Proc. Fla. State Hort. Soc. 117:112-117. 2004.

IMPROVING EFFICIENCY OF FIELD TRIALS: OPTIMIZING DESIGN BASED ON RESULTS FROM SIMULATED EXPERIMENTS

ED STOVER[1]
University of Florida
Indian River Research and Education Center
2199 S. Rock Rd.
Ft. Pierce, FL 34945-3138

KENNETH PORTIER
University of Florida
Dept. of Statistics
P.O. Box 110339
Gainesville, FL 32611-0339

Additional index words. experimental design, pseudoreplication, randomized complete block design, RCBD, statistical analysis

Abstract. **Simulations were used to contrast the ability of a variety of experimental designs to distinguish defined treatment differences. In all cases, a single experimental treatment was compared to a control with each applied to twelve trees. Analyses were compared in which data was collected on individual tree experimental units or multi-tree experimental units. Trees were blocked according to a specified factor, which was quantified but could represent a qualitative factor such as spatial position. The probability of rejecting the null hypothesis was computed for a range of situations including small and large values for the following parameters: treatment response, standard deviations of the response, blocking factor effects, and blocking factor standard deviations. In all cases, power of the test was significantly greater when data was collected on individual tree experimental units, and decreased as the number of trees pooled per data point increased. Analyses indicated that use of one or two independent trees per treatment per block provides the greatest efficiency in distinguishing treatment effects, and that two trees per treatment per block is superior to one, when there is a significant treatment × block interaction. When the blocking factor had a spatial relationship (neighboring trees were more similar than two randomly selected trees in a block) there was a high probability of incorrectly finding differences significant when trees in multi-tree blocks were treated as though they were independent experimental units or pseudoreplicates.**

Field experiments will remain critical in addressing issues important to commercial citrus production. Proper design of these experiments will increase the efficient use of resources and should aid in identification of commercially significant outcomes. Since field-grown plants often display variations in performance that can be identified prior to experimentation, the randomized complete block design (RCBD) is one of the most frequently used experimental designs for agricultural field studies (Lin et al., 1993). Because of the high degree of variability when measuring parameters in field-grown tree species, some researchers have adopted the strategy of pooling data on multiple tree experimental units in an effort to reduce variability, and instead use fewer but larger experimental units (many scientists, pers. comm.). The objective of this research was to test the hypothesis that use of such multi-tree plots will significantly increase the power of our tests when variability due to random factors or a blocking variable is sufficiently high.

Materials and Methods

Monte Carlo simulations designed in SAS® (SAS Institute Inc., 1999, Cary, N.C.) allowed input of specific values for design factors and all general linear model parameters, permitting comparison of treatments in a wide range of scenarios. In each case, we used simulations to compare a control group to a single experimental group. Prior to analysis, experimental units were blocked as though by spatial orientation or tree size. A Type I error rate of $P < 0.05$ was used for all analyses, except where noted. There were 5000 runs conducted for each experiment.

Models. The general linear model used to generate the responses is defined by the following relationship:

Response variable X = (control mean+ mean treatment effect) + (effect of blocking factor + interaction between blocking factor and treatment) + residual error. The control group was specified as having a mean value of 100 for response variable X. The experimental treatment was specified as increasing the mean value of the response variable X by 0, 5, or 10 units. The mean effect of the blocking factor was also specified as 0, 5, or 10 unit increase above the untreated mean. The block by treatment interaction effect was specified as being either not present (0), additive of the treatment effect (+5 or +10 units) or antagonistic of the treatment effect (-10 or -5 units). The standard deviations selected for the blocking factor and the residual error ranged from 5 to 20, generating a range of individual values for X ranging to extremes of ~60-140 for the control.

The seven block structures that we studied are described in Table 1. When each treatment was applied to more than one tree in each block, we compared use of multi-tree experimental units as in a standard RCBD to multiple single-tree experimental units which were independently randomized within the block, a design known as a Generalized RCBD (Fig. 1). In all cases the total number of "trees" was 24 (12 per treatment), permitting comparison of blocking strategies with the same amount of effort expended on data collection.

Simulations tested. The range of comparisons run in our simulations is described in Table 2. Most comparisons were made with random variability between trees within each block so that adjacent trees were no more similar on average than any two trees in a block. Some comparisons were made in which the blocking factor was arranged in an escalating fashion such that adjacent trees tended to be more similar to each other than to more distant trees in the same block, which is likely to be representative of many field experiments with spatial blocking.

Testing. In the simulations, two scenarios were compared. In the first scenario, each tree was used as an experimental

---

This research was support by the Florida Agricultural Experiment Station and approved for publication as Journal Series No. R-10313.

[1] Corresponding author; e-mail: ewstover@ifas.ufl.edu.

Table 1. Factors common to all simulations.

Treatments	Number of simulated experimental iterations	Values for the mean for response variable in controls	Values for the standard error of response variable	Values for the standard error of blocking factor	Number of blocks	Trees/ treatment/ block	Error term for F-test with collection of data from individual trees	Error term for F-test with pooled collection of data from groups of trees
Control and treated	5000	100	5, 10, & 20	5, 10, & 20	12	1	block × tree MS	not applicable
					6	2	MSE[z]	block × tree MS[y]
					4	3		
					2	6		

[z] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block.
[y] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block.

unit, as in a generalized RCBD. In this design each treatment is individually assigned to a separate tree, even if there are multiple trees of the same treatment in each block, and the residual error is properly estimated by tree-to-tree variation pooled across each block by treatment. Therefore, the treatment mean squares (MS) is divided by the error MS from the ANOVA table to generate the appropriate F test. In the second scenario, treatments were assigned randomly to groups of trees within each block (e.g., a spatial block of 1,2 3, or 6 trees) and data were pooled across all trees in each experimental unit. In this case the proper F test is generated by dividing the treatment MS by the block × treatment MS from the ANOVA. The results of these two approaches are compared.

Results and Discussion

These studies evaluated many combinations of variables and produced an extensive array of data. Only a small subset, demonstrating the most important relationships, is presented in this paper.

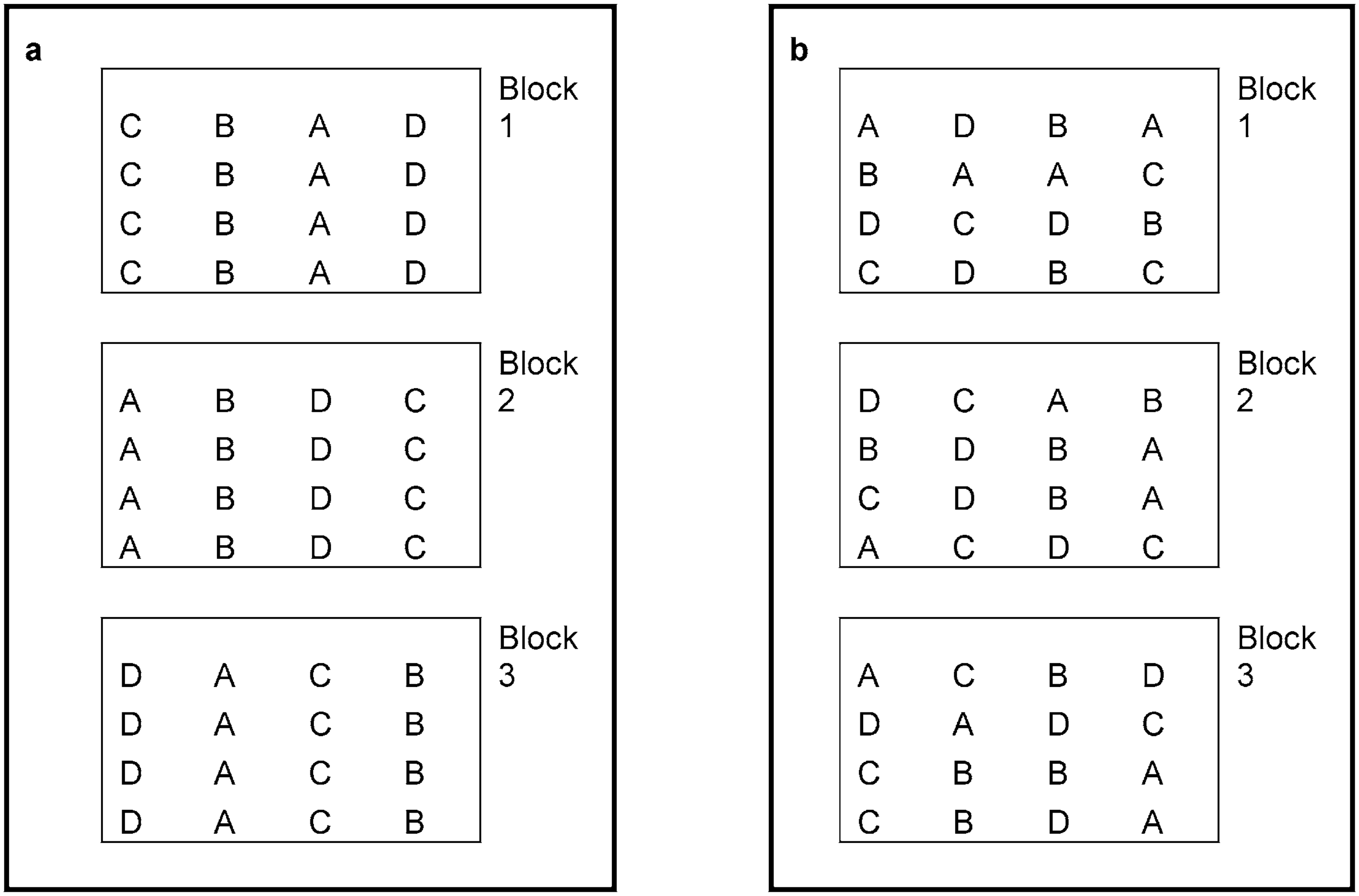

Fig. 1. Depiction of two approaches to a Randomized Complete Block Design with multiple trees receiving each of four treatments (A, B, C, and D) in each of three blocks. Fig. 1a. illustrates the standard RCBD, in which a single four-tree experimental unit receives each treatment. The proper error term for the F-test is Block × Treatment with 6 degrees of freedom [(#Blocks-1) × (#Treatments -1)]. Fig. 1b. illustrates the generalized RCBD, in which four single-tree experimental units are independently randomized within each block for each treatment. The proper error term for the F-test is the default MSE, which is Tree (Block × Treatment) which has 36 degrees of freedom [(#Blocks) × (#Treatments) × (#Trees per treatment per block – 1)].

Table 2. Factors tested in each experiment.

Experiment	Values for the effect of treatment on the response variable	Values for the slope of the response variable vs. the blocking factor	Values for the effect of treatment on the slope of the response variable vs. the blocking factor	Relationship between value of blocking factor covariate and assignment of experimental units	Table(s) listing results
1—no effect of treatment on response variable	0	0, 0.5, 1.0	0	Each block is uniformly variable and adjacent trees are likely no more alike than any two trees in the same block	data not shown
2—no blocking factor effect but treatment increases response variable by 5 and 10%	5,10	0	0	As above	Table 3
3—blocking factor vs. response variable has moderate to large slope, treatment increases response variable by 5 and 10% but doesn't influence slope of response variable vs. blocking factor	5, 10	0.5, 1.0	0	As above	Tables 4 & 5
4—blocking factor vs. response variable has moderate slope, treatment increases response variable by 5% but also influences slope of response variable vs. blocking factor	5	0.5	-0.5, -0.1, 0.1, 0.5	As above	Table 6
5—blocking factor vs. response variable has no, moderate, or large slope, treatment increases response variable by 0, 5, or 10% but doesn't influence slope of response variable vs. blocking factor	0, 5, 10	0, 0.5, 1.0	0	Effect of blocking factor generally increases from one end of the block to the other, so that adjoining trees are more similar to each other.	Table 7

Responses from these studies included many that one would expect in a properly functioning model system. When there was no effect of treatment on response variable X, the *Pr*(Reject) averaged 0.050 which was the type I error rate (probability of incorrectly rejecting the null hypothesis) designated for this analysis (data not shown). When the treatment effect was set as zero and the variability within each block was random, there were no apparent differences among different blocking strategies or other experimental parameters (data not shown). Also, as one would expect, the probability of correctly rejecting the null hypothesis was greater as the effect of treatment on response variable increased relative to response variable standard error (data not shown).

When the blocking factor is not significant, the number of blocks has minimal effect if single tree experimental units are used, but the power of the test declines as a fixed number of trees are allocated into fewer multi-tree experimental units (Table 3). However, when the blocking factor significantly influences the response variable, there is often a considerable benefit to using more blocks even when single tree experimental units are used. In this scenario, with no significant block × treatment interaction, more and smaller blocks always have higher power than fewer and larger blocks of multi-tree units. This is true whether the association between the blocking factor and the response variable is moderate (Table 4) or strong (Table 5).

When the treatment influences the slope of the blocking factor vs. the response variable (meaning there is an interaction between the block and the treatment), the effect of experimental design was strongly influenced by degree of slope alteration. Small treatment influences on slope resulted in similar power for the designs with 12 blocks of one tree per treatment and 6 blocks of two trees per treatment (data not shown). However, when treatment substantially altered the slope of response variable vs. blocking factor, there was a major advantage in collecting data on multiple single-tree exper-

Table 3. Simulation results with no blocking factor effect and experimental increment of 5 percent.

				12 blocks of 1 tree/treatment	6 blocks of 2 trees/treatment		4 blocks of 3 trees/treatment		2 blocks of 6 trees/treatment	
				P of rejecting null hypothesis						
					Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data
Slope of response variable vs. blocking factor	Treatment effect on response variable	Standard error of response variable	Standard error of blocking factor	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]
0	5	5	5	0.62	0.51	0.62	0.39	0.62	0.14	0.64
			10	0.61	0.50	0.61	0.39	0.63	0.15	0.64
			20	0.61	0.49	0.60	0.40	0.63	0.16	0.65

[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block.
[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block.

Table 4. Simulation results with moderate blocking factor effect and experimental increment of 10 percent. There is no treatment effect on slope between blocking factor and response variable.

				12 blocks of 1 tree/treatment	6 blocks of 2 trees/treatment		4 blocks of 3 trees/treatment		2 blocks of 6 trees/treatment	
				P of rejecting null hypothesis						
					Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data
Slope of response variable vs. blocking factor	Treatment effect on response variable	Standard error of response variable	Standard error of blocking factor	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]
0.5	10	5	5	0.99	0.95	0.99	0.83	0.99	0.22	0.93
			10	0.98	0.88	0.95	0.63	0.90	0.15	0.61
			20	0.90	0.61	0.70	0.33	0.53	0.08	0.22

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block.

[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block.

imental units for each treatment in a block (Table 6). Data on more than one tree per treatment per block allows the separation of block × treatment interaction and subsequent removal of this source of variability from the error term (Schabenberger, 2001). However, there was never a scenario in which more than two trees per treatment per block was advantageous. For maximum efficiency, the minimal amount of within block replication needed to estimate a true residual error should be used.

In many field experiments with spatial blocking, adjacent trees may be more alike than two more distant trees within the same block. This was simulated by arranging the blocking factor in an escalating fashion for treatment assignment. When the effect of treatment on the response variable was zero: *Pr*(Reject) was around 0.05 when the blocking effect was also set at zero, but when there was a blocking effect, the *Pr*(Reject) was quite high when trees in multi-tree blocks were treated as though they were true replicates and the inappropriate SSE error term was used (number 5, Table 2). Therefore, the null hypothesis is incorrectly rejected in a rather high proportion of trials (up to 67%) with larger multi-tree experimental units (Table 7).

Use of individual tree data from trees which were not independently assigned to treatments is known as pseudoreplication (Heffner et al., 1996). This practice violates proper statistical practice, and can result in inaccurate statistical inference when similarity between adjacent plants results in a marked underestimate of tree to tree variability. When pseudoreplicates are used with a blocking effect which varies across the block as well as between blocks, and there is a treatment effect, use of multi-tree experimental units does increase the likelihood of finding the treatments significantly different (Table 7), but this is largely a spurious effect from using the wrong error term, and reflects the probability of erroneously finding even a non-significant treatment effect to be significant. The danger of accidently analyzing samples as pseudoreplicates can be avoided by conducting the analysis on the means for multi-tree experimental units, since there should be no difference in the F-test for use of means or individual samples in a properly conducted test.

Using this Information. In this study we have compared the power of the statistical test for treatment effects under the constraint of fixed total size of the experiment. This has allowed direct comparison of design effects without other con-

Table 5. Simulation results with large blocking factor effect and experimental increment of 10 percent. There is no treatment effect on slope between blocking factor and response variable.

				12 blocks of 1 tree/treatment	6 blocks of 2 trees/treatment		4 blocks of 3 trees/treatment		2 blocks of 6 trees/treatment	
				P of rejecting null hypothesis						
					Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data
Slope of response variable vs. blocking factor	Treatment effect on response variable	Standard error of response variable	Standard error of blocking factor	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]
1	10	5	5	0.98	0.87	0.95	0.64	0.90	0.39	0.59
			10	0.90	0.59	0.70	0.34	0.51	0.08	0.22
			20	0.56	0.27	0.29	0.14	0.19	0.06	0.09

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block.

[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block.

Table 6. Simulation results with medium blocking factor effect and medium treatment effect on slope between blocking factor and response variable and experimental increment of 5 percent.

					12 blocks of 1 tree/ treatment	6 blocks of 2 trees/treatment		4 blocks of 3 trees/treatment		2 blocks of 6 trees/treatment	
					P of rejecting null hypothesis						
						Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data	Group of trees pooled for data	Individual tree data
Slope of response variable vs. blocking factor	Change in slope of response variable vs. blocking factor	Treatment effect on response variable	Standard error of response variable	Standard error of blocking factor	Block × trt as error for F test	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]	Block × trt as error for F test[z]	MSE as error for F test[y]
0.5	0.5	5	5	5	0.23	0.06	0.50	0.02	0.45	0.01	0.28
				10	0.07	0.01	0.37	0.00	0.29	0.00	0.14
				20	0.02	0.01	0.28	0.00	0.20	0.01	0.10

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block.

[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block.

founding effects. These analyses demonstrate that designs in which individual trees are the experimental units are almost always superior to pooling data across several trees for a fixed number of trees per treatment. Our results also suggest that there is no significant advantage to using an experimental design with more than 2 experimental units per treatment per block. Even in the situation where blocking is important and there is a block by treatment interaction, only the minimal replication needed to allow estimation of both treatment × block and pure error effects is needed. However, where there is no treatment × block interaction, and a blocking factor with substantial effect can be used, considerable power is lost by not maximizing the total number of blocks used.

The experimenter's knowledge or expectations of the blocking factor effect and likely blocking factor × treatment interaction should be used to select the best experimental design. If these factors are totally unknown to the experimenter, perhaps two experimental units per treatment per block should be used. Clearly the presence of a blocking factor × treatment interaction greatly increases the power of this design vis-à-vis a single experimental unit per treatment per replication and if there is no interaction, the loss of power is modest compared to use of a single tree per treatment per block.

Certainly there may be advantages to increasing the total number of trees on which data are collected by choosing to sample multiple trees using a similar number of independent experimental units. In some cases logistics of treatment application or plot characteristics may make it preferable to use multiple tree experimental units for more accurate application, avoidance of nearest neighbor effects, or efficient use of

Table 7. Simulation results with ordered assignment to simulate spatial blocking, in which consecutive trees form multi-tree experimental units and neighboring trees are more like each other than more distant trees in the same block. Experimental increments of 0 and 5% are compared with no or modest block effect on response variable. There is no block × treatment interaction.

				12 blocks of 1 tree/treatment	6 blocks of 2 trees/treatment		4 blocks of 3 trees/treatment		2 blocks of 6 trees/treatment	
				P of rejecting null hypothesis						
Slope of response variable vs. covariate	Treatment effect on response variable	Standard error of response variable	Standard error of covariate	Block × trt as error for F test	Block × trt as error for F test[z]	SSE as error for F test[y]	Block × trt as error for F test[z]	SSE as error for F test[y]	Block × trt as error for F test[z]	SSE as error for F test[y]
0	5%	5	5	0.62	0.51	0.63	0.39	0.64	0.15	0.64
			10	0.62	0.51	0.61	0.40	0.63	0.15	0.66
			20	0.60	0.51	0.62	0.39	0.64	0.16	0.65
0.5	5%	5	5	0.58	0.41	0.58	0.22	0.55	0.09	0.54
			10	0.52	0.26	0.52	0.12	0.52	0.11	0.67
			20	0.37	0.11	0.44	0.08	0.50	0.13	0.64
0	0%	5	5	0.05	0.05	0.049	0.05	0.05	0.05	0.05
			10	0.05	0.05	0.046	0.05	0.06	0.05	0.04
			20	0.05	0.05	0.05	0.05	0.05	0.06	0.05
0.5	0%	5	5	0.05	0.06	0.08	0.05	0.13	0.02	0.47
			10	0.06	0.05	0.14	0.07	0.29	0.12	0.57
			20	0.05	0.05	0.26	0.09	0.45	0.12	0.59

[z] Appropriate for analyzing data from an RCBD in which multiple trees in a group are assigned the same treatment within each block.

[y] Appropriate for analyzing data from a generalized RCBD in which multiple trees are independently assigned the same treatment within each block, in this case it is the wrong error term and treats individual samples in multi-tree experimental unit as a pseudoreplicate, spuriously rejecting the null hypothesis at a high rate.

limited trees with buffer trees between treatments. These results are not intended to suggest that such uses are not appropriate, but rather that there is no inherent advantage to use of multi-tree blocks to buffer treatment variability.

A survey of the ecological literature revealed that a large proportion of the published studies used pseudoreplication (reviewed in Heffner et al., 1996). It is possible that some horticultural researchers who routinely use multi-tree experimental units in field trials have found that such designs are more likely to successfully distinguish treatment differences because each tree is treated as a pseudoreplicate. The analyses presented here underscore that such practices are contrary to sound statistical analysis and may result in misleading and unfounded conclusions.

Literature Cited

Heffner, R. A., M. J. Butler IV, and C. K. Reilly. 1996. Pseudoreplication revisited. Ecology 77:2558-2562.

Lin, C. S., M. R. Binns, H. D. Voldeng, and R. Guillemette. 1993. Performance of randomized block designs in field experiments. Agron. J. 85:168-171.

Schabenberger, O. 2001. The difference between the GRBD and the RCBD with subsampling. 15 June 2004. <http://www.stat.vt.edu/~oliver/stat5616/handouts/GRBDvsRCBDSS.pdf>.