United States
Department of
Agriculture

Forest Service

Pacific Northwest
Research Station

Research Paper
PNW-RP-455
December 1992

Genotype * Environment Interaction: A Case Study for Douglas-Fir in Western Oregon

Robert K. Campbell

Author

ROBERT K. CAMPBELL is a principal plant geneticist, Forestry Sciences Laboratory, 3200 S.W. Jefferson Way, Corvallis, OR 97331.

Abstract

Unrecognized genotype × environment interactions (g•e) can bias genetic-gain predictions and models for predicting growth dynamics or species perturbations by global climate change. This study tested six sets of families in 10 plantation sites in a 78-thousand-hectare breeding zone. Plantation differences accounted for 71 percent of sums of squares (15-year heights), replications an additional 4.4 percent, families 1.9 percent, the first principal component of interaction effects 3.5 percent, and the second principal component 1.2 percent. Results in this study and in a larger survey (87 sets in 10 breeding zones) were similar: 51 percent of sets indicated significant g•e. In 46 percent of sets, the g•e interaction-family variance ratio was greater than 1; in 35 percent, greater than 1.5; and in 10 percent, greater than 5.

Summary

If effects of genotype • environment interaction (g•e) contribute substantially to the performance of individual trees or families, considerable research on intrinsic causes may be justified. Ignored interactions may bias genetic-gain predictions and models for predicting growth dynamics or species perturbations by global climate change. Questions about g•e in coast Douglas-fir led to this study, which included six sets of about 30 families each, tested in 10 plantations within one small breeding zone in western Oregon. Families in sets resulted from open-pollination of parents of various and loosely clustered origins. Except for one set, all parents were native to the breeding zone. Data were analyzed by ANOVA, Gauch's AMMI, and Eberhardt-Russell models. The g•e variance ratios ([variance g•e]/[variance family effects]) obtained by ANOVA for height changed from age 5 to 15; ratios became greatly smaller (<1) for two sets and greatly larger (>1) for four sets. The AMMI model analysis of 15-year data suggested a larger contribution of g•e than did the ANOVA model. In the AMMI model, averaged over sets, plantations accounted for 71 percent of sums of squares, replications an additional 4.4 percent, families 1.9 percent, the first principal component of interaction effects 3.5 percent, and the second principal component 1.2 percent. Interaction effects therefore contributed about 2.5 times more variation in plot means than did family effects. As suggested by eigenvector coefficients and Eberhardt-Russell slope coefficients, interactions involved families with greatly different capabilities for growing in plantations of low or high site productivities. A significant fraction of g•e effects was associated with parent tree origin, thereby suggesting an adaptational aspect of interaction. Because this study indicated amounts of g•e larger than previously reported for coast Douglas-fir, sets in other breeding zones were surveyed to place results in a larger context. Results in this study and in the survey (87 sets in 10 breeding zones) were similar: 51 percent of sets indicated significant (P<0.05) g•e. In 46 percent of sets, the g•e variance ratio was greater than 1, in 35 percent greater than 1.5, and in 10 percent greater than 5. Genotype • environment interaction therefore seems to be ubiquitous, variable, and an important aspect of family performance in Douglas-fir in western Oregon. Possible reasons for amounts of interaction and variation among sets are discussed.

Introduction

Genotype · environment interaction (g·e) complicates prediction of individual tree and stand performance, and standard ANOVA analysis procedures often underestimate its importance in this respect. Interactions in forest trees are widely perceived as being relevant only for tree breeding; indeed, forest geneticists and tree breeders have supplied virtually all the current estimates. But tree breeders emphasize genetic gain, and for a number of reasons, genetic-gain calculations often indicate negligible effects of g·e. Consequently, g·e is commonly considered as being equivalent to noise in the data set. In fact, if ignored, interactions may significantly bias models of silvicultural and site-productivity dynamics or of species perturbations by global climate change. If interaction effects contribute substantially to performance, considerable research on intrinsic causes with respect to their importance in tree breeding also may be justified. Needed first is an estimate of g·e's relative importance in predictions of performance in plantations.

Current tree breeding efforts in the Pacific Northwest provide a rare opportunity: genotypes with a variety of known parental origins are planted in several environments, and interactions therefore can be examined for variation among sets and parental origins. The many small Tree Improvement Cooperatives in the Pacific Northwest usually enter the first stage in tree breeding with a test of the wind-pollination progeny of about 300 trees in from 8 to 10 plantation sites in a breeding zone (Silen and Wheat 1979). Approximately the same design is used in each cooperative: 10 or more sets of 25 to 30 families tested in all plantations. By design, sets are evaluated independently of one another, but by chance, the origins of the parents of families in a set may be loosely clustered in many or a few localities within the zone.

This paper presents a case study of g·e interaction in one breeding zone of the Snow Peak Cooperative, one of the many cooperatives (Silen and Wheat 1979) in the region. A yield analysis routinely done at age 15 for all cooperatives indicated substantial variation among sets in the high- and low-elevation zones of the Snow Peak Cooperative. Although the seven sets tested in the high zone had smaller trees and larger error variances than the six sets in the low zone, they also had much larger estimates of additive genetic variance. Family-mean correlations among plantations in the high zone were about twice as large as in the low zone. These results suggested complications due to g·e interaction, especially in the low zone. This degree of variation among sets is not unique among cooperatives. The 15 sets tested in the Mololla Cooperative, which borders Snow Peak on the north, exhibited a range of conditions similar to that found in Snow Peak.

Several questions about interaction in the Snow Peak low-elevation zone led to this study: Was it stable with age? What was its contribution to family performance in predictions at a range of sites? Was interaction a deeply embedded property of coast Douglas-fir *(Pseudotsuga menziesii* (Mirb.) Franco var. *menziesii)* or characteristic of a few families or plantations only? Was it associated with parent origin or mainly a feature of within-population genetic variation?

Table 1—Plantation characteristics

Plantation	Township	Sec.	Range	Elevation	Planting date	15-year mean Height	15-year mean Survival
				Meters		Meters	Percent
209	11 S.	32	01 E.	260	1974	9.02	90.6
210	11 S.	09	01 E.	320	1974	10.45	85.7
212	10 S.	29	01 E.	366	1974	10.37	81.0
214	11 S.	04	01 E.	427	1974	10.04	94.0
215	12 S.	34	01 E.	442	1974	9.56	94.9
216	11 S.	20	02 E.	442	1974	4.33	83.8
309	11 S.	28	01 E.	224	1974	9.65	88.5
318	10 S.	29	03 E.	550	1974	8.71	96.4
518	13 S.	08	02 E.	550	1974	7.74	72.0
520	14 S.	21	01 E.	610	1976	5.25	84.1

Materials and Methods

The Snow Peak Cooperative includes about 78 000 hectares partitioned into two breeding zones, one for high elevation and one for low (Silen and Wheat 1979). Yield trials in both were begun in 1973 to test open-pollination progeny from 180 parent trees in the low zone and 210 in the high zone. Thirty families usually made up a set, each family tested by 120 trees as four individuals per replication per plantation. Seedlings in families were planted randomly (noncontiguous plots) within replications at 3-meter spacing. Tree heights were measured at ages 5, 10, and 15 years.

In the low zone, parents were tested in 6 sets in 10 plantations and in the high zone, in 7 sets in 9 plantations. Each plantation included three replications of a set. Experimental design did not differ between zones. Characteristics of plantations are given in table 1. In the low zone at age 15, the average height across plantations was 8.5 meters and the average survival was 87 percent. Some families nevertheless were not included in analyses because of missing plots or missing information about parent origin, which left a range (27-30) of family numbers in the sets.

Due to numerous constraints, parents and plantations did not provide a representative sampling of the breeding zone. Although all parents came generally from a common region, the sets of parents sampled somewhat different geographic areas (fig. 1). Notably, parents of families in set 6 spread across a much larger area than did the parents in other sets. Plantations tended to sample mainly the eastern one-quarter to one-half of the area sampled by parents (fig. 1).
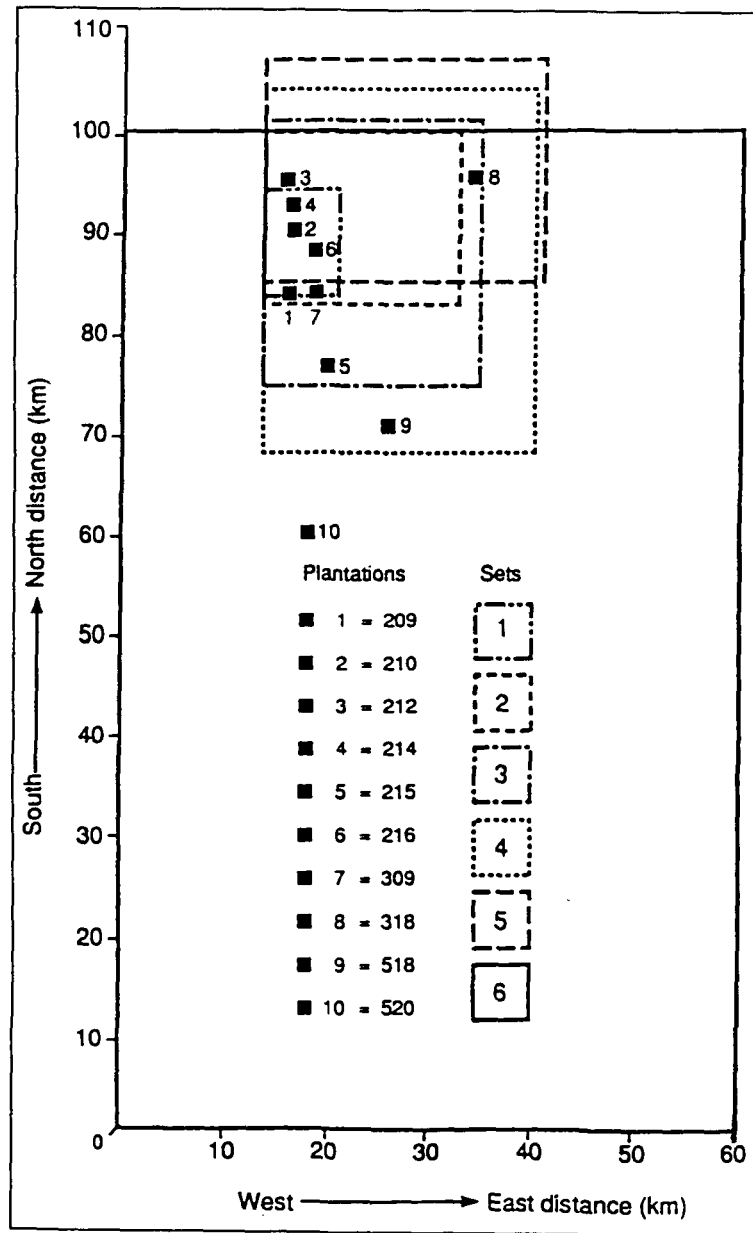
Figure 1—Geographic dimensions of plantation and parent locations. Given for each set is a rectangle conforming to maximum latitudinal and longitudinal extremes of locations of the parents of families tested within a set.

Data were transformed to common logarithms to help equalize variances across plantations and then were analyzed by three models. The first (analysis of variance [ANOVA] model), applied to 5- and 15-year data, is used as an initial analysis in most yield trials:

$$Y_{erg} = u + B_r + E_e + e1_{er} + G_g + (EG)_{eg} + e2_{erg} ,$$

where

Y_{erg} = the yield of family g in replication r in plantation e,

u = the grand mean,

E = the plantation mean deviations,

B = the block mean deviations,

e1 = pooled deviations of replications within plantations,

G = the family mean deviations,

(EG) = interaction deviations, and

e2 = error deviations.

The second model (additive main effects—multiplicative interactions [AMMI]) was used only to analyze the 15-year data. In the AMMI model, interaction effects are partitioned into principal components (Gauch 1988). A limitation of the ANOVA model is that its interaction term conveys none of the response patterns of individual genotypes or environments. All interaction effects are pooled in a g·e sums of squares (SS), often with many degrees of freedom (G-1)(E-1). The importance of a large SS in predicting family performance may be discounted when the mean square (and consequently the g·e component of variance) is small or statistically nonsignificant. Data obtained in several small replicates of a yield trial over many plantation sites can be expected to be especially noisy. Because of this, potentially useful information about the predictable performance of genotypes across plantation sites may be lost. To provide yield estimates closer to true means (and therefore of more value for prediction), patterns must be distinguished from noise. The ANOVA model does not have this capability. Several methods, including the E-R model discussed in the next paragraph, have been developed to partition g·e in attempts to minimize this shortcoming of the first model (Lin and others 1986). The AMMI model was chosen here because it makes fewer assumptions about the structure of interactions than do other models (Gauch 1988). The model is:

$$Y_{erg} = u + B_r + E_e + e1_{er} + G_g + \Sigma \lambda_n a_{gn} b_{en} + R_{ge} + e2_{erg} ,$$

where

λ_n = the eigenvalue of principal component (PC) axis n,

a_{gn} = the family PC vector coefficient for PC axis n,

b_{en} = the plantation PC vector coefficient for PC axis n,

n = the number of axes retained in the model,

R_{eg} = the residual interaction effect, and other terms are as in the first model.

The degrees of freedom assigned to PC axis n is (G + E − 1 − 2n) (Gauch 1988).

4

The third model (the well-known Eberhardt-Russell [E-R] [1966] modification of the Finlay-Wilkinson model [1963]) was used on the 15-year data expressly to provide estimates of two parameters, which then could be compared with results from the AMMI model. The purpose of the model was to regress each family in the test on an environmental index to provide slope and deviation parameters described in the following:

$$Y_{eg} = U_g + S_g I_e + D_{eg} ,$$

where

Y_{eg} = the family mean of the gth family in the eth plantation,

U_g = the mean of the gth family over all plantations,

I_e = an environmental index obtained as a mean of all families within a set at the eth plantation minus the grand mean,

S_g = the regression coefficient (slope) measuring the response of the gth family over the varying environmental indexes, and

D_{eg} = the deviation from regression of the gth family at the eth plantation.

For breeding endeavors, rank is an important attribute of a family, perhaps as important as its contribution to family and interaction variance. Ranking may or may not be affected by interaction, or ranking may be affected in only some portion of the families. Two common questions exist about the relation of rank to interaction: Do rankings change appreciably from plantation to plantation in highest ranked families even in sets of families exhibiting large amounts of interaction? If so, does this difference in variablity of plantation ranks reflect variability in degrees of genotype × environment interaction in upper and lower ranks? These questions were examined by analyzing variation among the plantation ranks of families in 15-year average height growth over a range of expected ranks. Plantation rank is defined here as the ranking of a family within a set tested in a plantation. Expected rank is rank of the family as determined from the average performance of the family over all 10 plantations. The first null hypothesis tested was as follows: variation among the 10 plantation ranks of a family did not differ regardless of the expected rank of the family. The second question led to a second null hypothesis: for families with given expected ranks, plantation ranks do not differ among plantations. The first hypothesis was tested by regression of the standard deviation of plantation ranks on expected ranks. The test was restricted to the top third of expected ranks in each set, because genetic selections usually are restricted to this upper group. The hypothesis would be disproved if a relation existed between expected rank and variability in plantation ranks. The second null hypothesis was tested by first grouping families of given expected ranks into classes (1-5, 6-10, 26-30, for example), then averaging plantation ranks across expected ranks. With each set acting as a replication, Friedman's nonparametric test of randomized blocks (Sokal and Rohlf 1969) was used to test variation in ranking among plantations. In the absence of rank-change interactions, plantation ranks should not differ.

For analysis of the association of g·e effects with parent-tree location, families within sets were clustered geographically (latitude and departure) by the SAS (1987) procedure, CLUSTER.

Results

As measured by the relative contributions to plot variance of average family effects (σ_g^2) and g*e interaction effects (σ_{ge}^2), response to environment was a highly variable property of Douglas-fir in breeding zones of the Snow Peak Cooperative. Estimates of variance depended strongly on the combination of parents and plantations involved in tests. Identically designed tests similarly installed (same year and crews) in high- and low-elevation zones, for example, exhibited greatly different expressions of genetic variation in 15-year heights in the two zones. In six of seven sets (~30 families per set) in the high zone, variances among families were higher than in any of the six sets in the low-zone test (fig. 2). The probability of this occurring by chance is 0.004. Though trees in high-zone sets were shorter (by 4 percent) and test error variances were larger (by 23 percent) than in the low-zone sets, interaction contributed considerably more to total genetic effects ($\sigma_g^2 + \sigma_{ge}^2$) in low-zone sets. The experimental error was small in either case: the average coefficient of variation,

$$CV = (\text{error standard deviation} \times 100)/\text{mean log height}),$$

was 2.4 percent for the low zone and 2.8 percent for the high zone. Low-zone sets (Snow Peak low) also differed greatly from high-zone sets (Snow Peak high) in the ratios of interaction variance to family variance (σ_{ge}^2/σ_g^2). In low-zone sets, the ratio varied from much less than one to much more than one. In high-zone sets, it never exceeded one. This variation in genetic expression among low-zone sets is not unusual, however. It represents the range of genetic effects seen in many cooperatives in the Pacific Northwest. Further analyses were restricted to these six low-zone sets.

The large σ_{ge}^2/σ_g^2 ratios found in some low-zone sets implied substantial changes in family rank across plantations. The rank changes, however, were not necessarily characteristic of all families. To determine if ranks change appreciably in families with high expected ranks, two null hypotheses were tested. The first, that variation among plantation ranks did not differ regardless of expected rank of a family, was disproved in within-set tests. Regressions of standard deviation of plantation ranks on expected ranks were not significant within sets. But, when data from all six sets were combined and analyzed, variability in plantation ranks was found to increase by 0.20 units of standard deviation for each unit decrease in expected rank from 1 to 10 (P=0.009, R^2=0.11). Highest ranked families therefore showed more stability in plantation rank than did families of lower expected rank. The second null hypothesis (that for families with given expected ranks, plantation ranks do not differ) was disproved for families with high and low expected ranks. In the absence of rank-change interactions, plantations should not differ, beyond error, from the average for all plantations. In fact, the five families of highest rank (mean=3) consistently deviated from expected ranks (P=0.004). They ranked higher than the average in some plantations (table 2; for example, plantations 214 and 216) and lower than the average in others (518 and 520). The five families of lowest expected rank (mean=28) ranked higher than the average in plantations 518 and 520 and lower than the average in 209 and 318 (P=0.003). For intermediate expected ranks, the null hypothesis was not disproved (P=0.76). Families with intermediate rankings, from 6-10 (mean=8) for example (table 2), did not show consistent ranking differences among plantations. Rank changes were substantial even among families with the high average expected rank of three. In some plantations, ranks were almost twice as large as in other plantations (table 2). This occurred even though three of the sets (1, 2, and 4) included in the average plantation ranks exhibited no apparent g·e interaction by the ANOVA model (fig. 2, table 3).
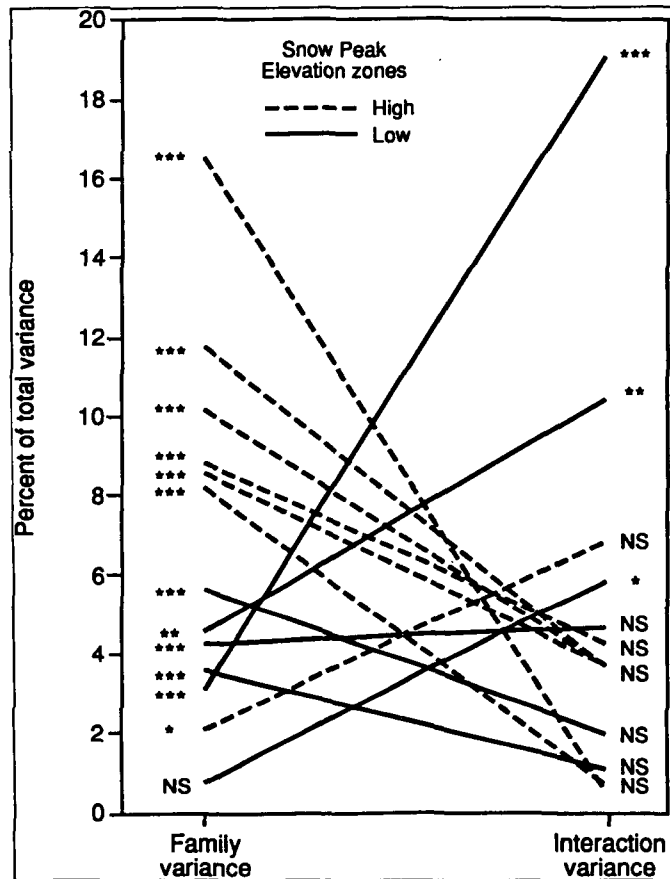
Figure 2—Distribution of genetic variances within the 13 sets from Snow Peak high- and low-elevation breeding zones. Total variance is the sum of variances among families, family • plantation interaction (g•e) effects, and experimental error. Probabilities that the effects creating the variances are due to chance are NS=P>0.05, *=P<0.05>0.01, **=P<0.01>0.001, ***=P<0.001.

Relative values of family and interaction components of genetic variance were not stable over time. From age 5 to age 15, the variance structure, as estimated by the ANOVA model, appeared to change radically in some sets (fig. 3). At age 5, significance tests suggested strong genetic differentiation among families in all sets and interaction in four of the sets. At age 15, the relative importance of g•e had decreased in sets 1 and 2, but the decrease in g•e had been accompanied by a larger relative decline in family variance in set 2 than in set 1. In sets 3, 5, and 6, g•e variance increased dramatically from age 5 to age 15, apparently at the expense of family variance.

The variance trends shown in figure 3 imply that interaction in heights at age 15 accounted for a small fraction of genetic variation in three of the six sets. A very different picture emerged when g•e effects were partitioned into principal components in the AMMI model (table 3). Effects associated with the first component (PC1), as measured by mean squares, were larger than average family effects in five of six sets. Although effects attributed to PC2 were statistically significant for five sets, they were larger than family effects only for set 5.

Table 2—Average actual family ranks within sets and plantations for families of given expected rank (rank based on 15-year heights as averaged over all plantations) for 3 classes of expected rank, highest, intermediate, and lowest

Expected rank	Set	Plantation[a]										
		210	212	214	309	215	209	318	519	520	216	x̄
3.0	1	7.4	7.0	5.8	7.2	7.4	9.8	10.0	10.2	5.8	6.0	7.7
	2	6.8	5.2	11.0	10.6	9.6	8.2	6.8	12.2	13.8	8.8	9.3
	3	6.6	11.6	10.0	11.2	8.6	8.4	10.0	12.4	11.4	7.6	9.8
	4	4.6	8.4	4.2	7.0	10.4	6.4	13.2	14.6	11.2	7.4	8.7
	5	9.8	11.6	5.2	7.6	10.6	12.4	12.6	16.4	17.4	6.2	11.0
	6	8.6	13.8	5.2	9.6	8.4	6.0	5.8	13.8	17.8	10.8	10.0
	x̄	7.3	9.6	6.9	8.9	9.2	8.5	9.7	13.3	12.9	7.8	9.4
8.0	1	17.0	13.2	9.0	6.0	14.6	13.8	7.6	9.0	21.2	12.4	12.4
	2	11.2	7.6	13.4	15.0	12.0	10.6	13.2	19.2	14.6	7.6	12.4
	3	10.2	9.0	11.2	12.4	14.0	14.4	10.2	10.4	11.6	10.0	11.3
	4	16.8	11.0	12.6	13.2	11.6	13.8	8.4	13.0	5.6	16.4	12.2
	5	10.0	14.4	15.4	13.4	7.8	10.4	9.0	15.4	15.2	18.2	12.9
	6	15.6	15.6	13.0	17.2	9.2	13.8	17.8	16.0	15.0	10.6	14.4
	x̄	13.5	11.8	12.4	12.9	11.5	12.8	11.0	13.8	13.9	12.5	12.6
28.0	1	25.6	23.8	22.8	22.8	16.4	23.6	24.0	18.8	16.2	18.8	21.3
	2	23.2	21.6	26.8	19.6	26.6	28.0	20.6	17.8	20.0	23.4	22.8
	3	21.8	24.2	22.6	24.0	21.4	26.2	25.8	24.6	16.8	20.6	22.9
	4	25.2	28.0	22.4	22.2	25.2	25.6	27.0	11.6	17.0	16.8	22.1
	5	21.2	15.6	20.0	24.2	23.0	21.6	24.6	21.0	16.4	20.0	20.8
	6	21.8	19.8	24.6	22.8	22.8	23.2	25.4	·19.6	20.6	18.2	21.9
	x̄	23.1	22.2	23.2	22.6	22.6	24.7	24.6	18.9	17.8	19.6	21.9

[a] Plantations are ordered by productivity as measured by average 15-year height, from highest (left) to lowest (right).

The above reports the expression of genetic variation in terms of mean squares and variance components. Variance components, especially, are part of the vocabulary of the classification model for describing variation in experiments. Tree breeders rely on estimates of components for help in evaluating alternative breeding designs. Results from regression models provide, however, a more useful basis for judging potential effects of interaction in respect to growth modeling. In regression models, sum squares can be partitioned so their relative importance in predicting performance at plantation sites can be judged. In this paragraph, the importance of the various design variables is reported in a regression context. In orthogonal designs, the fraction of the sums of squares accounted for by a statistically significant variable provides an estimate of the relative importance of the variable in describing variation

Table 3—Mean squares and degrees of freedom (in parentheses) associated with sets of families, main effects (plantations, replications, and families), and with interactions, including principal components (PC1-PC3) of the family × plantation interaction

Source variation	Sets					
	1	2	3	4	5	6
Plantations (E)	(9)2.1806***	(9)1.5183***	(9)1.2742***	(9)1.4722***	(9)1.5516***	(9)2.4606***
Reps in E	(20).0197***	(20).0420***	(20).0246***	(20).0359***	(20).0725	(20).0512***
Families (G)	(28).0145*	(28).0147	(28).0117**	(28).0149	(29).0074*	(26).0132***
G x E:	(252).0053ns***	(252).0070ns	(252).0056***	(252).0067ns	(261).0060**	(234).0069***
PC1—	(36).0144*	(36).0186	(36).0220	(36).0237	(37).0133	(34).0191***
Clusters (CL)[a]	(12).0238**	(16).0140ns	(12).0203ns	(13).0247ns	(10).0075***	(15).0397
Within CL	(24).0098	(20).0223*	(25).0219***	(23).0231*	(27).0155**	(19).0029ns
PC2—	(34).0064ns	(34).0127**	(34).0077*	(34).0097***	(35).0115	(32).0104
CL	(12).0056ns	(16).0192	(12).0107ns	(13).0231	(10).0005ns	(15).0060ns
Within CL	(22).0063ns	(18).0063ns	(23).0054ns	(21).0013ns	(25).0147	(17).0128***
PC3—		(32).0091ns	(32).0033ns	(32).0160ns	(33).0071ns	(30).0087
Residual 1[b]	(216).0038ns	(216).0051ns	(216).0029ns	(216).0039ns	(224).0048ns	(200).0048ns
Residual 2[c]	(182).0033ns	(182).0037ns	(182).0020ns	(182).0028ns	(189).0036ns	(168).0039ns
Residual 3						(138).0027ns
Error	(560).0050	(560).0067	(560).0041	(560).0058	(580).0051	(520).0040

ns = P>0.05.
* = P<0.05>0.01.
** = 0.01>P>0.001.
*** = P<0.001.

[a] Variation among parent-tree clusters tested by variance within clusters. Variation within clusters tested by error.
[b] Residual mean squares of interaction effects after removing PC1.
[c] Residual mean squares of interaction effects after removing PC2.

among plot means. In this context, R^2 (the coefficient of determination = sum squares for effects/total sum squares) measures the proportion of the total sums of squares "explained" by design variables and their interactions. Averaged over all sets, plantations accounted for 71 percent ($R^2=0.71$) of sums of squares (fig. 4). Replications within plantations contributed an additional 4.4 percent, families 1.9 percent, PC1 3.5 percent, and PC2 1.2 percent. Set 1 was not included in the calculation of the average fraction added by PC2. Genetic effects (including g·e) explained about 9 percent as much variation as did plantations and about 1.5 times more than did replications. Note that g·e effects accounted for almost 2.5 times more variation than did average family effects. Even in sets 1 and 2, in which interactions were statistically nonsignificant in ANOVA analyses, the contributions of PC1 and PC2 were as large as that of families.
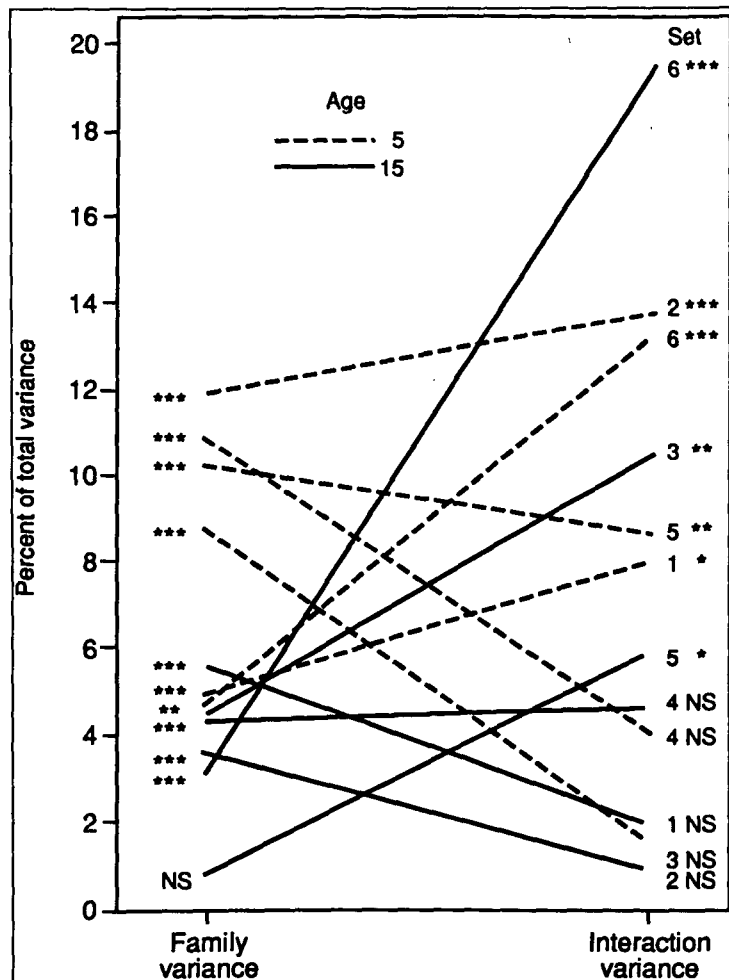
Figure 3—Distribution of genetic variance, at two ages, within the six
sets for the Snow Peak low-elevation zone. Total variance is the sum
of variances among families, family • plantation interaction (g•e) effects,
and experimental error. Probabilities that the variation in effects creating
the variances is due to chance are NS=P>0.05, *=P<0.05>0.01,
**=P<0.01>0.001, ***=P<0.001.

The estimated importance of interaction differed from set to set; because families
differed from set to set but all were tested in the same plantations, families ac-
counted for most of that variation. Plantations contributed too, though, because
interaction is a multiplicative response of genotypes to environments. Correlations
among family means (not shown) at the 10 plantations indicated more stability of
family performance in some plantation pairs than in others. The more productive
sites tended to group together (higher correlations) as did the less productive ones.
The three most severe plantation sites (216, 518, 520), as indicated by shorter
average tree heights and lower survival (table 1), were most strongly involved in the
interactions. For these plantations, eigenvector coefficients usually were larger, and
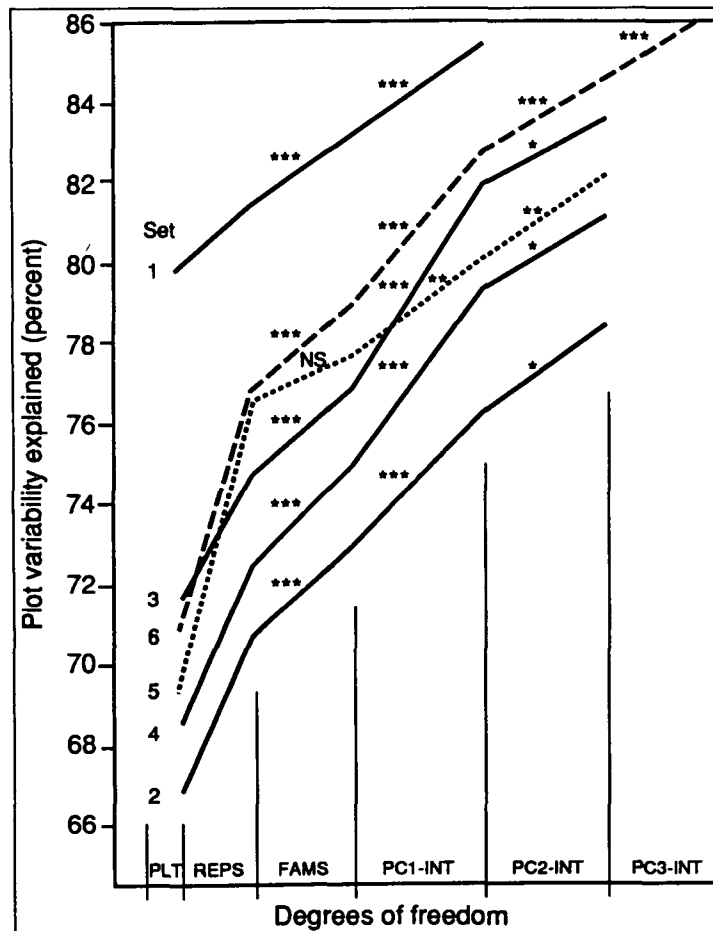
Figure 4—For each set, the percentage of total variability (total sums of squares) in plot means explained by plantation (PLT), replications within plantations (REPS), families (FAMS), and the first (PC1-INT), second (PC2-INT), and third (PC3-INT) principal components of family • plantation (g•e) interaction. Dimension on the abscissa is measured in degrees of freedom associated with each design category; the steeper the line, the larger the mean square for the effect. The probabilities that effects are due to chance are NS=P>0.05, *=P<0.05>0.01, **=P<0.01>0.001, ***=P<0.001. Effects of plantations and replications are all significant at P<0.001.

within sets, usually were of opposite sign than for the faster growing plantations (table 4). Although these three more severe plantations produced trees of about the same average heights, families in different sets did not react identically to the three plantation environments. The plantation contributions to interaction appeared to be quite different depending on the sets involved. When plantation coefficients for PC1 and PC2 are compared within sets, the pattern of differences in sign and magnitude is seen to range greatly among sets (table 4).

Table 4—Eigenvector coefficients of the 1st (PC1) and 2d (PC2) principal components of the within-set genotype.* environment interaction effects associated with plantations

	Plantation																			
	209		210		212		214		215		216		309		318		518		520	
Set	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
1	0.10	—	0.21	—	0.24	—	0.09	—	0.02	—	-0.45	—	0.21	—	0.26	—	0.05	—	-0.75	—
2	-.13	-0.02	-.07	0.07	-.11	0.76	-.08	-0.12	-.08	0.05	-0.26	-0.59	-.04	0.01	-.12	-0.15	-.04	0.08	.93	-0.10
3	.19	-.13	.10	-.05	.10	-.21	.18	.02	.12	-.01	-.17	.88	.17	-.18	.17	.14	.04	-.22	-.90	-.24
4	.26	.29	.18	.22	.19	.11	.13	.06	.17	.06	.05	-.12	.06	-.14	.05	-.02	-.87	.36	-.22	-.83
5	.14	-.29	.01	-.19	-.31	-.11	.04	-.09	.07	-.03	.06	-.32	.09	.02	-.16	-.17	.67	.83	-.62	.62
6	-.25	.11	-.16	-.01	-.43	.03	-.04	-.02	-.05	-.05	.29	.65	-.18	.09	-.06	-.01	.12	-.73	.76	-.06

— = not calculated because mean squares for this factor was not statistically significant ($P > 0.05$).

Families also contributed variously to interaction effects; eigenvector coefficients (PC1 and PC2) of the four most reactive families in each set are given in table 5. A disadvantage of the AMMI model is that it does not provide readily interpretable family means. Fortunately, in several sets, the family eigenvector coefficients apparently measured about the same features of interaction as are described by the E-R slope coefficients (table 5). An E-R slope above 1 denotes a family that performs relatively better in the more productive plantations than in the less productive ones, when compared to the performance of other families in the set. In sets 3 and 6, larger eigenvector coefficients were associated with larger slope coefficients. In sets 1 and 2, on the other hand, negative eigenvector coefficients were associated with the larger E-R slope coefficients. The latter relation demonstrates the fact that the sign of an eigenvector coefficient is not inherently associated with the size of the slope coefficient. The E-R slopes therefore cannot be used as substitutes for eigenvector coefficients in performance evaluation.

Eigenvector coefficients in tables 4 and 5 document the presence of plantations and families that strongly participated in interactions. These very reactive entities seemed to account for most of the interaction in sets 1, 2, and 3, but interaction appeared to be more deeply imbedded in other sets. When components of variance were estimated from the original data set by ANOVA, the g·e component was larger than the family component in sets 3, 4, 5, and 6 (compare figs. 5 and 6). After removing the most reactive plantation (for example, 520 in set 1, table 4), or the two most reactive families (for example, 225 and 201 in set 1, table 5), the g·e component was no longer larger than the family component in set 3. When the two most reactive plantations or the four most reactive families (two each from PC1 and PC2) were removed, g·e variance was larger than family variance in only two sets. Even after removing four families, which resulted in negative ANOVA estimates of g·e components of variance for sets 1, 2, and 3, the AMMI analysis suggested a remaining component of interaction within the sets. Mean squares for PC1 were still significant and the proportions of the sum squares explained by PC1 were still almost as large as those explained by family effects (fig. 7).

Table 5—Eigenvector coefficients of 1st and 2d principal components (PC1, PC2) and Eberhardt and Russell (1966) slope and deviation coefficients for reactive families in genotype-environment interaction[a]

| Set	Family	PC1 Vector coefficent	PC1 Slope	PC1 Deviation	PC2 Vector coefficient	PC2 Slope	PC2 Deviation
1	225	0.419	0.65(28)	0.0006(24)			
	211	.277	1.12(27)	.0033(2)			
	206	−.219	1.16(5)	.0001(17)			
	201	−.426	1.43(1)	.0017(7)			
2	240	.257	1.17(4)	.0046(3)	−0.216	1.17(4)	0.0046(3)
	239	.256	1.05(10)	.0029(4)			
	257	−.266	1.29(2)	.0012(16)			
	248	−.689	1.25(3)	.0122(1)			
	261				.816	.640(29)	.0114(2)
	233				.197	.887(24)	.0011(17)
	252				−.311	1.449(1)	.0011(19)
3	296	.882	1.87(1)	.0133(1)			
	267	.227	1.14(4)	.0026(2)			
	276	−.108	.99(12)	.0012(8)			
	269	−.193	.72(29)	.0069(17)			
	283				.264	.828(27)	.0005(25)
	291				.250	.779(28)	.0003(27)
	264				−.347	1.185(3)	.0011(9)
	295				−.590	1.326(2)	.0022(3)
4	308	.237	.72(28)	.0021(9)			
	331	.212	.94(16)	.0027(6)			
	313	−.434	1.15(7)	.0086(2)	.347	1.15(7)	.0086(2)
	322	−.469	1.22(3)	.0086(1)	.319	1.22(3)	.0086(1)
	300				−.280	1.06(11)	.0019(11)
	299				−.363	1.39(1)	.0044(3)
5	409	.448	1.01(14)	.0049(1)			
	346	.338	1.13(4)	.0038(3)			
	401	−.297	1.14(3)	.0036(4)			
	352	−.335	1.06(8)	.0030(7)			
	344				.382	1.04(10)	.0045(2)
	351				.246	.98(17)	.0023(9)
	406				−.316	1.01(15)	.0032(5)
	334				−.353	1.15(2)	.0030(6)
6	517	.567	1.42(1)	.0078(1)			
	504	.285	1.37(2)	.0005(26)			
	416	−.317	.72(29)	.0038(2)			
	511	−.456	.48(30)	.0024(8)			
	413				.417	.89(24)	.0033(4)
	503				.416	.92(29)	.0032(2)
	509				−.208	.98(15)	.0011(19)
	411				−.363	1.24(5)	.0019(9)

[a] Rankings within sets are in parentheses.
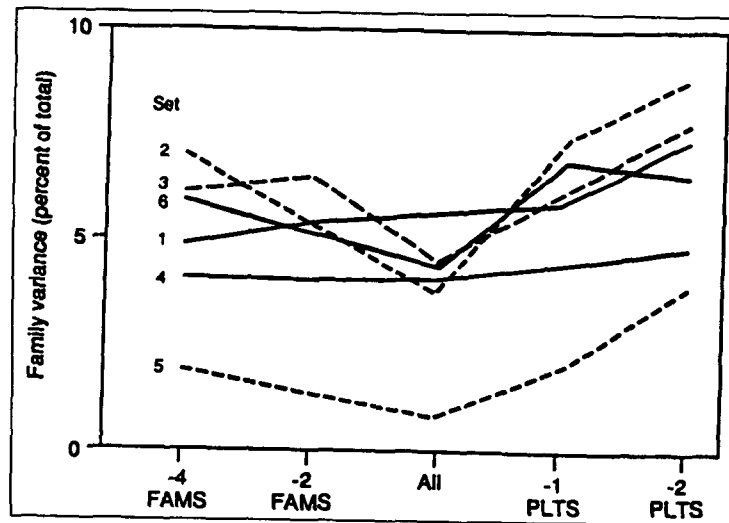
Figure 5—Estimated variances among families within sets as they are affected by removing the most reactive (in interactions) two to four families (FAMS) or one to two plantations (PLTS) from the data set before analysis. "ALL" designates estimated variances in the complete data set. Total variance includes family variance, plantation · family interaction (g·e), and experimental error.

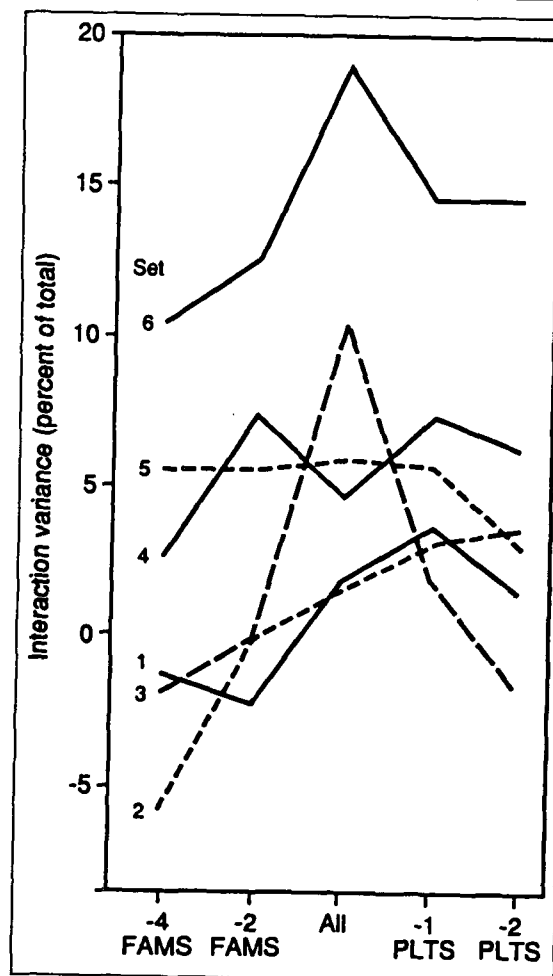Figure 6—Estimated variances of family · plantation interaction effects (g·e) within sets as they are affected by removing the most reactive (in interactions) two to four families (FAMS) or one to two plantations (PLTS) from the data set before analysis. "ALL" designates estimated variances in the complete data set. Total variance includes family variance, plantation · family interaction (g·e), and experimental error.
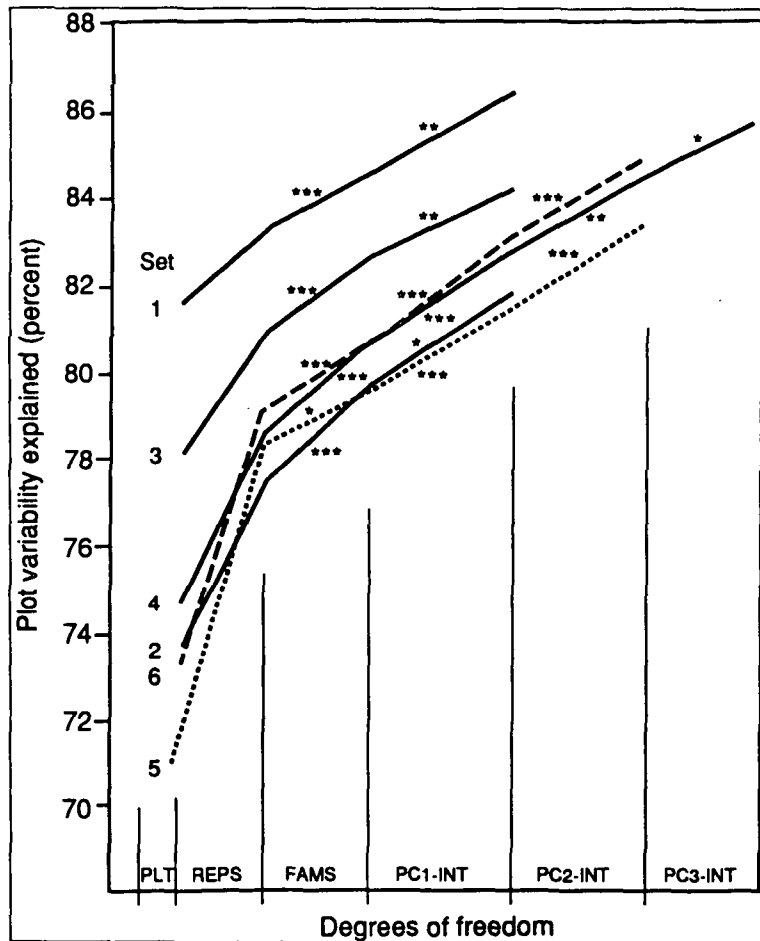
14

Figure 7—For each set, the analyzed data did not include plot means for the four most reactive families (in interactions) in each set. Given are the percentages of total variability in plot means explained by plantation (PLT), replications within plantations (REPS), families (FAMS), and the first (PC1-INT), second (PC2-INT), and third (PC3-INT) principal components of family · plantation (g·e) interaction. Dimension on the abscissa is measured in degrees of freedom associated with each design category. Thus, the steeper the line, the larger the mean square for the effect. The probabilities that effects are due to chance are NS=P>0.05, *=P<0.05>0.01, **=P<0.01>0.001, ***=P<0.001. Effects of plantations and replications are all significant at P<0.001.

Table 6—Percentage of family mean height and interaction variances (slope, deviations, and PCs) accounted for by clusters of parent trees within sets[a]

| Set | Numbers of clusters | Mean family height | Eberhardt-Russell | | Principal components | |
			Slope	Deviations	PC1	PC2
1	5	0(0.68)	18.8(0.09)	0.2(0.42)	(0.03)	(0.57)
2	9	9.2(.30)	0(.48)	0(.91)	(.82)	(.01)
3	5	15.6(.12)	27.2(.03)	12.2(.17)	(.54)	(.08)
4	6	4.6(.34)	26.9(.05)	0(.81)	(.43)	(.00)
5	3	0(.67)	15.9(.09)	.1(.38)	(.89)	(.99)
6	8	17.8(.11)	34.7(.04)	62.1(.00)	(.00)	(.95)

[a] Cluster mean squares were tested against within-cluster mean squares. Probabilities of obtaining cluster mean squares by chance are given in parentheses.

A significant fraction of the g·e interaction was associated with parent tree origin. Parent trees had been chosen haphazardly by location in loosely ordered clusters. Clustering procedures grouped families within sets into three to nine clusters based on the geographic proximity (latitude and departure) of the parents (table 6). In sets 1 and 6 (PC1) and in sets 2 and 4 (PC2), interaction effects were significantly larger among clusters than within clusters (table 3). When cluster mean squares were significant, within-cluster mean squares usually were not and vice versa. Eberhardt-Russell partitioning of interaction effects also suggested an association of interaction with clustering. About 20 percent of the variance among families in E-R slopes was contributed by clustering (table 6). Although there seemed to be an association of slope and eigenvector coefficients in table 5, slopes and PC interaction effects were not interchangeably associated with parent clusters. In sets 3 and 5, for example, PC effects apparently did not cluster by parent origin (table 3), whereas E-R slopes apparently did (table 6). In set 6, clustering accounted for 62 percent of the variance in deviations from the regression lines (table 6). In other sets, deviations were no more closely associated with clusters than were mean family heights (table 6), in which clustering effects approached statistical significance only in sets 3 and 6. Parent origin apparently influenced interactions (as expressed in PCs, slopes, and deviations) more than it did the family average heights.

Discussion

This study reports the occurrence of appreciable amounts of interaction within one breeding zone, more than has been indicated in previous reports for coast Douglas-fir (Adams and Joyce 1990, Kitzmiller 1990, Stonecypher 1990). This raises the question whether Snow Peak sets fairly represented conditions in the Pacific Northwest. Fortunately, data are available that can be used to place results within a larger context. Analyses based on the ANOVA model exist (15-year data) for 10 breeding zones (including Snow Peak, high and low) in western Oregon.[1] Eight of the zones cover most of the lands from lat. 43°20' N. to 45°20' N. along the eastern slopes

[1] Unpublished data. Pacific Northwest Tree Improvement Cooperatives. On file with Forestry Sciences Laboratory, 3200 S.W. Jefferson Way, Corvallis, Oregon 97331.

of the Coast Range. The other two zones extend from lat. 44°30' N. to 45°30' N. along the west side of the Cascade Range. The zones range in size from about 30 to 80 thousand hectares. Progeny tests in the zones include 87 sets, usually of 28 to 30 families, each tested in from 5 to 10 plantations. Fifty-one percent of these sets exhibited significant (P<0.05) levels of g·e. In 46 percent of the sets, the g·e component of variance was larger than the family component; that is, the ratio σ_{ge}^2/σ_g^2 (the g·e ratio) was greater than 1. In 35 percent of the sets, the ratio was greater than 1.5; in 21 percent, greater than 2; and in 10 percent, greater than 5. Interaction seemed to be more prevalent in the Coast Range than in the Cascades; in the Coast Range, the ratio was greater than 1 in 51 percent of sets and in the Cascade Range in 36 percent of sets. Of the six sets making up the sample for the present study, the ANOVA model indicated 50 percent had significant g·e (P<0.05) and 50 percent had a σ_{ge}^2/σ_g^2 ratio greater than 1, and 33 percent a ratio greater than 5. Sets 5 and 6, with ratios greater than 5, therefore over-represent the proportion of sets with very large ratios in the larger sample. In other respects, Snow Peak low seems to adequately represent conditions adjacent to the Willamette Valley in central and northern parts of western Oregon.

The interaction in Snow Peak seemed to reflect mainly a differential response of families to plantations of differing productivity. The interactions indicated not only variation in relative performance among families, but also changes in rank. The latter manifested itself in larger g·e ratios and in lower family mean correlations among plantations in Snow Peak low compared with Snow Peak high. Taller families in more productive plantations often performed less well in less productive plantations. Furthermore, families from parents with similar origin tended to exhibit similar interaction effects. These results jointly suggest that families from some locations cannot flourish in conditions leading to low site productivity, whereas others do well in that situation. This was not consistently the case, however. Families with highest expected rankings did poorly in plantations 518 and 520 (table 2); but the same families ranked above average in plantation 216, which was even less productive than 518 and 520. On the other hand, families with the lowest expected ranks performed better than the average at all poorest sites (table 2, plantations 518, 520, and 216). Families with intermediate expected ranks performed inconsistently with respect to plantations, though there can be little doubt that these families also experienced rank changes; intermediate families exhibited more variability in plantation rank than did families at extreme expected ranks. Intermediate families may be better fitted for growth in the less extreme plantation sites. Or they may be in a period of transition and at age 15 are still undergoing sorting into appropriate ranks. The changes in g·e with age as seen in figure 3 suggest that this is a likely possibility.

The relatively large amounts of g·e within breeding zones may simply reflect variable environments and variation among Douglas-fir populations within the zones. In western Oregon, the major determinants of site productivity for Douglas-fir are soil depth and drainage, and cold- or drought-conditioned growing-season length. Many local topographical factors, such as slope, cold-air drainage, or aspect, may achieve importance because they influence temperature and available moisture at critical times. Rain shadow is one such local factor that figures prominently in the distribution of precipitation in the region (Froehlich and others 1982). Climate in western Oregon is generally Mediterranean and therefore moisture limiting; vegetation and site productivity closely reflect variation in annual precipitation (Franklin and Dyrness 1973). Snow and rain arrive in storms from the Pacific Ocean, and amounts are affected by prominent ridges and mountains. Highest annual precipitations are found on highest mountains and lowest precipitations in adjacent interior valleys (Froehlich and others 1982). It therefore is not surprising that plantations differed considerably in untransformed heights at age 15 even within breeding zones; within the surveyed zones, shortest plantations averaged only 57 percent of heights in the tallest plantations. Furthermore, previous studies have shown strong gradients of geographic genetic variation in Douglas-fir that is associated with rain shadows (Campbell 1986; Griffin 1978; Sorensen 1979, 1983). Interactions may be a reflection of variable population responses in the particularly variable environments associated with rain shadows. This would account for the greater prevalence of interaction in breeding zones in the lee of the Coast Range.

Genotype · environment interaction in amounts indicated by this study could represent a serious source of error in models developed to predict growth in plantations or in native stands in a changing environment. In Snow Peak low, statistically significant g·e effects accounted for 19 percent of that proportion of variation among plot means not explained by plantation differences. These effects represent potential error in any global-climate model that assumes a consistent response in Douglas-fir to environmental change. The error could be larger or smaller depending on the degree of environmental response and the genotypes involved. The interaction seen here also should preclude using seed-orchard seed in breeding zones outside the ones in which parents have been tested. Interactions of considerable magnitude occur within existing breeding zones and intuitively we might expect family performance to be even less predictable in larger zones; evidence exists to suggest that genotype is a very large component of site productivity within a regional setting (Monserud and Rehfeldt 1990). The main reason for advising caution stems, however, from questions concerning an understanding of the measurement, dynamics, and causes of g·e in Douglas-fir.

The measurement and interpretation of g·e effects have received sustained attention over the years (Allard and Bradshaw 1964, Eberhardt and Russell 1966, Finlay and Wilkinson 1963, Freeman 1973). Until very recently efforts have been concentrated on the within-trial accuracy of a model's fit to its own data: the measure of "post-dictive" success as contrasted to "predictive" success (Gauch 1988). The latter measures the fit between a model developed from part of the data and a separate validation set of data. It is not clear whether models recommended by statistical tests for Snow Peak sets would be predictively accurate for other trials of the same sets even in the same breeding zones. Gauch (1988) employed a data-splitting validation procedure to evaluate a soybean yield trial and found that only the first PC axis improves prediction. In his case, the first PC accounted for 16 percent of the plot

variation exclusive of plantation effects. Including PC1 in the model reduced the sums of squared differences between model building and validation data by 12 percent. Because plot error precludes perfect prediction of validation data, PC1 undoubtedly explained more than 12 percent of the sums of squared "predictable" differences. The model that included the first PC predicted validation data better than did all the data used to construct the model. Noise in the data set obviously obscured true yield patterns existing within the experiment. One test is not, of course, sufficient basis for recommending inclusion of only one, or of even one, significant PC in a prediction model. The test does point out, however, that best postdictive models are not necessarily the best predictive models. The optimum structure of predictive models for Douglas-fir obviously bears investigation.

Another source of uncertainty over our ability to predict growth performance in Douglas-fir arises from questions about the consistency of interaction effects. In Snow Peak low, the sets differed greatly in estimated amounts of g·e, as did sets in the larger survey sample. And removing reactive families or plantations from the Snow Peak data sets affected the estimated relative amounts of g·e variously in the different sets (figs. 5-7). Furthermore, sets showed two entirely different dynamics of g·e expression with tree age; in two sets, the σ_{ge}^2/σ_g^2 ratio decreased dramatically with increasing age, but in the other four sets, the ratio increased just as dramatically.

The variability of g·e estimates in space and time reported here and previously (Adams and Joyce 1990, Kitzmiller 1990, Stonecypher 1990) for coast Douglas-fir, lead to a host of questions about its cause. Of first importance are the physiological mechanisms involved (Wheeler and others 1990). Is interaction a function mainly of genetic variation in growth potential or of growth timing? If growth-timing effects exist, are they matters of timing within years or between years? Some other questions relate to the structure of sets. Would the nonreactive "stable" families within a present set retain their stabilities if tested within a different family matrix? Does the within-set distribution itself (of family heights, for example) affect estimates of interaction; is it possible to preclude interaction within a set by inadvertent or purposeful choice of a distribution of families, half with very poor growth rates and half with very fast growth rates, for example? Would interactions be essentially eliminated if families within sets all originated from one local stand; is interaction characteristically a feature of comparisons of genotypes from somewhat different populations?

Genotype · environment interaction seems to be ubiquitous and of troublesome magnitude in Douglas-fir in western Oregon, even in breeding zones that seldom exceed 80 thousand hectares. There are many questions about it, of which the above questions are a small sample. Considerable work on the measurement, dynamics, and cause of interaction therefore seems justified. Interaction must be understood and measurable in support of increasing precision of models for predicting stand development, yield, and growth responses in a changing climate. The work also is needed to improve selection procedures for tree breeding. For the present, however, genetic improvement could probably be aided by efforts to reconfigure breeding zones. Available data within the region should be examined to determine whether a significant fraction of g·e is indeed associated with site productivity, as it seems to be in Snow Peak. If it is, partitioning the area into low and high site fractions may reduce interaction and increase variance among average family effects. By so doing, it also is possible that zone sizes could be increased significantly without sacrificing genetic gains or adaptation (Stonecypher 1990).

Literature Citations

Adams, W.T.; Joyce, D.G. 1990. Comparison of selection methods for improving volume growth in young coastal Douglas-fir. Silvae Genetica. 39: 219-226.

Allard, R.W.; Bradshaw, A.D. 1964. Implications of genotype-environment interactions in applied plant breeding. Crop Science. 4: 503-507.

Campbell, Robert K. 1986. Mapped genetic variation of Douglas-fir to guide seed transfer in southwest Oregon. Silvae Genetica. 35: 85-96.

Eberhardt, S.A.; Russell, W.A. 1966. Stability parameters for comparing varieties. Crop Science. 6: 36-40.

Finlay, K.W.; Wilkinson, G.N. 1963. The analysis of adaptation in a plant-breeding programme. Australian Journal of Agricultural Research. 14: 742-754.

Franklin, Jerry F.; Dyrness, D.T. 1973. Natural vegetation of Oregon and Washington. Gen. Tech. Rep. PNW-8 Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station. 417 p.

Freeman, G.H. 1973. Statistical methods for the analysis of genotype-environment interactions. Heredity. 31: 339-354.

Froehlich, Henry A.; McNabb, David H.; Gaweda, Frank. 1982. Average annual precipitation, 1960-1980, in southwest Oregon. Ext. Serv. Misc. Publ. 8220. Corvallis, OR: Oregon State University. 7 p.

Gauch, Hugh G., Jr. 1988. Model selection and validation for yield trials with interaction. Biometrics. 44: 705-715.

Griffin, A.R. 1978. Geographic variation in Douglas-fir from the coastal ranges in California. II: Predictive value of a regression model for seedling growth variation. Silvae Genetica. 27: 96-101.

Kitzmiller, Jay H. 1990. Genetic variation and adaptability of Douglas-fir in Northwestern California. In: Proceedings of the joint meetings of the Western Forest Genetic Association and IUFRO working parties; S2.02-05, 12, and 14; 1990 August 20-24; Olympia, WA. Rept. 2. 181. Olympia, WA: [publisher unknown]. 13 p.

Lin, G.S.; Binns, M.R.; Lefkovitch, L.P. 1986. Stability analysis: where do we stand? Crop Science. 26: 894-900.

Monserud, R.A.; Rehfeldt, G.E. 1990. Genetic and environmental components of variation in site index in inland Douglas-fir. Forest Science. 36: 1-9.

SAS Institute Inc. 1987. SAS/STAT™ guide for personal computers, version 6 edition. Cary, NC: SAS Institute Inc. 1028 p.

Silen, R.R.; Wheat, J.G. 1979. Progressive tree improvement program in coastal Douglas-fir. Journal of Forestry. 77: 78-83.

Sokal, Robert R.; Rohlf, F. James. 1969. Biometry. San Francisco: W.H. Freeman and Company. 776 p.

Sorensen, F.C. 1979. Provenance variation in *Pseudotsuga menziesii* seedlings from the *menziesii*-var. *glauca* transition zone in Oregon. Silvae Genetica. 28: 96-103.

Sorensen, F.C. 1983. Geographic variation in seedling Douglas-fir (*Pseudotsuga menziesii*) from the western Siskiyou Mountains in Oregon. Ecology. 64: 696-702.

Stonecypher, R.W. 1990. Assessing effects of seed transfer for selected parents of Douglas-fir: experimental methods and early results. In: Proceedings of the joint meetings of the Western Forest Genetic Association and IUFRO working parties, S2.02-05, 12, and 14; 1990 August 20-24; Olympia, WA. Rept. 2.290, Olympia, WA: [publisher unknown]. 23 p.

Wheeler, N.C.; Stonecypher, R.W.; Jech, K.S. 1990. Physiological characterization of select families to aid in source movement decisions: supplementing long-term performance trials. In: Proceedings of the joint meetings of the Western Forest Genetic Association and IUFRO working parties, S2.02-05, 12, and 14; 1990 August 20-24; Olympia, WA. Rept. 2.362. Olympia, WA: [publisher unknown]. 16 p.

Unrecognized genotype x environment interactions (g·e) can bias genetic-gain predictions and models for predicting growth dynamics or species perturbations by global climate change. This study tested six sets of families in 10 plantation sites in a 78-thousand-hectare breeding zone. Plantation differences accounted for 71 percent of sums of squares (15-year heights), replications an additional 4.4 percent, families 1.9 percent, the first principal component of interaction effects 3.5 percent, and the second principal component 1.2 percent. Results in this study and in a larger survey (87 sets in 10 breeding zones) were similar: 51 percent of sets indicated significant g·e. In 46 percent of sets, the g·e interaction-family variance ratio was greater than 1; in 35 percent, greater than 1.5; and in 10 percent, greater than 5.

U.S. Department of Agriculture
Pacific Northwest Research Station
333 S.W. First Avenue
P.O. Box 3890
Portland, Oregon 92708

Official Business
Penalty for Private Use, $300

do NOT detach label