PREDICTING PROVENANCE BY PLANTATION INTERACTION BY THE LATITUDE AND THE EFFECT OF THE SEED SOURCE IN WHITE ASH

Fan H. Kung and Knud E. Clausen[1]

Abstract: Height growth of 19 white ash provenances in 4 plantations showed significant provenance by plantation interaction. The interaction terms were obtained through a "mean polish procedure." Regression models indicated that fast-growing seed sources would gain extra growth in Illinois but would suffer extra loss in Wisconsin. Extra gain or loss of growth in Louisiana and Ohio were random and unpredictable in relationship to the mean performance of the seed sources. Provenances IN6795, TN6728 and KY6792 had below average stability but were specifically adapted to favorable environments. The best model for predicting interaction combined the contribution of the latitude and the effect of the seed source. About 86 percent of the variation in the interaction was explained by the model.

Additional Keywords: Analysis of adaptation, genotypic stability. *Fraxinus americana,* regression analysis

INTRODUCTION

The genotype-environment (GE) interaction is the failure of genetic entries to maintain the same relative rank and level of differences when tested in different environments (Snyder 1972). It is necessary first to determine whether interactions are present and then to consider their importance and effect on subsequent work. From the viewpoint of factorial experiments in forest genetics, GE interaction is detected by a significant F-value for that term. However, a significant F value may just tell one side of the story; the interaction may not be significant if the data are transformed (Campbell and Wilson 1973). In white ash (*Fraxinus americana* L.), we also found that the plantation by year interaction was significant if heights were measured in cm but was not significant if heights were transformed to the logarithmic scale (Kung and Clausen 1983).

When interactions exist, the approach by forest geneticists may be different from that of tree breeders. The tree breeders wish merely to minimize the effect of the GE interaction on their field trials while the forest geneticists wish to understand the causes of interactions.

---

[1] The authors are, respectively, Professor, Department of Forestry, Southern Illinois University at Carbondale, Carbondale, Illinois 62901 and Principal Plant Geneticist, North Central Forest Experiment Station, Forest Service, Forestry Sciences Laboratory, Carbondale, Illinois 62901.

ACKNOWLEDGEMENT

The authors would like to thank Drs. John R. Toliver, Daniel B. Houston, and Hans Nienstaedt for plantation maintenance and data collection.

In this paper we will use an example from white ash to illustrate a regression method which should help forest geneticists to characterize the nature of genotype by environment interaction and at the same time would help tree breeders to identify some of the genotypes that are specifically adapted to favorable environments, or some of the plantation locations that are specifically useful for genetically improved planting materials.

MATERIALS AND METHODS

Seeds were collected from up to 10 native parent trees per stand throughout the natural range of white ash during 1973-1974. The seeds were sown in the Illinois Division of Forestry Nursery at Jonesboro, Illinois and plantations established in 1976 as randomized complete blocks with 5-tree plots and 5 replications. Plantations in Louisiana (LA), Illinois (IL), Ohio (OH) and Wisconsin (WI), which had 45 wind pollinated families from 19 provenances in common, were used in this study.

Total height after 3 and 5 years in the field (ages 4 and 6 from seed) were recorded for each tree. Because families within provenance variance was small (Clausen, et. al. 1981), provenance means were used as observations in the data analysis.

In order to test whether or not any GE interaction was due to scaling, natural logarithmic transformation of height data was also used in the analysis of variance. A fixed effects model was assumed for testing interactions while a random effects model was assumed for variance component (VC) analysis. Data analysis was processed through the Statistical Analysis System (SAS ' 1982).

The provenance by plantation interaction was highly significant both in the original measurements and after the natural logarithmic transformation, but the provenance by age interaction was not significant. Therefore we will use only the data at age 6 for this study (Table 1).

The mathematical model for interaction is defined as

$$X_{ij} = Y_{ij} - M - S_i - P_j$$

where X_{ij} = Interaction of the ith seed source at the jth plantation

Y_{ij} = Height growth of the ith seed source at the jth plantation

M = Population mean

S_i = Effect of the ith seed source

P_j = Effect of the jth plantation

The interaction term, X_{i}., can be calculated easily by a "mean polish procedure" (Kung 1981a). The calculated interactions are presented in Table 2 as observed interaction terms.

Once the interaction terms were obtained, regression analysis using interaction as dependent variable could be run with other independent

Table 1.--Height growth of white ash provenances in 4 plantations at age 6 from seed

Seed Source			Height in Plantation				Mean
State	Stand	Lat.	LA	IL	OH	WI	
	No.	Deg.		cm - -			
TX	6768	30.3	157	279	208	40	171
LA	6738	30.5	154	243	208	42	162
MS	6737	30.8	116	252	246	34	162
MS	6740	33.4	135	322	275	43	194
AL	6733	34.5	89	229	232	38	147
TN	6728	35.3	92	244	319	40	174
TN	6871	35.5	114	263	302	38	179
KY	6734	36.9	88	265	314	35	176
KY	6792	37.3	78	250	352	32	178
IL	6721	37.7	100	278	323	44	186
IN	6795	38.3	73	271	342	46	183
WV	6778	38.9	55	141	278	65	135
IL	6771	39.0	89	228	361	52	183
CT	6794	41.3	64	173	267	64	142
VT	6782	43.9	52	140	320	83	149
ME	6785	44.9	57	149	237	70	128
MI	6779	45.2	52	158	261	55	132
WI	6723	45.7	62	162	299	80	151
MI	6736	46.6	54	146	261	67	132
Plantation mean			88	221	284	51	
Population mean							161
Plantation Latitude			30.4	37.5	40.0	45.6	

variables. We considered the following three models within each plantation.

Model 1: Linear provenance effect, S_i,

$$X_{ij} = a + b(S_i) + e_i$$

Model 2: Linear and quadratic effect of the latitude of the seed source, Z_i

$$X_{ij} = a + b(Z_i) + c(Z_i)^2 + e_i$$

Model 3: Combination of above two models

$$X_{ij} = a + b(S_i) + c(Z_i) + d(Z_i)^2 + e_i$$

By interchanging the seed source with plantation, the three previous models become:

Model 4: $X_{ij} = a + b(P_j) + e_j$

Model 5: $X_{ij} = a + b(Z_j) + c(Z_j)^2 + e_j$

Table 2.--Observed and expected provenance by plantation interactions in white ash. Expected values were predicted by latitude and provenance effects.

Seed State	Seed Stand	Source Lat.	Source Effect	Observed (Expected) Interaction in Plantation LA	IL	OH	WI
		Deg.	-cm-		cm		
TX	6768	30.3	10	59(52)	48(45)	-86(-74)	-21(-24)
LA	6738	30.5	1	65(53)	21(37)	-77(-76)	-10(-14)
MS	6737	30.8	1	27(49)	30(35)	-39(-71)	-18(-14)
MS	6740	33.4	33	14(7)	68(48)	-42(- 8)	-41(-47)
AL	6733	34.5	-14	15(17)	22(2)	-38(-26)	1(6)
TN	6728	35.3	13	- 9(-1)	10(21)	22(2)	-24(-23)
TN	6871	35.5	18	8(-5)	24(25)	0(7)	-31(-28)
KY	6734	36.9	15	-15(-13)	29(16)	15(19)	-31(-22)
KY	6792	37.3	17	-27(-16)	12(16)	51(24)	-36(-24)
IL	6721	37.7	25	-13(-22)	32(21)	14(32)	-32(-31)
IN	6795	38.3	32	-37(-23)	28(15)	36(35)	-27(-27)
WV	6778	38.9	-26	- 7(-5)	-54(-28)	20(6)	40(27)
IL	6771	39.0	22	-21(-26)	-15(12)	55(39)	-21(-26)
CT	6794	41.3	-19	- 5(-15)	-29(-31)	2(20)	32(25)
VT	6782	43.9	-12	-24(-18)	-69(-35)	48(26)	44(26)
ME	6785	44.9	-33	2(-7)	-39(-56)	-14(10)	52(52)
MI	6779	45.2	-29	- 7(-8)	-34(-53)	6(12)	33(49)
WI	6723	45.7	-10	-16(-15)	-49(-38)	25(23)	39(30)
MI	6736	46.6	-29	- 5(-3)	-46(-57)	6(6)	45(54)
Plantation latitude				30.4	37.5	40.0	45.6
Plantation effect				-73	60	123	-110

Model 6: $X = a+b(P_j) + c(Z_j) + d(Z_j)^2 + e.$

where P_j is the effect of the j th plantation, and Z_j is the latitudinal effect of the j th plantation.

RESULTS AND DISCUSSION

In the Illinois and the Wisconsin plantations the interaction can be explained by the seed source effect (Table 3). The slope of the regression line was positive for the Illinois plantation but was negative for the Wisconsin plantation. The slope was level for both the Louisiana and Ohio plantations. The interpretation is that a fast growing seed source will gain extra growth in Illinois, but will suffer extra loss in Wisconsin. Gain or loss of growth in Louisiana and Ohio are random and unpredictable in terms of the mean performance of the provenances. About 86 and 60 percent of the variation in the interaction were accounted for by the first model in, respectively, the Wisconsin and the Illinois plantations.

The linear and quadratic effect of the seed source latitude (Model 2)

Table 3.--GE interaction (Xij) as explained by provenance effect (Si) and latitudinal effect (Zi) and fitting error (ei)

Plantation	Regression	F	r^2	Error -cm-
	Model 1: With provenance effect, S_i			
LA (j = 1)	$X_{ij} = 0.221 - 0.040\ S_i + e_i$	.0	.00	28
IL (j = 2)	$X_{ij} = -0.960 + 1.448\ S_i + e_i$	25.3	.60	26
OH (j = 3)	$X_{ij} = 0.190 + 0.078\ S_i + e_i$	.0	.00	42
WI (j = 4)	$X_{ij} = 0.076 - 1.487\ S_i + e_i$	113.4	.86	13
	Model 2: With latitude effect, Z_i			
LA (j = 1)	$X_{ij} = 1081.12 - 53.73\ Z_i + 0.654\ Z_i^2 + e_i$	34.6	.81	12
IL (j = 2)	$X_{ij} = -13.72 + 7.05\ Z_i - 0.172\ Z_i^2 + e_i$	21.2	.73	22
OH (j = 3)	$X_{ij} = -1546.91 + 76.88\ Z_i - 0.935\ Z_i^2 + e_i$	22.2	.73	22
WI (j = 4)	$X_{ij} = 479.95 - 30.24\ Z_i + 0.454\ Z_i^2 + e_i$	19.7	.71	19
	Model 3: Provenance and latitude effects, S_i and Z_i			
LA(j=1)	$X_{ij} = 921.62 - 0.441\ S_i - 44.25\ Z_i + 0.518\ Z_i^2 + e_i$	34.1	.87	10
IL(j=2)	$X_{ij} = 292.60 + 0.846\ S_i - 11.15\ Z_i + 0.089\ Z_i^2 + e_i$	23.9	.83	18
OH(j=3)	$X_{ij} = -1300.41 + 0.681\ S_i + 62.23\ Z_i - 0.724\ Z_i^2 + e_i$	19.3	.79	20
WI(j=4)	$X_{ij} = 84.58 - 1.092\ S_i - 6.74\ Z_i + 0.116\ Z_i^2 + e_i$	82.2	.94	9

explained 71 to 81 percent of the GE interaction in the four plantations. All four regression curves were highly significant. The curves for the mid-latitude plantations (Illinois and Ohio) were convex, while the curves for the two extremes (Louisiana and Wisconsin) were concave. The curves for interaction follow closely the curves for height growth over latitude in a previous study (Kung 1981b). Within the latter two plantations, the southern white ash grew better in the Louisiana plantation and the northern trees were better in Wisconsin. In Table .2 we can also see that interaction favors local seed sources in white ash.

Prediction of GE interaction may be further improved by combining both the seed source effect and the latitudinal effect (Model 3). The predicted interactions calculated from Model 3 are listed in Table 2 for comparison with the observed values. The standard error of curve fitting ranged from 9 cm in the Wisconsin plantation to 20 cm in the Ohio plantation. The coefficient of determination for the four plantations averaged 86 percent.

Comparison between Model 1 and Model 3 in Table 3 showed that the direction of the coefficient for the provenance effect, S_i, remained unchanged in the two models. However, when Model 2 and Model 3 were compared, the coefficients for the latitudinal effect in the Illinois plantation were not as consistent as those in the other three plantations.

Which one of the three models is the best? The answer probably depends on the criterion used for evaluation. Based on the F-value, Model 1 is the best for the Illinois and the Wisconsin plantations. Based on the coefficient of determination, Model 3 is superior to both Model 1 and Model 2. The same conclusion is drawn if the standard error of curve fitting is used as the criterion for model selection. Model 3 had the smallest error.

Because of orthogonality, the sum of coefficients for a given dependent variable among the four plantations must be equal to zero. This restraint can be used in checking for computational or rounding errors. For example, the sum of the intercepts in Model 3 is 921.62 + 292.60 - 1300.41 + 84.58 = -1.61, while the sum of the coefficients for seed source effect is -0.441 + 0.846 + 0.681 - 1.092 = -.006. The arithmetic sum is less than one percent of the sum of the absolute values for the four coefficients.

Because there were only four plantations for each provenance, regression in the form of Model 6 cannot be processed for GE interactions, which were within a provenance and across the plantations. To do so would leave no error term for testing the goodness of fit. Although Model 5 can be run with one degree of freedom for the error term, none of the 19 regression curves were significant.

Only three seed sources were significant according to Model 4. They were IN 6795, TN 6728 and KY 6792. On the average, 91 percent of the variance in GE interaction within provenance can be accounted for by the model.

The slopes of the linear regression in the three provenances were all positive. As shown in Table 2, the observed interactions for IN 6795, TN 6728 and KY 6792 followed closely the ups and downs of the plantation effect. These three provenances would grow better on good sites and worse on poor sites. According to Finlay and Wilkinson (1963), these provenances would have below average stability but would specifically adapt to favorable environments.

The three provenances with below average stability were from the central portion of the natural distribution of the species. This finding is contrary to the phenotypic stability of height growth in jack pine provenances (Morgenstern and Teich 1969). The central provenances of jack pine were stable but the northern and southern provenances were unstable.

There were no significant trends in the seed source x plantation interaction for the remaining 16 provenances. The failure to detect trends may be due to the small number of plantations ($n = 4$) used in this study.

How many plantations are needed for a stability study? Or more generally, how many plantations are needed in order for a multiple regression model to show the interaction trend? The answer depends on the number of parameters in the model and the degree of determination of the model. For example, using

Model 4 for the present stability study, we cannot be sure about the linear trend unless more than 90 percent of the sum of squares were accounted for by the seed source effect (Table 4). By the same token, the present study cannot support the quadratic regression based on the latitudinal effect unless more than 99 percent of the interaction can be attributed to it.

Because we have more provenances than plantations, naturally we are more successful in interpreting interaction within a plantation than within a provenance. If we also had 19 plantations, we could obtain more statistically significant regression models once the contribution to the sum of squares due to the seed source effect was above 21 percent in Model 4, or the latitudinal effect was above 31 percent in Model 5, or the combined seed source and latitudinal effect was above 40 percent in Model 6.

Table 4.--Sample size and minimum value for the coefficient of determination needed for a significant regression model

Sample Size	No. of parameters		
	1	2	3
		r^2	
4	.90	.99	-
5	.77	.95	.99
6	.66	.86	.97
7	.57	.78	.90
8	.50	.70	.83
9	.44	.63	.76
10	.40	.58	.70
11	.36	.53	.65
12	.33	.49	.60
13	.31	.45	.56
14	.28	.42	.53
15	.26	.39	.49
16	.25	.37	.47
17	.23	.35	.44
18	.22	.33	.42
19	.21	.31	.40
20	.20	.30	.38
21	.19	.28	.36
22	.18	.27	.35
23	.17	.26	.33
24	.16	.25	.32
25	.16	.24	.30
30	.13	.20	.26
40	.10	.15	.19
60	.06	.10	.13

Table 4 was developed using the relationship that

$$F = \frac{r^2/p}{(1-r^2)/(n-p-1)} \quad \text{---------- (Rao, 1973, p. 273)}$$

Where n is the sample size, p is the number of regressors, F is the F-value, and r^2 is the coefficient of determination. F values significant at the 5 percent level for p and n-p-1 degrees of freedom were used to solve for the r^2. Table 4 can be used for sample size determination for other regression experiments.

RECOMMENDATIONS

1. Superior provenances of white ash should be planted in Illinois so that the extra gain from seed source x plantation interaction can be realized. Average gains are expected in Louisiana and Ohio.
2. In Wisconsin, recommendations for seed sources to plant should not be based on the average performance in the four plantations. Instead, local sources should be used.
3. A combination of provenance effect and latitudinal effect of the seed sources (Model 3) is the best model to describe interaction. However, other environmental variables such as maximum and minimum temperature, moisture availability, soil fertility level of the seed source may be important for characterizing the genotype by environment interaction (Abou-El-Fittouh, Rawlings and Miller 1969).
4. Provenances IN 6795, TN 6728 and KY 6792 have below average stability but are specifically adapted to favorable environments. These provenances should perform well with improved culture methods like site preparation, fertilization and weed control because the selected genotypes will take advantage of such environmental improvement (Owino and Zobel 1977).
5. The number of environments (plantations) needed for a genotypic stability or adaptation analysis should range from 8 to 12 so that a linear or a quadratic model would be significant if more than half of the variation in the GE interaction can be accounted for by the model.

LITERATURE CITED

Abou-El-Fittouh, H. A., J. 0. Rawlings, and P. A. Miller. 1969. Genotype by environment interactions in cotton - their nature and related environmental variables. Crop. Sci. 9:377-381.

Campbell, R. K., and B. C. Wilson. 1973. Space-genotype interaction in Douglas-fir. Silvae Genet. 22:15-20.

Clausen, K. E., F. H. Kung, C. F. Bey, and R. A. Daniels. 1981. Variation in white ash. Silvae Genet. 30:93-97.

Finlay, K. W., and G. N. Wilkinson. 1963. The analysis of adaptation in a plant breeding programme. Aust. J. Agric. Res. 14:742-754.

Kung, F. H. 1981a. Exploratory data analysis for a multiple plantation forest genetics study though a simple mean polish procedure. North Cent. Tree Improv. Conf. Proc. 2:159-167.

Kung, F. H. 1981b. Delineating seed collection zone based on multi-plantation provenance tests. South For. Tree Improv. Conf. Proc. 16:336-342.

Kung, F. H., and K. E. Clausen. 1983. Provenance x environment interactions in white ash. North Cent. Tree Improv. Conf. Proc. 3:228:232

Morgenstern, E. K. and A. J. Teich. 1969. Phenotypic stability of height growth of jack pine provenances. Can. J. Genet. Cytol. 11:110-117.

Owino, F., and B. Zobel. 1977. Genotype x environment interaction and genotypic stability in Loblolly pine. Silvae Genet. 26:21-26.

Rao, C. R. 1973. Linear Statistical Inference and Its Applications. 2nd Ed. Wiley. NY, NY, 625 pp.

SAS 1982. SAS User's Guide: Statistics 1982 Edition. SAS. Cary, North Carolina. 584 pp.

Snyder, E. B. 1972. Glossary for forest tree improvement workers. South For. Exp. Sta., For. Ser. USDA, 22 pp.