

Using the Factor Relationship Diagram to Identify the Split-plot Factorial Design

prepared by
Wendy A. Bergerud

Research Branch
B.C. Ministry of Forests
P. O. Box 9519 STN Prov. Govt.
Victoria, B.C.

Introduction

Split-plot experimental designs can be hard to identify and understand. This paper will show how to correctly identify split-plot designs and will contrast them to simple factorial designs. The process of developing a factor relationship diagram (FRD) can help elucidate the correct design. An example will be used to show how the FRD works and the appropriate SAS code to obtain the correct ANOVA F-tests.

Factors and their relationships

A *treatment factor* is an independent variable that may affect the response variable. It is controlled by the researcher and is often of specific interest to the researcher. Its levels can be randomly assigned to the experimental material. A *classification factor* is similar to a treatment factor except that its levels are not assignable because it is an inherent property of the experimental material. Examples include gender, occupation, and tree species.

The experimental material or subject of the experiment is usually arranged in a hierarchy of differently sized nested units (the unit hierarchical structure). For instance, trees may be arranged into rows or groups of, say, 20 trees. These rows may then be grouped into blocks. Measurements may be taken by subsampling the trees. For example, the number of cones may be counted on only two of the many branches on each tree. Another example would be a study of school children in which the children are arranged into classrooms, classrooms into schools and schools into school districts. This is a nested hierarchical structure, with each level a different *unit factor*.

While unit factors are generally nested within each other, treatment and classification factors are often crossed. When two factors are *crossed* this means that the levels of one factor occur in the experiment with all the levels of the other factor. The treatment and classification factors, together with their relationships, form the treatment structure of an experiment. We develop the full description of the experimental design when we assign the different treatment and classification factors to appropriate unit factors in the unit hierarchical structure.

Factor Relationship Diagram Construction

The factor relationship diagram is used to display the relationships between the various treatment, classification and unit factors of an experimental study. It can be constructed using the following three steps.

Step 1: Identify all the treatment, classification and unit factors in the study. Assign unique numbers to each unique level for each factor. This will help clarify whether factors are crossed or nested.

Step 2: Draw the unit structure with the largest grouping of experimental material placed on the top row and successively smaller groupings placed below.

Step 3: Determine which unit factor is the experimental unit for each treatment and classification factor. Then place the treatment and classification factors above their experimental unit. If two or more factors have the same experimental units then they can be placed in any order. The actual random assignment of factor levels to their experimental units is omitted to assist in the identification of the factor relationships.

Factorial Designs

The general factorial design has two or more treatment or classification factors which are crossed with each other. In the simple factorial, all the combinations of factor levels are randomly assigned to the elements of just one unit factor. This design is known as the Completely Randomized Design (CRD).

Suppose a forestry experiment is to be done with three different types of fertilizer (treatment factor F) and two different amounts of a boron supplement (treatment factor A) to be applied to orchard trees. If these treatments are crossed then there will be six different treatments in all. Suppose that the experimental material are young trees arranged into rows of two trees each. The unit structure has two factors: individual trees which are nested within the second factor of rows. A simple, completely randomized design (CRD) results if all six treatment levels are randomly assigned to the rows of 2 trees each. The rows are the experimental units. Trees are the subsamples, since measurements for most response variables must be taken on each tree individually.

The FRD for this design is shown in Figure 1, and the SAS program is shown below. Note that the test statement is necessary to get the correct F-tests. With subsampling, the default error term is not the correct error term for the treatment factors.

ANOVA Table:

<u>Source of Variation</u>		<u>df</u>	<u>Test</u>
Fertilizer	F	2	R(FA)
Boron	A	1	R(FA)
F x A		2	R(FA)
Rows	R(FA)	6	E(RFA)
Trees	E(RFA)	12	--
Total		23	

SAS Program:

* variable names: f for fertilizer types a for boron supplement
row for row of trees y for the response variable;

proc glm;

class f a row;

model y = f | a row(f*a);

test h = f | a e=row(f*a); *<== required because of the sub-sampling;

run;

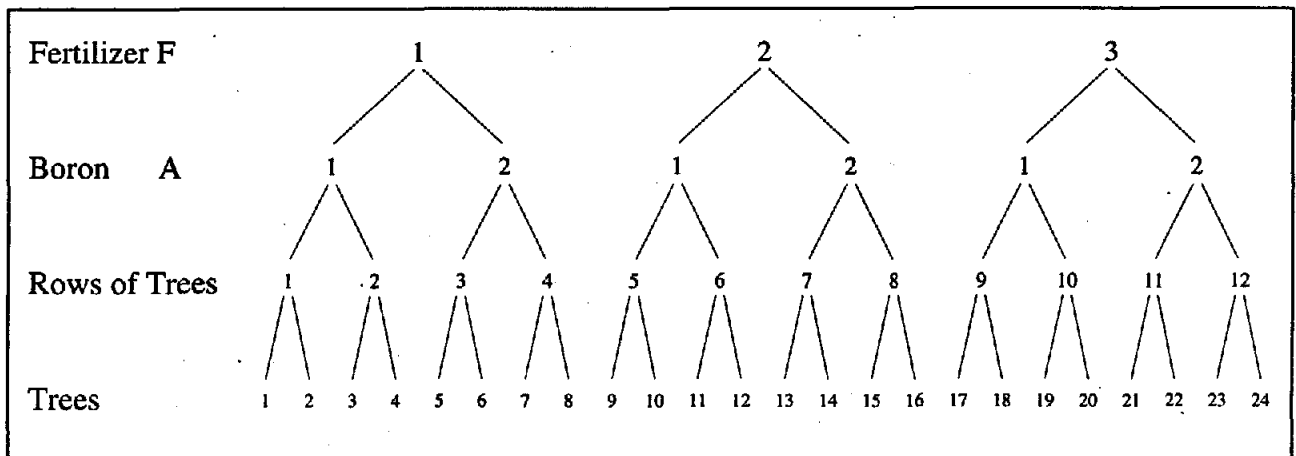


Figure 1. Factor Relationship Diagram for a two-way factorial design with subsampling

Split-plot Designs

Split-plot designs are a specific form of a factorial design, where different treatment or classification factors are assigned to different levels in the unit hierarchy. To show how this works, let's change the example. Suppose that the three different fertilizers are randomly assigned to the rows of two trees as before, but that the boron supplement is now assigned to individual trees within each row. This is shown in Figure 2. Notice that there is no subsampling in this design. Thus the default error is the split-plot error term and does not need to be specifically mentioned in the SAS program.

ANOVA Table:

<u>Source of Variation</u>	<u>df</u>	<u>Test</u>	
Fertilizer	F	2	R(F)
Rows	R(F)	9	--
Boron	A	1	A x R(F)
F x A		2	A x R(F)
A x Rows	R(FA)	9	E(RFA)
Trees	E(RFA)	0	--
Total		23	

SAS program.:

* variable names: f for fertilizer types a for boron supplement
 row for row of trees y for the response variable;

```
proc glm;
```

```
  class f a row;
```

```
  model y = f | a row(f);
```

```
  test h = f e=row(f);
```

```
run;
```

Suppose that tree response was determined by counting the number of cones produced on two branches of each tree. In this case, these branch responses are subsamples at the split-plot level. The SAS program must then include the subsampling and appropriate test statements.

SAS program.:

* variable names: f for fertilizer types a for boron supplement
 row for row of trees y for the response variable;

proc glm;

class f a row

model y = f | a row(f) a*(row*f);

test h = f e=row(f);

test h = a f*a e = a*(row*f); *<== required because of the sub-sampling ;

run;

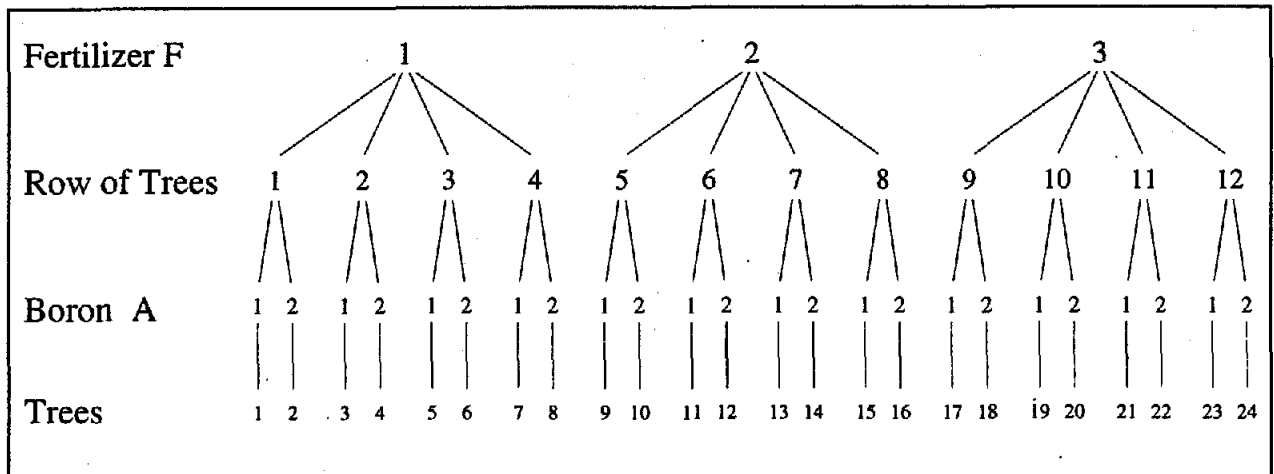


Figure 2. Factor Relationship Diagram for a split-plot design without subsampling

This example split-plot design was derived from the completely randomized factorial design whose experimental units or mainplots (rows of 2 trees) were split into individual trees for the split-plot unit factor. Another common split-plot design arises when a randomized block factorial design (RBD) is the starting point instead. In this case, the experimental units or mainplots which are assigned at least one treatment or classification factor are grouped into blocks. The split-plot design arises when these experimental units are split into smaller units. These smaller units or split-plots are then either randomly assigned the levels of another treatment factor or their level of another classification factor is observed.

Bibliography

Bergerud, W.A. 1996. *Displaying factor relationships in experiments*. The American Statistician, 50(3): 228-233.

Milliken, G.A., and Johnson, D.E. 1984. *Analysis of messy data, Volume 1: Designed experiments* Lifetime Learning Publications, Belmont, CA.