

Statistical Consulting Report for
Cari Sasser Furiness
Department of Forestry and Environmental Resources

June 21, 2009
By Fikret Isik
Statistical Liaison of CNR

Objective: Perform greenhouse trial to measure phenotypic variation in loblolly pine seedlings of known families in response to stress conditions and determine correlation with growth potential from progeny tests.

Treatments:

- 1) low light/low fertility
- 2) low light/high fertility
- 3) high light/low fertility
- 4) high light/high fertility

Main hypotheses:

- 1) Seedlings of different families will show different growth response to stress conditions in the greenhouse.
- 2) Seedling growth response to stress conditions will correlate with growth potential from progeny field tests.
- 3) Minor hypothesis: Light will have a greater effect on variation in growth response than fertilization.

We have available to us 300 seeds from 20 loblolly families. We should have space on the tables in the greenhouse for 4000 test pots + border pots, possibly up to 4800 if we extend one table--we're looking into that.

My thinking was that we would treat the families as blocks. We would shade basically 1/2 of the greenhouse tables to impose the light treatment, and within each light treatment, 1/2 the area would get low or no fertilization and half would get high fertilization. Once the seedlings germinate (for each family, we were planning to plant single seeds into 200 pots and plant 100 seeds into trays for potential transplanting), we would randomly assign 50 (possibly 60) pots into each of the four treatment areas. I've sketched out a basic schematic in the attached powerpoint.

Does this sound like a good design? Or is it too simple? Can we treat each of our seedlings of each family in each treatment as a rep, so we have 50 (or 60) reps per family per treatment? Or do we need to set up some replication of the whole design? Or treat groups of seedlings as reps? Are there some confounding factors we should be taking into account?

I'm definitely rusty on setting up controlled comparative experiments and the resultant statistics, so would really appreciate your advice going into this.

COMMENTS:

I suggest a classical randomized complete block split-plot design for the study to test the major hypotheses.

Within each r blocks (replication), the a levels of the between factor LIGHT are randomized to the main units (whole or main plots). Imposing light treatment to greenhouse tables would be easier. Because the 2 levels of LIGHT treatment are randomized to the main plots, they are called the main treatments.

Within each main plot, the 2 levels of FERTILITY are randomized to the 2 subunits (sub-plots). The levels of Fertility are called sub treatments.

The FAMILY is another factor with 20 levels. Each family can be represented by 1 seedling or 15 seedlings per sub-plot (assuming 300 seeds are available). Seedlings are randomly distributed within each sub plot.

Assuming you have 300 seeds per family, the experiment should be replicated (5 blocks). Replication will increase the power of the test. If there are logistic limitations to have 5 reps, then the number of reps can be reduced to 3. In that case each family will have 25 seedlings in each sub-plot.

Using the PLAN procedure of SAS, the following design was generated. See the appendix for the details of the experimental design.

Block 1

Shade High Fertility	Shade Low Fertility
Light High Fertility	Light Low Fertility

NOTE:

- A block can be set up on different tables in a greenhouse. They do not have to be side by side.
- Each family may have different number of seedlings in each Fertility sub-plot. They should be randomly assigned to sub plots in a block. Linear models are robust to handle unbalanced data.
- One or two border seedlings should be planted around the Fertility sub-plots to make sure homogenous conditions for all families.

F tests for different factors should be constructed correctly. I simulated a data set and analyzed using the MIXED procedure of SAS to show appropriate F tests for different factors as follows.

Class Level Information

Class	Levels	Values
Block	5	1 2 3 4 5
Light	2	1 2
Fertility	2	1 2
family	20	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Number of Observations

Number of Observations Read	6000
Number of Observations Used	6000
Number of Observations Not Used	0

Source	DF	Mean Square	Error Term	Error DF	F Value	Pr > F
Light	1	0.000001860	MS(Block*Light)	4	0.00	0.9968
Fertility	1	0.175099	MS(Residual)	5969	2.10	0.1469
Light*Fertility	1	0.004650	MS(Residual)	5969	0.06	0.8131
Block	4	0.205796	MS(Block*Light)	4	2.01	0.2583
Block*Light	4	0.102568	MS(Residual)	5969	1.23	0.2945
family	19	0.099774	MS(Residual)	5969	1.20	0.2473
Residual	5969	0.083191				

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Light	1	4	0.00	0.9968
Fertility	1	5969	2.10	0.1469
Light*Fertility	1	5969	0.06	0.8131

APPENDIX: Split plot experimental design

If there are 300 seeds per family:

Note: 15 seedlings per family should be randomly assigned to each sub-plot.

block	1				2				3			
light	2		1		2		1		2		1	
fertility	2	1	2	1	1	2	1	2	2	1	2	1
family	6	19	11	20	14	18	1	15	12	15	5	8
	13	20	19	15	20	19	10	12	13	1	8	18
	2	13	8	2	18	11	5	8	2	18	20	12
	16	6	10	19	19	16	9	14	16	3	17	4
	15	5	14	5	13	17	19	7	10	20	2	15
	19	1	16	6	16	2	11	18	18	14	16	9
	9	4	5	17	15	20	20	17	8	2	15	3
	3	9	13	9	3	8	16	13	3	19	10	1
	17	16	18	3	7	10	14	6	15	6	7	16
	8	7	6	11	17	7	15	10	20	8	4	17
	20	10	9	10	4	13	18	11	4	16	13	5
	5	18	3	18	10	15	13	19	19	11	3	20
	11	17	12	7	2	4	3	2	1	7	6	10
	7	2	20	4	6	1	12	1	9	10	11	14
	12	11	4	13	8	9	7	9	11	13	19	19
	18	15	7	8	1	3	4	4	14	4	1	2
10	14	2	1	11	14	6	3	7	9	14	6	
1	8	17	14	5	6	2	16	6	5	18	11	
4	12	1	12	9	5	17	5	17	17	9	7	
14	3	15	16	12	12	8	20	5	12	12	13	

block	4				5			
light	1			2			1	
fertility	1	2	2	1	2	1	2	1
family	10	19	3	8	11	4	1	1
	15	2	13	6	13	17	3	10
	14	8	11	7	16	13	2	20
	11	6	4	5	2	18	5	18
	9	1	9	14	14	10	7	15
	19	14	15	10	12	8	14	2
	4	5	18	3	7	2	19	9
	18	9	7	18	18	5	17	3
	6	4	19	4	6	9	12	14
	7	17	5	19	5	12	11	17
17	16	2	20	19	19	4	11	

	20	18	20	16	1	7	9	12
	8	15	16	11	8	3	15	8
	12	7	10	1	9	14	8	13
	3	10	12	12	15	15	10	19
	1	3	14	9	3	16	16	4
	13	13	8	17	20	1	6	7
	16	12	6	2	10	11	18	16
	2	20	1	15	4	20	20	6
	5	11	17	13	17	6	13	5

Number of seedlings per sub-plot is 300. Each family has 15 seedlings in each sub-plot.

Block	Light	Fertility	N
1	1	1	300
		2	300
	2	1	300
		2	300