GENOTYPE-BY-ENVIRONMENT INTERACTION AND DEPLOYMENT CONSIDERATIONS FOR FAMILIES FROM FLORIDA PROVENANCES OF LOBLOLLY PINE

Victor Sierra-Lucero1, Dudley A. Huber2, Steve E. McKeand3, Timothy L. White3 & Donald L. Rockwood2.

1 Currently, Manager of Tree Improvement, Forestal Mininco S.A., Avda. Alemania 751 Los Angeles, Chile.  
2 School of Forest Resources and Conservation, P.O. Box 110410, University of Florida, Gainesville, FL 32611-0410, U.S.A.  
3 Department of Forestry, N.C. State University, Raleigh, NC 27695-8002, U.S.A.  
4 Corresponding author – Phone:352-846-0898; E-Mail: dahuber@ufl.edu  

Received June 24, 2002; accepted May 21, 2003

ABSTRACT

Open-pollinated families of Pinus taeda L. from Marion County and Central Florida provenances were planted at seven sites in the lower coastal plain of the southeastern United States. Data for mean annual increment in volume at age 15 were analyzed. The results indicated that these families were more interactive with environment than previously reported for families from other loblolly pine provenances. Multivariate restricted maximum likelihood was used to define deployment zones with little genotype by-environment interaction within a zone and high interaction between zones. The two statistically defined zones were biologically pleasing with one zone consisting of the western-most tests and the other zone of the tests along the Atlantic Coast. Two sites with aberrant soils failed to group with either zone. Evaluation of deployment and selection options across and between zones indicated that considerable gain in mean annual increment for volume was available through use of these zones. Further research is required to understand the causes and nature of these genotype-by-location interactions.

Keywords: Pinus taeda, Florida source, genotype-by-environment interaction, type B genetic correlation, deployment.

INTRODUCTION

Loblolly pine (Pinus taeda L.) is the most important timber species in the southeastern United States making up 80% of the 10 million ha of forest plantations in this region (Kellison 1992). Several studies have examined genotype-by-environment interactions for loblolly pine and found differences among genotypes across various environments for traits such as growth rate, rust resistance, and sweep (Wells & Switzer 1971; Owino 1977; La Farge & Kraus 1981; Yeiiser et al. 1981; Li & McKeand 1989; Douglass et al. 1993).

When genotype-by-environment interactions are significant, it is important to define how many families contribute to the interaction, the nature of the interaction, and also how economically significant it is, before making any decision about future breeding, testing and deployment (Burdon 1977; Lindgren 1984; Aastveit & Aastveit 1993; Cooper & DeLacy 1994). A better understanding of genotypic stability or conversely, responsiveness, of selected families over various locations can increase genetic gain by allowing breeders to deploy families to specific environments to maximize forest productivity (Li & McKeand 1989). In addition, breeding and progeny testing programs could be altered to produce greater gain if genotype-by-environment interactions are important.

Several statistical methods for assessing genotype-by-environment interactions have been proposed in the field of agriculture and forestry such as variance components, regression, genetic correlations, multivariate analyses and cluster techniques (Shelbourne 1972; Barnes et al. 1984; Skroppa 1984; Lin et al. 1986; Pswarayi et al. 1997). In forestry the primary methods that have been used are genetic correlations (Burdon 1977; Hodge & White 1992; Pswarayi et al. 1997) and regression analysis (Owino 1977; Barnes et al. 1984; Li & McKeand 1989). Both of these types of methods have important advantages (such as being easy to compute and having a strong statistical basis).

Genetic parameters such as heritabilities and interactions of family and environment have not been well studied for the Florida provenances compared to
provenances from other geographic areas of the loblolly pine natural distribution, and these Florida provenances are important both in the southeastern USA as well as in other parts of the world (SCHULTZ 1999; BRIDGWATER et al. 1997). There is also a strong indication that the Florida provenances perform better than those from the Atlantic Coastal Plain and Gulf Coastal Plain when planted in the Lower Coastal Plain (SIERRA-LUCERO 1999). This paper focuses exclusively on 31 families from two Florida provenances, Marion County (MC) and Central Florida (CF).

The objectives of the paper are, for mean annual increment for volume at age 15, to:

1. Estimate genetic parameters including heritability, type B genetic correlations and trait-trait genetic correlations;
2. Examine the family-by-environment interactions within the Florida sources and develop deployment and testing regions homogeneous for genetic performance to improve genetic gains;
3. Compare deployment gains for a potential region based on parental performance in the same and other regions.

MATERIALS AND METHODS

The data

Data were available from seven provenance-progeny tests located in four states: Alabama, Florida, Georgia, and Mississippi (Figure 1, Table 1). The provenances planted in these experiments are from three Coastal Plain areas in the southeastern USA (see Figure 1): South Carolina and Georgia, Florida (CF and MC), and Alabama and Mississippi. Open-pollinated seed-orchard seed was collected from approximately 15 families from each of the four sources (about 60 open-pollinated families). Within each provenance the lightly-rogued or unrogued grafted orchards consisted of wild selections from within the specific provenance and the orchards were physically located within the zone of origin. At each field location, 38 to 60 families were planted in a randomized complete block design in a split-plot experiment with 3 to 5 complete blocks. Provenances were arranged as whole plots and each family within a whole plot was represented by a 10-tree sub-plot. Families from the same provenance were grouped together in a whole-plot within each block to minimize competition among provenances; however, there were no internal border rows. In total, the seven

Figure 1. Test numbers and locations for the seven provenance-progeny studies used in this experiment and zone of origin for the Florida sources.
sites included approximately 14,500 trees.

Only the age-15 data from the two Florida provenances were used in these analyses, i.e. data from Georgia and South Carolina, and Mississippi and Alabama sources were deleted. Each tree was measured for height in meters (HT), diameter at breast height in centimeters (DBH), rust presence or absence, and survival. Tree volume in cubic meters (VOL) (LAMBERTH et al. 1983; HODGE et al. 1996) and mean annual increment for volume in cubic meters per hectare per year (MAIV) (SIERRA-LUCERO 1999) were computed as:

\[
\text{VOL (m}^3\text{)} = (0.25)(3.14)(\text{DBH}^2)(1.37 + 0.33(\text{HT} - 1.37))
\]

\[
\text{MAIV (m}^3\text{ha}^{-1}\text{yr}^{-1}) = ((\text{VOL})(\text{Survival})(\text{Trees per hectare})/\text{Age})/\text{Age}
\]

where:
- Survival = the decimal equivalent of survival of the family row plot at age 15;
- Trees per hectare = planting density of the tests;
- Age = measurement age in years, 15.

Before analysis, data were checked for errors by examining suspected outliers in each block of each site. Outliers are data points outside the normal range of measurement that produce inflated variances (MAGNUSSEN 1993; WHITE 1996). A total of 2.2% of the observations were deleted as suspected outliers for various reasons such as recording errors during data collection or being runts that were likely due to selfing (WHITE et al. 1986). Individual MAIV's were standardized by the square root of the mean square error of each associated test-block combination to remove scale effects and homogenize variances across sites (HODGE et al. 1996; WHITE 1996; DUTILLEUL & CARRIERE 1998; LU 1999). Hence the error variance of the standardized data was approximately 1.0 for all sites, and when pooled across sites.

### Statistical analyses and genetic parameter estimates

The univariate linear model used in the statistical analysis was:

\[
y_{ijkl} = \mu + S_i + B(S)_j + P_k + PS_{ik} + PB(S)_{ijk} + f(P)_{il} + f(P)S_{il} + f(P)B(S)_{iilk} + w_{ijkl}
\]

where:
- \(\mu\) is an overall mean;
- \(S_i\) is the fixed effect of the \(i\)th site, \(i = 1\) to \(7\);
- \(B(S)_j\) is the fixed effect of the \(j\)th block within the \(i\)th site, \(j = 1\) to \(5\);
- \(P_k\) is the fixed effect of the \(k\)th provenance, \(k = 1\) to \(2\);
- \(PS_{ik}\) is the fixed interaction between provenance and site;
- \(PB(S)_{ijk}\) is the fixed interaction between provenance and block within site;
- \(f(P)_{il}\) is the random effect of the \(l\)th family within the \(k\)th provenance, \(l = 1\) to \(15\), \(\sim \text{NID}(0, \sigma^2_{f(P)})\);
- \(f(P)S_{il}\) is the random interaction between family within provenance and site, \(\sim \text{NID}(0, \sigma^2_{f(P)S})\);
- \(f(P)B(S)_{iilk}\) is the random interaction between family within provenance and block within site, \(i.e.\) plot, \(\sim \text{NID}(0, \sigma^2_{f(P)B(S)})\); and
- \(w_{ijkl}\) is random error, \(\sim \text{NID}(0, \sigma^2_w)\).

Variance component estimates from the above model were used to calculate individual-tree heritabilities and type B genetic correlations (YAMADA 1962) for
Unbiased individual tree heritabilities \( h^2 \) were estimated from pooled sites measured at 15 as:

\[
h^2 = V_a / V_p
\]  

where: \( V_a = \) additive genetic variance estimated as \( 4\sigma^2_{(P)} \) from equation 3. \( V_p = \) phenotypic variance within provenance estimated as from equation 6. The coefficient of relationship was assumed to be 4. This implies that the open-pollinated seed-orchard seed were produced by unrelated and non-inbred parents, although more conservative values for the coefficient of relationship for open-pollinated seed have been described (Squillace 1974; Sorensen & White 1988).

The Yamada (1962) method was used to estimate type B genetic correlations by families \( r_{g(p)} \) for each of the regions (various groupings of the seven sites). This method estimates (using a univariate linear model) the genetic correlation of the same trait in different environments (Fernando et al. 1984; Lu 1999).

\[
r_{g(p)} = \frac{\sigma^2_{f(P)}}{\sigma^2_{f(P)} + \sigma^2_{f(P)B(S)}}
\]  

When \( r_{g(p)} = 1 \), there is no family-by-site interaction, and when \( r_{g(p)} = 0 \), family-by-site interaction is strongly indicated. Shelbourne (1972) suggested a limiting value of 0.67 to indicate the presence of severe genotype-by-environment interaction, i.e. values lower than 0.67 compromise the genetic gain and the breeder should consider regionalisation (Matheson & Raymond 1984; Hodge 1996).

The major focus of this paper was to examine the family-by-site interactions for the Florida families. Many indices has been suggested to identify relatively interactive or stable families across sites. These indices and their interrelationships were described by Lin et al. (1986) and Weber et al. (1996), and some of these indices have been used in forestry (Li & McKeand 1989; Ades & Garnier-Géré 1996).

Because of the data imbalance in this study, a stability index, \( p_i \), was used to indicate reactive and stable families across sites. These indices and their interrelationships were described by Lin et al. (1986) and Weber et al. (1996), and some of these indices have been used in forestry (Li & McKeand 1989; Ades & Garnier-Géré 1996).

Genetic correlations between MAIV measured in different regions were estimated using MTDFREML as:

\[
\tau_\sigma(\text{region } 1, \text{region } 2) = \frac{\sigma_{f(P)\text{region } 1, \text{region } 2}}{\sqrt{\sigma^2_{f(P)\text{region } 1} \cdot \sigma^2_{f(P)\text{region } 2}}}
\]  

where \( \sigma_{f(P)\text{region } 1, \text{region } 2} \) is the estimate of the covariance for families within provenance between regions 1 and 2; \( \sigma^2_{f(P)\text{region } 1} \) is the estimate of family within provenance variance for region 1 or 2, respectively from eq. 3.

**Genetic gain estimates**

Deployment gains that quantify the results of using the
best parents were calculated using (1) estimated marginal family values (MFV), which are the estimated family performance across all sites and regions, and (2) estimated family values (FV), which are the estimated family performance in a specific region. The MFV for each family was estimated, using the results from eq. 3, as:

$$MFV_{ij} = LSM_i + f(P)_{ij}$$  \[8\]

where:

- $LSM_i$ is the least square mean for MAIV for a provenance (either MC or CF) across all sites;
- $f(P)_{ij}$ is the BLUP prediction for family 'i' within provenance 'k' across all sites.

The FV, which is specific to a particular family-site combination, was calculated, using the results from analysis using eq. 3, as:

$$FV_{ij} = LSM_i + f(P)_{ij} + f(P)S_{ij}$$  \[9\]

where:

- $LSM_i$ is the least squares mean for a site-provenance combination;
- $f(P)_{ij}$ is the BLUP prediction for family performance across all sites; and
- $f(P)S_{ij}$ is the BLUP prediction of the family within provenance by site interaction, which assumes the interaction is repeatable.

A regional value for a family was calculated as the mean of the FV’s for the family across all sites in the region.

The above least square means and family values were used to predict genetic gains across all sites and also for regional deployment. Deployment gains across all sites were estimated as the average MFV’s of the five best marginal parents over the mean of the two provenance means, converted to a percentage [eq. 8]. Regional deployment gains were calculated as the average of the top five FV’s for a region divided by the mean of the two provenances for the region, converted to a percentage [eq. 9].

### RESULTS AND DISCUSSION

#### Family-by-environment interaction for Florida provenances

In by-site analyses for MAIV, family within provenance was highly significant for all sites except 280 (Table 1). Analysis across the seven sites, Eq. 3, yielded significant site by family within provenance interaction, not shown, and very low heritability (Table 2). The type B genetic correlation for families within provenances across all sites [eq. 5] was 0.27 indicating significant rank changes for the Florida families across the seven sites in this study (Table 2). Due to the importance of family-within-provenance by site interactions for MAIV for the Marion County and Central Florida provenances, analyses were conducted to aid in understanding these genotype-by-environment interactions.

First, graphical analyses (not shown) were performed by plotting the FV’s for each provenance separately [eq. 9], arrayed by site according to the site index. The graph of the MC source showed strong rank changes across sites, while families of the CF source showed fewer rank changes.

The second analysis used $p_i$ [eq. 6] to identify the most reactive families. Eight very reactive families (out of a total of 31 families) with $p_i > 1.6$ accounted for 48% of the sum of the weighted squared family-within-provenance by site predictions for MAIV at age 15 [eq. 6]. Five of the eight reactive families were from the MC provenance and these accounted for 34% of the weighted squared predictions. This proportion of highly reactive families, 26 %, is greater than other authors have reported for loblolly pine (Li & McKean 1989; McKean et al. 1990; Stonecypher et al. 1996; Pswarayi et al. 1997).

When these eight reactive families were dropped from the analysis, the type B genetic correlation increased substantially from 0.27 to 0.64, for the combined analysis across seven sites. Further, the individual heritability [eq.4] for MAIV increased from 0.08 to 0.13.

#### Regionalisation to capture family-by-environment interaction

The low value for heritability and the very low value for type B genetic correlation (Table 2) for the analysis
of all 31 families indicate that these families do not rank consistently across the seven sites. For this reason, it seemed important to find groupings of sites (called regions) such that there was consistent family performance.

Several different runs of MTDFREML were conducted with potential groupings of sites. In each multivariate run, MAIV from different sites within a grouping (region) was treated as the same trait, while MAIV was treated as a different trait between regions. Two groupings of sites were identified as potential deployment regions (Table 3).

Region 1 (R1) contained the two westernmost sites which were located in Alabama and Mississippi (281 and 288, see Fig. 1). Family performance was consistent across these sites with type B genetic correlation of 0.71. Region 2 (R2), located in the Atlantic Coastal Plain (Florida and Georgia), contained three sites (284-286-287, see Figure 1) with a type B genetic correlation of 0.72 (Table 3).

The remaining two sites (280 and 285, see Figure 1) failed to group in any of the defined regions or with each other. There were a priori reasons that may have caused these sites to fail to group with the others. The site in Nassau County, Florida was extremely phosphorus deficient to the point that parts of the test were discarded and block boundaries redrawn (McKEAND pers. comm.). The other site in Levy County, Florida is located on high pH calcareous soils. If these soil characteristics were sufficient to cause rank changes in family performance, then the fact that these two sites failed to group with any of the others seems reasonable.

Grouping sites into regions substantially improved both the narrow sense heritabilities (> 0.2 vs < 0.1) and the type B genetic correlations (> 0.7 vs < 0.3) over those from the seven site pooled analysis for MAIV (Tables 1 and 2). Multivariate analysis of this data structure also yielded a genetic correlation [eq.7] between the two candidate regions for MAIV of 0.37. This low correlation adds further support to considering MAIV as a single trait within each of these regions and a different trait between regions. These results strongly suggest that considerable effort should be given to developing a better understanding of the causes and nature of these genotype-by-environment interactions.

### Genetic gains for deployment

Having developed candidate regions with reasonable genetic correlations and heritabilities within a region and a low genetic correlation between the two regions, it seems appropriate to survey possible genetic gains for deployment from this regionalisation. Using equations 5 and 6, gains from deployment of this population of open-pollinated families were assessed by choosing the top five families (out of 31) for two scenarios: (1) Using the selections from the pooled analysis ignoring regions (MFV) and deploying them across the two regions (FV); and (2) Using the results from each of the regions to choose families to deploy in the other (FV). These results are summarized as follows: (1) Using the overall analysis to choose families resulted in a reduction of 2 to 3 % in MAIV per hectare in R1 and R2, respectively; (2) Selecting families in R1 for deployment in R2 resulted in an 8 % reduction in MAIV per hectare; and (3) Selecting families in R2 to deploy in R1 produced a 4 % loss in MAIV per hectare. These values further confirm that, based on this data, regionalisation did improve gains for MAIV at age 15.

### CONCLUSIONS

The 31 loblolly families from Central Florida and Marion County exhibited much more family-by-site interaction \( r_p = 0.27 \) than would have been anticipated based on previous reports and on the performance of Atlantic Coastal Plain and Gulf Coastal Plain provenances in these same experiments (SIERRA-LUCERO 1999). Further, a higher proportion of these families than expected were responsible for this interaction. These genotype-by-environment results prompted an examination of the data to determine if there were site groupings within which genotypes performed consistently.

The statistically defined regions which resulted from the groupings enquiry were biologically pleasing:

1. Two sites with aberrant soil types failed to group;
2. The western-most sites grouped together, R1; and
3. The Atlantic coastal sites formed a group, R2. While interesting, these groupings are based on

---

**Table 3. Estimates of type B genetic correlations \( r_p \) and heritabilities \( h^2 \) for Central Florida and Marion County source loblolly pine for mean annual increment for volume at age 15, when sites are grouped to raise the type B correlation within a group.**

<table>
<thead>
<tr>
<th>Region</th>
<th>Sites</th>
<th>( r_p )</th>
<th>( h^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>281,288</td>
<td>0.71</td>
<td>0.26</td>
</tr>
<tr>
<td>R2</td>
<td>284, 286, 287</td>
<td>0.72</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Note: Region 1 (R1) represent the two sites in Alabama and Mississippi and Region 2 (R2) contains the three sites in the Atlantic Coastal Plain (Florida and Georgia). The genetic correlation between regions was 0.37.
genotype-by-environment interaction predictions assumed to be repeatable. To the extent that these genotype-by-environment interactions are not repeatable, the pattern of regionalization would not be repeatable. Further, the results are based on few families (31) and few sites (7).

The Central Florida and Marion County materials have been shown to have high mean productivity (SIERRA-LUCERO 1999), but the optimal use of this material in the lower coastal plain seems hindered by a lack of understanding of the genotype-by-environment interactions which are present. Further study is needed to define the environmental factors causing repeatable interaction. When these causes are known, regionalisation can be assessed in a more elegant and practical manner.

ACKNOWLEDGMENTS

This research was supported by the Florida Agricultural Experiment Station, the NCSU Agricultural Research Service, members of the Cooperative Forest Genetics Research Program at the University of Florida and the N.C. State University – Industry Cooperative Tree Improvement Program, the School of Forest Resources and Conservation at the University of Florida, and the NCSU Department of Forestry and is approved for publication as Journal Series Number R-09505.

REFERENCES


