PRACTICAL USES OF BREEDING VALUES IN TREE IMPROVEMENT PROGRAMS AND THEIR PREDICTION FROM PROGENY TEST DATA

T. L. White and G. R. Hodge 1/

Abstract.--One of the main functions of genetic tests is progeny testing, that is prediction of breeding values of parents based on performance of their offspring. By definition, parents with high breeding values tend to produce offspring with superior performance characteristics. Breeding values for a specific trait are unobservable random variables that can never be known exactly. However, because the information is useful in so many ways in tree improvement programs, it is critical that they be predicted both accurately and precisely.

This paper is divided into three major sections. First, we highlight how breeding value predictions for all parents (selections) can be used at many stages of a program to increase genetic gain. Second, several factors are described that make it difficult to precisely predict breeding values (i.e. rank parents) from progeny test data: tests of different ages, tests of different precisions, parents represented in different number of tests. Finally, an analytical methodology called Best Linear Prediction is discussed which was developed by dairy breeders specifically to deal with "messy" progeny test data. This method and a related method (called Best Linear Unbiased Prediction) may well have a place in tree improvement programs that need to efficiently rank parents based on multiple sources of "messy" data.

Additional Keywords: Best Linear Prediction, Best Linear Unbiased Prediction, Messy Data, slash pine, selection index.

INTRODUCTION

Genetic tests are widely used in tree improvement programs for a variety of reasons (see Libby 1973, McKinley 1983). One important function of genetic tests is progeny testing: determining the values of parents based on the performances of their offspring (Allard 1960, p470). For example, the performances of open-pollinated offspring from first-generation selections planted in randomized, replicated field tests can be used to rank the selections. Selections whose offspring perform well for a specific trait, such as disease resistance, are predicted to have high breeding values for that trait. Breeding values are best thought of as unobservable random variables whose values can never be known exactly, but only predicted (Henderson 1973, 1977, 1984 p. 38). That is, based on an observed sample of offspring from a given parent (i.e. a selection) growing in progeny tests, we wish to predict

---

[1]/ Director and Geneticist, Cooperative Forest Genetics Research Program, Department of Forestry, University of Florida, Gainesville, FL, 32611.

how future offspring will perform under operational conditions (i.e. on different field sites and possibly with different cultural treatments than the progeny tests received).

Precise and accurate predictions of breeding values for all selections in a tree improvement program are essential for maximizing genetic gain; however, the "messy" nature of progeny test data from many programs complicates the analytical procedures required to develop these predictions. In this paper, we first highlight the many ways that accurate, precise predictions of breeding values can be used to further tree improvement programs. This is followed by a description of the types of "messy" data problems that arise and complicate the prediction of breeding values from progeny tests. Finally, an analytical method is discussed which has been developed to predict breeding values from "messy" progeny test data. Examples are given from a slash pine breeding program to compare the new method to one currently in use.

PRACTICAL USES OF BREEDING VALUES IN TREE IMPROVEMENT PROGRAMS

The precise, accurate predictions of breeding values are critical in tree improvement programs because the predicted values are used in many ways to increase genetic gains and further the breeding program (see McKinley 1983, Lindgren 1986, White 1987). The brief description of these uses (below) focuses on programs working with sexually-propagated species, though the concepts also apply to vegetatively-propagated species (much of this is adapted from White 1987).

Production Population

The production population in any given generation of breeding is that group of selections used to produce propagules for operational forestation. Clonal (grafted) seed orchards are the most common type of production populations in first-generation tree improvement programs. Increased yield and product quality from the harvesting of these improved forests are the primary realized benefits of most tree improvement programs. Once selections have been ranked based on their predicted breeding values, the production population can be managed to achieve maximum genetic gains.

1. Upgrading the Production Population: Low ranking members can be excluded from the production population and therefore not allowed to contribute genes to operational plantings. As examples, inferior clones can be rogued from first-generation orchards and/or mid-generation orchards (i.e 1.5 generation orchards) can be established which include only superior genotypes.

2. Tailoring Deployment Strategies: The predicted breeding values can also be used to enhance deployment of seed (or vegetative propagules) to specific planting sites. Many companies in the Southeast collect orchard seed and keep it in separate categories for deployment. For example, seed from clones known to be resistant to fusiform rust can be deployed to planting sites expected to experience a high level of exposure to the disease. Deployment strategies can become quite complex and may involve greater use of seed from clones with high predicted breeding values (Lindgren 1986).

Breeding Population

1. Upgrading the Breeding Population: If the intermating of selections is delayed until after progeny test results are available, low ranking selections can be excluded from the breeding population. This increases the genetic quality of the base population of offspring from the intermating, because superior genotypes contribute a higher proportion of the alleles. This potential increase in gain has to be weighed against the time delay of waiting for the progeny test results and the reduction in effective population size.
2. Enhancing Mating Designs: Predicted breeding values can be used to create more efficient mating designs. As only one example (see White 1987), parents with high predicted breeding values can be used in more crosses (Lindgren 1986).

Advanced Generations

1. Increasing Gain from Selection: Advanced-generation selections will typically be made in the pedigreed base populations generated by intermating the members of the breeding population. The predicted breeding values will be used as parental information to aid in finding the best families within which to make selections.
2. Other Uses: The predicted breeding values of the selections of a previous generation of breeding can continue to serve as ancestral information to be combined with the data obtained from the testing the current generation's selections. This enhances gain by increasing the precision of the current breeding value predictions.

MESSY DATA PROBLEMS IN BREEDING VALUE PREDICTION

Progeny tests are usually replicated across both years and diverse planting locations, but because of space limitations in any one field site, all selections are usually not represented in any given test. Data must be combined across multiple tests to develop breeding value predictions and some of the potential factors which can complicate this data analysis are briefly summarized below (see White et al 1986 for more details).

Test Age

Sometimes, tests are established over a wide range of years and test data ranging from say 3 to 15 years old must be combined into a single breeding value prediction for each trait (such as volume growth) for each selection. For growth traits, the older data will probably be more reliable in predicting relative growth performance at harvest (Lambeth 1980) and should therefore receive more weight in the prediction process. This may differ for different traits.

Test Precision

Family means (and rankings) vary in precision from test to test. Tests with small numbers of blocks, variable environmental and vegetation conditions, poor survival, etc, will be less effective in ranking families based on genetic values. In these "poor" tests, more of the variablility in family means is attributable to spurious environmental effects and, intuitively, it seems they should receive less weight in the development of breeding value predictions.

Parents in Different Numbers of Tests

If parents are represented in different numbers of tests and are ranked by the grand family means calculated across all tests in which each occurs, parents with the highest and lowest ranks will tend to be those in the fewest tests (White et al. 1986). This occurs because family means based on many tests will have a lower spread (variance) than those in few tests. This "variance" problem is well known in dairy breeding (Henderson 1973) and results in ranking errors since too many of the high ranking parents will tend to be the least tested (i.e. in fewer tests).

BEST LINEAR PREDICTION OF BREEDING VALUES

Most approaches to the analysis of "messy" progeny test data in forestry (Hatcher et al. 1981, Cotterill et al. 1983) treat family means as fixed effects (Class I model of Eisenhart 1947) and aim at developing family mean estimates of which are adjusted to be free of the complicating factors described above. An alternative analytical approach, developed by Henderson (1973, 1977, 1984), treats parental breeding values as unobservable random effects to be predicted from the observed sample of offspring in progeny tests (Class II model of Eisenhart 1947). This approach leads to two analytical methodologies, Best Linear Prediction (BLP) and Best Linear Unbiased Prediction, the latter of which is used extensively to predict breeding values in animal science. We believe that both of these methods can be useful in predicting breeding values from messy progeny test data in forestry. Both methods involve similar logic and rationale; here we first briefly describe the principles of BLP (see White et al. 1986 for a more complete description) and then summarize its application to open-pollinated slash pine progeny tests.

Properties of BLP

The formula for calculating Best Linear Predictions is (see Henderson 1973, 1977, 1984 for derivation):

$$\hat{g} = C'V^{-1}(y-\alpha).$$

Here, y is an nxl vector of observed data (e.g.,family means from open-pollinated progeny tests) being used to predict g, a pxl vector of breeding values for the p parents represented in the tests. C is cov(y,g'), an nxp matrix of genetic covariances between the observed data and the breeding values being predicted (for example, the covariance between an observed family mean for volume at 5 years with the breeding value of the parent for volume production at maturity). V is var(y), an nxn matrix of variances and covariances among the observations (e.g., the variance of family means in a given test and the covariances between family means in different tests).

Finally, $(y-a)$ expresses the data, y, as deviations from their means, a (e.g., family means from a given test expressed as deviations from the overall test mean). Predictions from this approach have many desirable properties if C, V, and a are known or well-estimated (see Henderson's work).

BLP uses observed data from any ages and traits to predict the breeding values for the traits of interest. This applies to a wide variety of cases including prediction of multiple traits (e.g. volume growth at maturity and rust resistance), different types of relatives, data of different ages and precision and prediction of traits not observed directly (indirect prediction). BLP is an application of selection index (Hazel 1943) that develops a different set of index coefficients for each parent. C^1V^{-1} can be thought of as developing multiple regression coefficients relating y to g (Thompson 1979).

The set of coefficients developed for a particular parent reflects the quality (age, precision) and quantity of data observed for that parent. Predicted breeding values will tend to spread out more (have larger variances among predictions) when there are lots of observed data of high precision that are closely correlated to the trait being predicted. Conversely, if the data are sparse, imprecise and poorly correlated with the trait being predicted (such as using juvenile progeny test volumes to predict mature volume breeding values), the predicted breeding values (of the mature trait being predicted) will cluster more closely around zero. Heuristically, if the deviations of observed data about their means, $(y-a)$, mostly reflect environmental as opposed to genetic causes and are not correlated to the trait being predicted, then they are shrunken (regressed) back towards the mean farther. BLP weights each observed family mean according to its precision and degree of correlation with the trait being predicted and thus handles all of the "messy" data problems described above in one step.

Application of BLP to Slash Pine Open-Pollinated Progeny Test Data

The Cooperative Forest Genetics Research Program (CFGRP) began in the mid-1950's with the mass selection of superior slash pine phenotypes from natural stands and plantations (Goddard 1981). Each cooperating organization made 50-200 selections (cooperative-wide total ..z 1500) which were immediately grafted into clonal seed orchards for the production of improved seed. As the the orchards began to produce seed, cooperators used this orchard open-pollinated seed to establish progeny tests of their selections. Tests were established over a long period of time (1963 to 1983) and to date over 362 tests have been established. Test designs vary, but an average test is a randomized complete block containing 30 families and 6 blocks with a seven to ten-tree row-plot representing each family in each block. Each tree in a test is measured at 5, 10, and sometimes 15 years old for rust incidence and volume. Any given parent is represented in from 1 to 22 such tests.

The current analytical method for analyzing these progeny test data calculates standardized average performances for each test and these are averaged over all tests in which a parent occurs (see Cotterill et al. 1983, Hatcher et al. 1981 for similar approaches). The standardized scores are expressed in standard deviation units from the mean; for example, a parent which was consistently superior for growth over several tests might have a score of 2 while a poor parent would be, say -2. These standard scores have been used to rank parents for the purposes of rogueing orchards,etc.

For BLP, breeding values to be predicted were rust incidence in a 50% rust incidence environment and volume growth at 15 years (White and Hodge 1987 present complete details of this BLP application). These roughly correspond to the rust and growth standard scores, but breeding values are always expressed in the units of measurement as deviations from the average. Thus parents with rust breeding values of say 20% (in a fifty percent hazard environment) and say 1.5 cubic feet (a deviation of 36% above the average of 4.21 cu ft. in 15 year old progeny tests) would be good.

A subset of 28 tests with data for both volume and fusiform rust incidence at all three measurement ages (5, 10 and 15 yrs) was used to estimate variances and covariances for the C and V matrices . Both variances and covariances varied dramatically from test to test and predictive regression models were developed to apply to test data not in the subset of 28. For rust, data from tests with very low or very high rust incidence levels had smaller variances of family means and smaller covariances with the trait being predicted (i.e. rust incidence in a 50% hazard environment). For growth, family means from young tests (especially 5 yrs old) and from tests with high coefficients of variation (large experimental error) had higher variances and smaller covariances with 15 year volume. Data from these types of tests get less weight and result in breeding values that are more "shrunken" or "regressed" back towards the mean.

To apply the formula for BLP, family means were calculated for each test and expressed as deviations from the estimated test mean (both checks and family means were used to estimate test means). Preliminary Best Linear Predictions of breeding values for rust resistance and volume at 15 years were then calculated for approximately 1300 parents and compared to the corresponding standard scores currently used. The coefficients of determination between standard scores and BLP breeding values were $r^2 = 0.77$ and $r^2 = 0.52$ for rust and volume, respectively. Also, for rust, the average absolute rank change from standard scores to BLP was approximately ±100, while for growth it was ±200. So based on these two measures of comparison, rankings of the two analytical systems are more similar for rust than for growth. Presumably, the fact that rust resistance is more heritable than growth and that the data are not so dependent upon measurement age makes the rust data less "messy" and therefore less benefit accrues from the use of BLP. For growth, large rank changes between the two systems occurred in certain situations. For example, a parent ranked high under the old system of standard scores falls considerably in the ranks of the BLP system if its progeny test data are from only a few, young, imprecise tests.

As described in the section on data problems, the standard score approach results in a tendency for parents in fewer tests to be at the extremes of the rankings. BLP has just the opposite trend. On the average in our slash pine data base, parents are in approximately 4 tests in which volume was measured. However, if the top 1% of the parents are selected to be used say in an orchard, those 13 parents (out of z 1300) are represented in an average of approximately 2.5 tests vs. 6 tests using standard scores vs. BLP. Thus, using the standard score approach, there is a tendency to choose as the better parents those that are in fewer tests, while with BLP the tendency is just the opposite: if a parent is good in several tests, it is predicted to have a high breeding value.

DISCUSSION

Best Linear Prediction treats progeny test data in an intrinsically different fashion than do most methods based on averaging family means or scores over tests. For data which is not very messy (tests of nearly the same precision, tests of similar ages, and all parents in similar or large numbers of tests), the benefits of BLP will not be of any practical importance. That is, rankings from most approaches will be similar. However, whenever data from a wide variety of sources (varying ages or precisions, different types of designs or relatives, greenhouse screening data, etc) need to be aggregated to rank parents, BLP (and/or BLUP) may offer significant advantages. For instance, in advanced generation breeding programs, it may be necessary to use data from several generations of testing to precisely predict breeding values. One distinct benefit of BLP is the ease in which multiple traits are handled. Multiple trait selection indices are easily formed after or at the same time as the predictions are made for the individual traits.

The accuracy and precision of BLP breeding values depend on the estimation of the variances and covariances in the C and V matrices. While more work is needed to assess the sensitivity of the predictions to errors in estimating these second moments, we currently believe that for most tree improvement programs they can be estimated well enough to make BLP or BLUP a substantial improvement over conventional methods of analyzing "messy" progeny test data.

LITERATURE CITED

Allard, R.W. 1960. Principles of Plant Breeding.John Wiley and Sons, New York,NY. 485p.

Cotterill, P.P., R.L. Correll and R. Boardman. 1983. Methods of estimating the average performance of families across incomplete open-pollinated progeny tests. Silvae Genet 32:28-32.

Eisenhart, C. 1947. The assumptions underlying the analysis of variance. Biometrics 3:1-21.

Goddard, R.E. 1980. The University of Florida Cooperative Forest Genetics Research Program. pp. 31-42 In Proc. 15th North Amer. Quant. For. Genet. Group. August 6-8. Coeur d'Alene, ID.

Hatcher, A.V., F.E. Bridgwater and R.J. Weir. 1981. Performance level-Standardized score for progeny test performance. Silvae Genet 30:184-187.

Hazel, L.N. 1943. The genetic basis for constructing selection indexes. Genetics 28:476-490.

Henderson, C.R. 1973. Sire evaluation and genetic trends. pp. 10-41. In Animal Breeding and Genetics Symposium in Honor of J. Lush. Animal Sci. Assoc. Amer., Champaign, Ill.

Henderson, CA. 1977. Prediction of future records. pp. 615-638. In Proc. Internat. Conf. Quan. Genet.(E. Pollack, O. Kempthorne and T. Bailey eds.). The Iowa State Univ. Press, Ames, IO.

Henderson, C.R. 1984. Applications of Linear Models in Anumal Breeding. Univ. Guelph Press. Guelph, Ontario, Can. 462p.

Lambeth, C.C. 1980. Juvenile-mature correlations in Pinaceae and implications for early selection. Forest Sci 26:571-580.

Libby, W.J. 1973. Domestication strategies for forest trees. Can J For Res 3:265-276.

Lindgren, D. 1986. How should breeders respond to breeding values?. pp. 361-372. In: Proc. IUFRO Conf. on Breeding Theory, Progeny Testing and Seed Orchards. October 13-17, Williamsburg, VA.

McKinley, C.R. 1983. Objectives of progeny tests. pp. 2-13. In Progeny Testing of Forest Trees. Southern Coop. Series Bull. 275. Texas A & M Univ., College Station, TX, 77843.

Thompson, R. 1979. Sire evaluation. Biometrics 35:339-353.

White, T.L. 1987. A conceptual framework for tree improvement programs. In Preparation.

White, T.L. and G.R. Hodge. 1987. Best linear prediction of breeding values in a slash pine tree improvement program. In Preparation.

White, T.L., G.R. Hodge and M.A. Delorenzo. 1986. Best linear prediction of breeding values in forest tree improvement. pp. 99-122 In Southern Coop. Series Bull. 324. Univ. of Florida, Gainesville, FL.