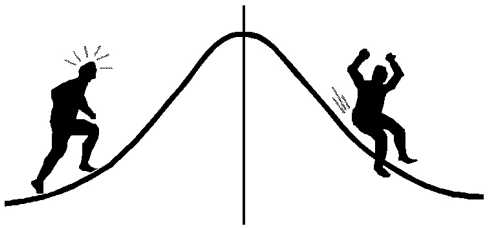


BIOMETRICS INFORMATION

(You're 95% likely to need this information)

PAMPHLET NO. # 34 DATE: September 13, 1991

SUBJECT: When are blocks pseudo-replicates?

A fundamental concern when designing experiments is proper replication of the treatments. Most would agree with this statement, but disagreements about replication can occur because, for instance, many consider replication and blocking to be synonymous terms. My definition of replication is more general, namely: *if an experiment has replication*, *then at least one treatment has been assigned to two or more treatment units*. Of course it is preferable for all treatments to be assigned to two or more treatment units. Blocking is one way to accomplish this replication but there are many other methods (see, for example BI#17: What is the Design?). This pamphlet is somewhat philosophical in nature with the intent of encouraging you to think about replication and to consider when blocks provide proper replication.

Before proceeding with the discussion I would like first to define my terms more precisely. A treatment is some set of experimental conditions which is either assigned to treatment units or which is observed about the treatment units. For instance, a fertilizer application of 200 kg/ha is a treatment which can be assigned, while noting that individual trees are either Df or Hw is an observed treatment if species is a factor in the experiment. While I use the term treatment loosely, I generally use it to refer to one combination of experimental conditions, such as an amount of fertilizer applied at the same time as a level of root-pruning. If I am referring to a class of treatments, such as a fertilizer applied at 4 different levels, I tend to call this a factor instead of a treatment. Hopefully it is clear from the context, if I mean one specific combination of treatments or a class of treatments (factor).

The definition of treatment units (t.u.'s) is also important. A treatment or experimental unit is some unit of experimental material which has a chance of receiving one treatment independent of what treatment another t.u. receives. As an example, suppose that rows of 50 trees are to be planted. Each row will be assigned either none, 100 kg/ha or 200 kg/ha of fertilizer. Each tree in a row receives the same fertilizer treatment and so trees cannot be the t.u. But if row 1 receives 100 kg/ha, row 2 could still receive any of the three treatments (assuming random assignment of treatments to t.u.'s), so rows are the t.u.'s. Pseudo-replication would occur if the trees are considered to be t.u.'s (see BI#5: Understanding Replication and Pseudo-replication).

Ideally, the t.u.'s used in an experiment would be a random sample of t.u.'s taken from some well-defined population. The trees in a seed orchard, say of one species and of similar age, would provide a very clearly defined population. The results of the experiments on the sample of t.u.'s can then be generalized back to the originally defined population so that we can infer something

about that population. If root-pruning a random sample of trees in the seed orchard encouraged those trees to flower when compared to a randomly selected control sample then we would infer that root-pruning would also encourage the other trees in the seed orchard to flower.

One of the problems we have in forestry research is that the population of interest is usually poorly defined. It is difficult to randomly sample from a poorly defined population. For that matter, even when the population of interest is well-defined, proper random sampling is difficult to accomplish in most forestry applications. Since the concept of randomly sampling for the t.u.'s to be used in experiment is rarely discussed in textbooks, or in courses on experimental design, it is not surprising that many researchers aren't aware of it. Nevertheless it is an important consideration when drawing valid inferences from our experiments.

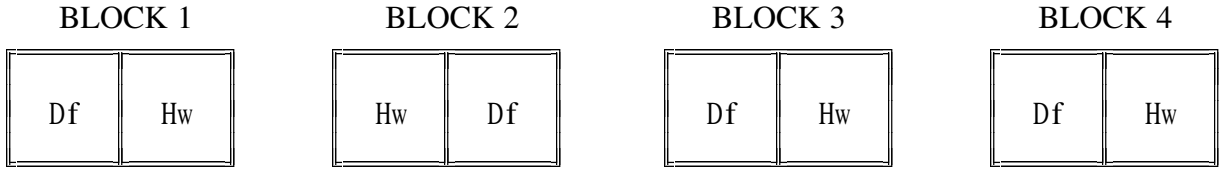
If a researcher has done an experiment on 2+0 Df stock readily found in a nursery, it is clear that the results of the experiment say something about the trees in the experiment. But to what other trees can the results be inferred? All the 2+0 Df stock for that year in that nursery? All the 2+0 Df stock for any year from that nursery? All 2+0 stock of any species for any year from any nursery? When the trees are not randomly selected from a well-defined population, we are unsure of what population of trees it is legitimate to extend our inferences.

Let's discuss another example. Suppose we want to do a species trial comparing the growth of planted Df and Hw. There are many questions that we need to answer before even starting to design the trial. From what population of Df and Hw nursery stock will the trees be selected for the experiment? Will we ignore questions regarding differences between nurseries, growing regimes, stocktypes, and seedlots, to name only a few? For instance, we may take 1+1 trees from one seedlot from each species grown that year in one nursery. How will this limit the inferences we can make about the results of our trial?

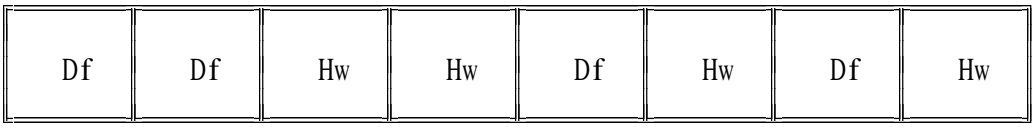
Suppose that we have limited our study to sites in the CWH subzone. On what sites will these trees be planted? Ideally, a random sample of suitable sites in the CWH subzone would be obtained and used. If this was practical, how many sites would we need? Can we replicate adequately on just one site? The answers to these questions depend on the inferences we want to make about the results of our experiment. Suppose our experiment indicates that Df has better growth than Hw. Will we want to say that we would expect this to occur at all other sites in the CWH subzone? If so, and our experiment occurs on only one site in the CWH, just how strong is the inference? Will our argument be readily subject to attack from those who disagree with our conclusions? I think so. Essentially, proper replicates or primary sampling units (p.s.u.'s) of the CWH subzone are needed if inferences are to be made to the whole of the CWH subzone.

Suppose that 8 plots (t.u.'s) will be used, each containing 50 trees. Let's look at some possible layouts for this trial.

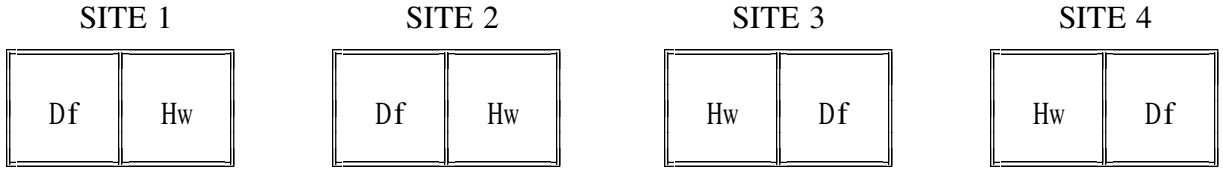
1) a randomized block layout at **one site**:



2) a completely randomized layout on **one site**:



3) one set of treatments laid out on each of **several sites**:



The accompanying ANOVA tables are:

Layout 1: RB at 1 site

Source of Variation	df	Error
Block B	3	--
Species P	***1***	***BxP***
BxP	3	--
Treatment units:		
E(BP)	0	--
Total	7	

Layout 2: CR at 1 site

Source of Variation	df	Error
Species P	***1***	***E(P)***
Treatment units:		
E(P)	6	--
	7	

Layout 3: Several sites

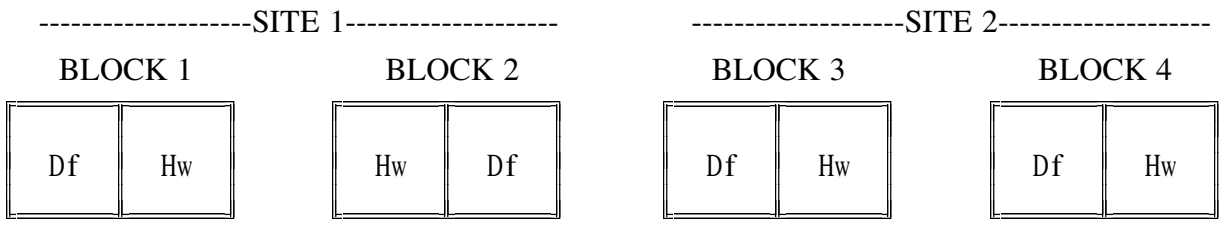
Source of Variation	df	Error
Site S	3	--
Species P	***1***	***SxP***
SxP	3	--
Treatment units:		
E(SP)	0	--
	7	

All three layouts have replication of species and all have F-tests available. Layouts 1 and 3 both have F-tests with 1, 3 df while layout 2 has 1, 6 df. Layout 1 has allowed for within-site variability by blocking, while layout 2 does not block within-site variability and assumes that it is adequately measured by the between plot (within species) variability. Layout 2 is rarely used by

researchers since they feel uncomfortable assuming that a large area is truly homogeneous. If a block design is placed in an homogeneous area then some df, and therefore power, is lost in the test for species differences. The pictures for layouts 1 and 3 don't look very different and the ANOVA tables appear identical. Site is acting like a block in the familiar RB design. Nevertheless, the inferential scope is quite different for the two layouts.

The design with 4 blocks at one site (layout 1) has done a fine job of measuring within-site variation, but there is no measure of between-site variation. If the results are to be generalized to the CWH subzone then only one realization or replicate of the CWH has been examined. How can we feel assured that what was observed at this one site would occur at other sites within CWH? Layout 3 on the other hand, will obtain very little information about within-site variability while obtaining much more about between-site variability. The CWH subzone has been sampled four times instead of once. If, for instance, Df has better growth on all four sites, then we are more confident that this pattern would occur at other sites within the CWH than we would be with similar results from layout 1. To see these differences more clearly, let's examine a layout that has replication of the CWH subzone along with within-site replication.

4) a randomized block design placed at two sites:



Let's compare the ANOVA for this new layout with those of layouts 1 and 3.

Layout 1: RB at 1 site

Source of Variation	df	Error
Block B	3	--
Species P	***1***	***BxP***
BxP	3	--
Treatment units:		
E(BP)	0	--
Total	7	

Layout 3: Several sites

Source of Variation	df	Error
Site S	3	--
Species P	***1***	***SxP***
SxP	3	--
Treatment units:		
E(SP)	0	--
	7	

Layout 4: RB at 2 sites

Source of Variation	df	Error
Site S	1	B(S)
Block B(S)	2	--
Species P	***1***	***SxP***
SxP	1	B(S)xP
B(S)xP	2	--
Treatment units:		
E(SP)	0	--
	7	

All three of these ANOVA's treat site and block as random factors. The ANOVA for the fourth layout looks complicated but it allows between-site differences to be tested. When layout 1 is used to make inferences to other sites within the CWH, the assumption is that between-site

differences are no greater than within-site differences. Layout 4 allows this assumption to be tested directly. Nevertheless, the most important point to note about layout 4 is that the ***correct*** test for species is the site by species interaction. The within-site replication provided by the blocks does not directly help the test for species differences. In fact, the species test for layout 1 must rely on the blocks as ***pseudo-replicates*** to obtain an F-test. This point is quite subtle but very important.

Unless within-site variation is of direct interest, layout 3 is the best from a statistical point-of-view. Of course, this layout is often too expensive or impractical to implement. But when research depends on case studies as implemented by layout 1, it is important for the researcher to understand the limitations of the design. Generalizations about species behaviour to other sites within the CWH has a much sounder statistical foundation with layouts 3 or 4 than with either layouts 1 or 2 (since layout 2 suffers from the same shortcomings as layout 1, namely the reliance on pseudo-replication to provide an F-test). Unless there is some good reason to study within-site variation, layout 3 is better than layout 4 since there are more proper replicates of the CWH subzone available. A possible practical advantage of layout 4 is that each site can be analysed separately using the layout 1 ANOVA.

Under the right circumstances, layout 3 is an ideal design for the establishment of operational trials with limited resources. Although each installation/site has unreplicated treatments, if several installations at different places and at different times can be established, then the various installations taken together form a complete design, with the advantage of replication over a wide range of situations. There must be consistency of installation establishment over space, time, and installation methods for this approach to work. For instance, all installations must have at least two treatments the same (one of which might be a control). The physical layout of plots, rows, number of trees per row, etc. must be the same at each installation and the application of the treatments must be identical. These sorts of conditions must be met or the various installations could not be considered blocks of one large experiment or trial. An excellent example of this approach is described in the protocol for PROBE prepared by Suzanne Simard of the Forest Sciences Section of the Kamloops Regional Office.

I hope that you have found this discussion thought-provoking. If you have comments you would like to send me I would be most interested in reading them, and if I receive a number of interesting comments I may put them together into a pamphlet. For that matter, I invite readers to not only suggest pamphlet topics to me, but to also consider writing their own pamphlets on topics that would be of wide interest. I would, of course, act as an editor with the appropriate editorial privileges.

CONTACT: Wendy Bergerud
387-5676

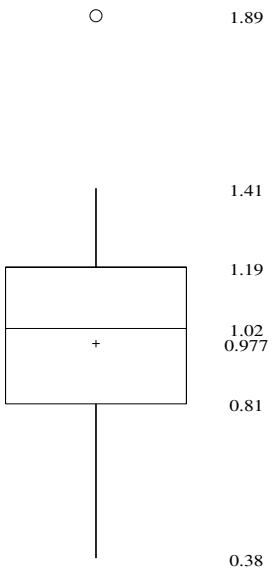
PROBLEM FROM BI#33

The data ranked in increasing order:

0.38 0.47 0.55 0.59 0.65 0.71 0.80 0.81 0.82 0.84

0.85 0.86 0.89 0.99 1.02 1.02 1.02 1.03 1.05 1.08

1.09 1.09 1.19 1.22 1.23 1.23 1.25 1.27 1.41 1.89



mean = 0.977;

median=1/2(15th data + 16th data) = 1.02

lower hinge = 8th data = 0.81;

upper hinge = 23rd data = 1.19;

H-Spread = 1.19 - 0.81 = 0.39;

Step = 1.5(0.39) = 0.57;

upper inner fence = 1.19 + 0.57 = 1.76;

upper outer fence = 1.76 + 0.57 = 2.33;

lower inner fence = 0.81 - 0.57 = 0.24;

lower outer fence = 0.24 - 0.57 = -0.33.

The upper whisker ends at the point 1.41 and the lower whicker ends at the point 0.38. There is one mild outlier at 1.89. The box plot for this set of data is shown on the left.