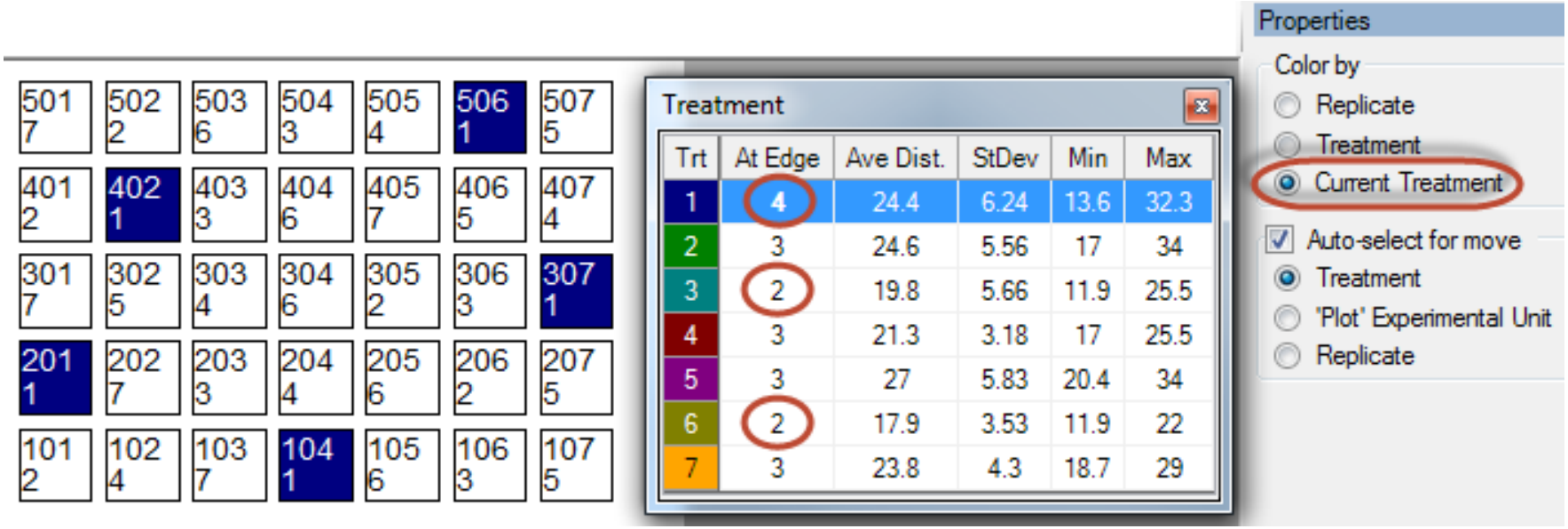
- Average distance from 17.9 to 24.6
- Ranges from 11.9(T3,T6) to 34(T2,T5)
- Error variances for treatments may not be homogeneous

Treatment

Trt	At Edge	Ave Dist.	StDev	Min	Max
1	4	24.4	6.24	13.6	32.3
2	3	24.6	5.56	17	34
3	2	19.8	5.66	11.9	25.5
4	3	21.3	3.18	17	25.5
5	3	27	5.83	20.4	34
6	2	17.9	3.53	11.9	22
7	3	23.8	4.3	18.7	29

Unbalanced "Edge effect"

- Treatment 1 occurs at edge 4 times, T2 and T3 at edge only 2 times



Balanced Treatment Distribution and Edge Effect

- Average distance from 21.3 to 24.4
- Distances range from 18.7 to 27.2
- "Edge effect" is balanced

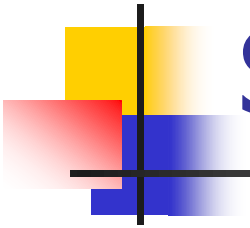
Treatment

Trt	At Edge	Ave Dist.	StDev	Min	Max
1	2	22	2.15	20.4	25.5
2	3	23.8	3.57	18.7	27.2
3	3	24.4	1.76	22	27.2
4	3	22.4	3.47	18.7	25.5
5	3	22	3.4	18.7	27.2
6	3	21.3	2.58	18.7	25.5
7	3	22.7	2.56	18.7	25.5

Post-hoc Power Analysis

- In example, LSD can distinguish 25% mean difference (largest existing difference is 18%)
- Current AOV Trt P(F) is 0.2979, so use 0.30+ significance level to separate treatment means
- Need 8+ replicates to reject null hypothesis at 0.05 significance

Crop Variety	CEZANNE
Trt No.	24
2	85.33 a
3	81.67 a
4	98.00 a
5	95.33 a
LSD P=.05 (% mean diff)	21.808 (25%)
Standard Deviation	10.915
CV	12.12
Grand Mean	90.083
Minimum Replicates (power = 80)	8
Largest Mean Difference (% mean diff)	16.333 (18%)
Treatment F	1.541
Treatment Prob(F)	0.2979



Summary

Software tools can help improve trial quality and efficiency:

- Plan appropriate number of replicates
- Improve quality of randomizations
- Analyze results to improve planning of follow-up experiments