TWO-STAGE EARLY SELECTION:
A METHOD FOR PRIORITIZATION AND WEIGHTING OF TRAITS

Cheryl B. Talbert"

Abstract--A simulation approach was used to evaluate the impact of a
two-stage early-screening-plus-field-test selection program on predicted
total selection-progress in a 'mature-stand' selection-goal. Early family-
screening , when followed by field testing for a specified time period, can
lead to more, less or the same expected gain than a standard 'field-test-
only' program, depending upon the correlations among the early-screening,
field and mature-stand traits chosen.   If the early-screening and field-
test criteria chosen are either strongly positively or strongly negatively
correlated, both total gain and gain per unit time may be less than if the
early screening had never been carried out.   This effect can be reduced by
appropriate allocation of selection intensity between the early-screening
and field steps.   Economic analysis will be necessary to evaluate whether
the benefits of early screening (in quality of the early-test environ-
ment(s), reduced field-test size, and/or larger family-size and greater
selection efficiency for a given field-test size) will outweigh its neg-
ative impacts under the conditions faced by a particular organization.

INTRODUCTION

Selection considerably prior to harvest-age is an operational   reality
in loblolly pine tree improvement, as a result of persistently high alter-
native rates of return and the crushing resource costs of carrying large
field tests over long periods of time.   The excellent research of the past
decade into methods of greenhouse, laboratory and nursery selection for
improvement of later field performance (reviewed by Lambeth, 1483 and by
Talbert and Lambeth, 1984) has not yet produced results conclusive enough
to cause operational programs to move away from conventional field tests of
4 to 8 years duration. However, a number of organizations are seriously
planning one- or two-year greenhouse, lab or nursery trials to 'screen out'
their poorest families for growth, quality and/or adaptability, in order to
reduce the size of their field-tests.

A likely scenario for such a program is selection in two or more
stages, where some proportion of a population of half-sib and/or full-sib
families would be discarded at some early age, or at several early ages, on
the basis of seedling traits. After this initial truncation, field tests
would be planted with reserve seed from the remaining families, and the
final set of orchard parents or selections for the subsequent generation
would then be chosen from those tests based upon survival, pest-response
and height or volume. This discussion will use the term early screening to
define truncation-selection based upon seedling traits.

Considerable theoretical work in the agronomy literature has shown
that, when different selection criteria are correlated, culling of the
population for one criterion can drastically impact gain-potential from
other criteria, and the total progress that can be achieved is strongly

1/ Scientist, Forestry Research Division, Weyerhaeuser Company WTC-2H2,
Tacoma, Washington 98477.

affected by the relative correlations of each criterion with the desired mature-stand traits, and by the relative economic importance of those mature-stand traits (Jain and Amble, 1962; Namkoong, 1970; Cunningham, 1975). Multi-stage culling can actually reduce total 'value gain' if the ordering of traits and the selection pressure applied to each trait is not properly matched to the genetic and economic characteristics of the population. At the same time, early family-screening can provide a number of advantages over conventional 'field-only' testing, which could offset these disadvantages (Talbert and Lambeth, 1984) - advantages in quality of the test environment(s) (for example, the ability to evaluate rust resistance in a high-inoculum environment), in reduced field-test size, and/or in larger family-size (and greater selection efficiency) for a given field-test size.

This report will explore the impact of early truncations on total gain from a fairly simple two-stage 'early-screening + field-test' selection program, and will explore alternatives for maximizing gain from such a program in populations with a variety of characteristics.

The Quantitative Basis for Two-Stage Selection

Selection is usually justified by the expectation that some desired change will occur as a result, in one or more correlated traits (the simplest example would be genetic value for the original trait). By the same token, any early truncation of a population will influence the mean and variance of later-stage, correlated traits, thereby affecting the progress achievable from selection on those later-stage traits. The impact of sequential truncations on predicted gain can be quantified based upon the characteristics of the multivariate-normal distribution (Eisen, 1983). The general theory is adapted to the current example below.

Because early selection is most commonly carried out to improve average genetic value for one or more harvest-age traits, gain in a harvest-age trait M resulting from early selection for the same or different trait J is appropriately described using an equation for indirect selection:

$$\Delta G_{M,J} = i_J Cov(M,J) / \sigma_{p_J} , \qquad (1)$$

where i_J = the selection intensity practiced on J,

$Cov(M,J)$ = the covariance between genetic value for the harvest-age trait and phenotypic value for the early trait, and

σ_{p_J} = the phenotypic standard deviation for J.

For individual selection this formula reduces to the familiar form:

$$\Delta G_{M,J} = i_J h_J h_M r_{G_{MJ}} \sigma_{p_M} , \qquad (2)$$

where h = square root of the heritability, and

$r_{G_{MJ}}$ = genetic correlation between the mature trait M and the early trait J.

Now, instead of one early-selection trait J, consider the case of two early-selection criteria J and F. In the current context, J would be the greenhouse/lab/nursery trait and F would be a field-test trait; therefore, J must be assessed prior to F. If a proportion of families p(J) is selected based upon J, retaining a reduced population having a standardized selection differenti,11 the variance of the field trait F is reduced whenever J and F are correlated:

$$V_F* = V_F - \frac{[Cov(J,F)^2 \times i_J(i_J-c_J)}{V_J}, \qquad (3)$$

where   *   = designation of adjustment for 1st-stage selection,

     V_F = the variance of F prior to selection on J,

     Cov(J,F) = the covariance between J and F prior to selection on J,

and c_J = the standardized truncation point for J.

In addition, the selection intensity which can be applied to F after selection on J, i_F*, is reduced, regardless of the correlation between J and F, due to the fact that the proportion left for field selection out of a fixed total proportion P to be selected, is reduced.

The covariance between F and M are also affected by prior selection whenever either F or M are correlated with J:

$$Cov(F,M)* = Cov(F,M) - \frac{[Cov(J,F) \times Cov(J,M) \times i_J(i_J-c_J)]}{V_J}$$

$$\qquad (4)$$

The correlated gain in M resulting from the selection on J can be predicted by equation (1). However, the gain which can be obtained from subsequent selection on F will be altered by the first-stage truncation, because of the impact of the truncation on i_F, on the variance of F, and on the correlation of F with M. Therefore, gain in M resulting from two-stage truncation selection on J and F will be:

$$\Delta G_{M,J+F} = \Delta G_{M,J} + i_F Cov(F,M)* / \sigma_{p_F*} .$$

The Simulation Analysis

To illustrate the impacts of first-stage selection on total progress from a two-stage selection program, several simulated populations were carried through equations 1-5. A number of combinations of genetic correlations between J, F and M were used, chosen to represent a range of combinations rather than to model any specific biological scenario (Table 1). For purposes of the simulation, individual-tree narrow sense heritabilities are assumed to be 0.2 for J, F and M, and the phenotypic variances for J, F and M are held at 1, 10 and 100, respectively, for all of the simulations. Finally, it is assumed that each population consisted of 100 half-sib families, with a constant family-size of 200 for the early screening and 60 in the field. Half-sib family selection is used in the early-screening stage, and a combined half-sib-family and within-family index is used in the field stage. It is important to note that, since reserve seed

's used to plant the field-test in our example, none of the same trees would be measured for both J and F, although half-sibs would be.

Table 1. Ranges of selection intensity and inter-trait genetic correlations used in the simulation analysis.

Parameter	Definition	Range
i_J	selection intensity for J	0.35 - 1.73
p_J	proportion of families selected for J	80% - 10%
rG_{JF}	genetic correlation between J and F	-0.4, 0.1, 0.4, 0.8
$rGJM$	genetic correlation between J and M	-0.4, -0.2, 0.2, 0.4
rG_{FM}	genetic correlation between F and M	-0.2, -0.4, 0.4, 0.2

RESULTS

Figures 1a and 1b illustrate the impact of increasing intensity of early half-sib family screening for a trait J on the phenotypic variance of a field-trait index F, and on the correlation between genetic value for M and the phenotypic index value for F, when the genetic correlation between J and M and between F and M prior to the early screening are equal in absolute value at rG = +0.4 or -0.4.

Regardless of the genetic correlations between J and H and between F and M, when the genetic correlation between J and F is strongly positive, the variance of the field-trait index F decreases to less than two-thirds of its original ($rG(J,F)=0$) value as the proportion of families selected for j reaches 20 percent (or 20 of 100 families selected). If the genetic correlations between S. and H, F and M and J and F prior to the early-screening are of the same sign, selection for j results in a decrease in the absolute value of the correlation between F and M [2]/. On the other hand, if the genetic correlations between J and M and between F and M are of the same sign and the correlation between J and F is of the opposite sign, then early screening for J will decrease the absolute genetic correlation between F and M if $rG(JM)$ is positive, and will increase that absolute correlation if $rG(Jf)$ is negative. Finally, if $rG(JM)$ and $rG(FM)$ are of opposite sign, then early screening for J will increase the absolute value of $rG(FM)$ if the genetic correlation between J and F is positive, and will decrease the absolute value of $rG(FM)$ if $re(JM.F)$ is negative.

[2]/ The importance of changes in $rG(FM)$ result from the direct proportionality of $rG(FM)$ with expected second-stage gain. If expected gain from selection on a given variable turns out to be negative, then selection on the opposite end of the scale for that variable will make the gain positive, and of the same magnitude. Therefore, absolute correlations and expected-gain values are the quantities of interest.
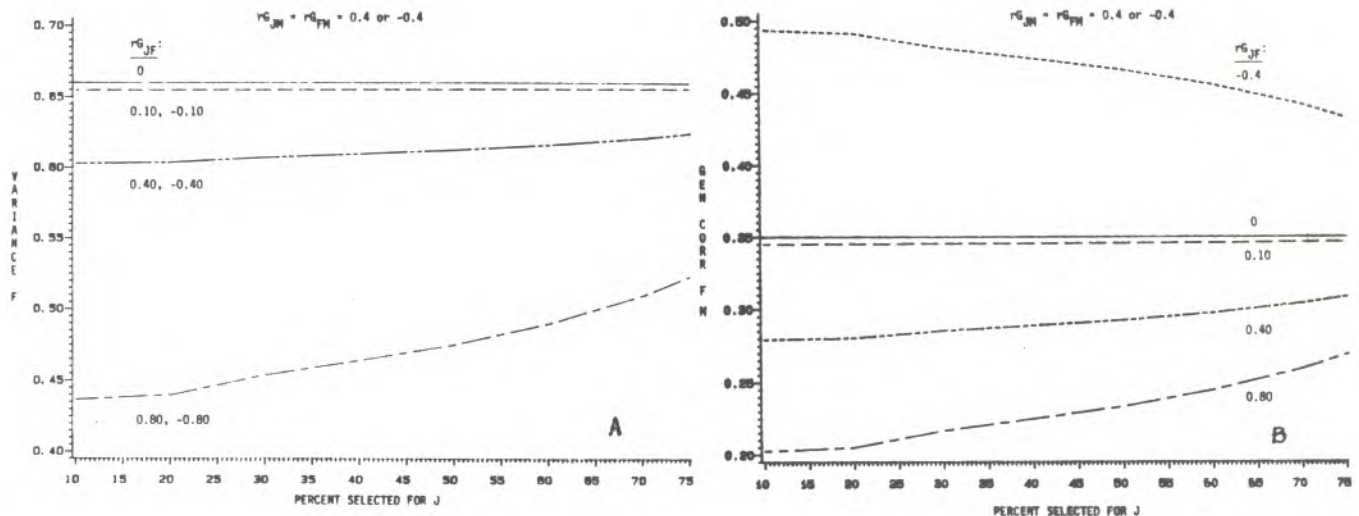
FIGURE 1. The impact of early screening for a trait J on (A) the variance of a correlated field-trait F, and (B) the correlation of F with a selection goal M. (Assumptions and definitions in text and in Table 1)

How can selection for one variable actually _increase_ the correlation between two other variables? This can happen because, under certain conditions, removal of entries on the low end of the scale for J means removal of entries which are simultaneously low for M (due to a positive correlation of J and M) and high for F (due to a negative correlation of J with F). An opposite but equivalent example exists if J and M are _negatively_ correlated and J and F are _positively_ correlated. The outcome is an upward pressure on the absolute value of the correlation between F and M.

The combined impact of these effects is illustrated in Figures 2a-2c, in terms of the predicted progress in the mature trait M which would be expected from the two-stage selection program relative to that which would be expected from selection for the field-trait index alone, when the genetic correlations of J with M and of F with M are again equal in absolute value at +0.4 or -0.4.

Regardless of the proportion selected in the early-screening step, when the genetic correlation between J and F is less than 0.4 or negative and the correlations of J and F with the selection goal are both either +0.4 or -0.4, selection for both J and F yields more expected gain (in M) than does field selection alone. In fact, under these conditions, if the genetic correlation between J and F is negative the expected gain in M from the second-stage, field-selection step alone is greater than the total expected gain would have been from a single-stage field selection on F. On the other hand, if J and F are strongly positively correlated (>0.4), and both r6(JM) and rG(FM) are either +0.4 or -0.4, two-stage selection yields less expected gain in M than does one-stage selection for F alone.

The opposite behavior occurs when the genetic correlations between and M and F and M are opposite in sign and equal in magnitude at :0.4:. In
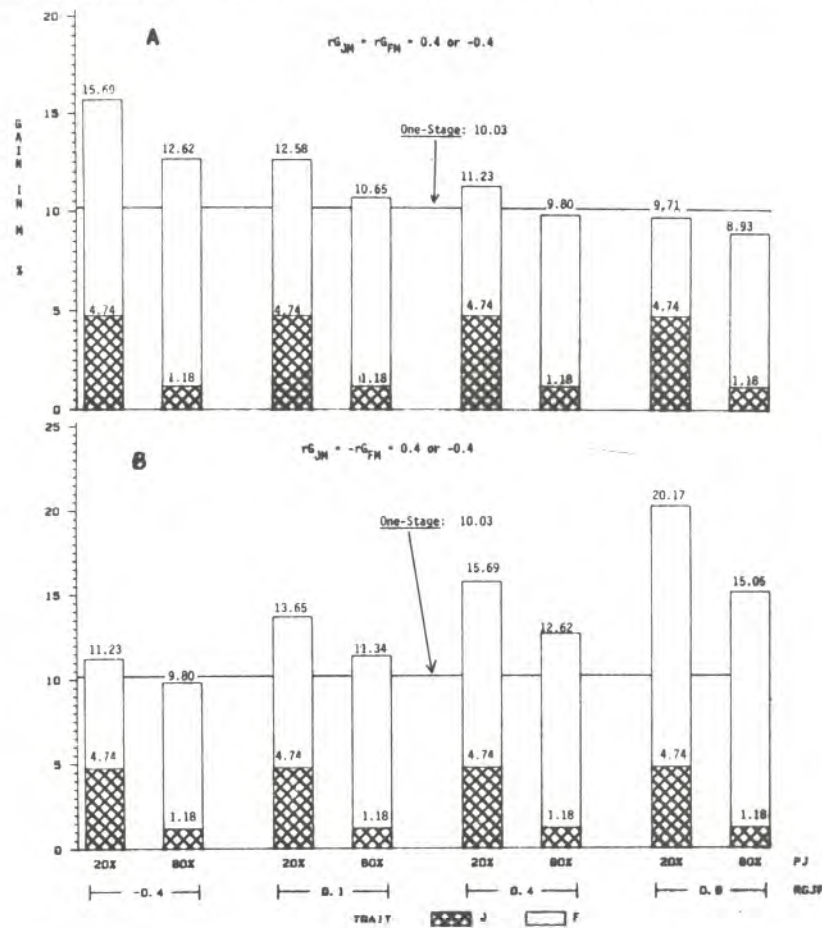
111

FIGURE 2. Predicted gain in a mature-stand trait M resulting from two-stage 'early-screening-plus-field' selection for a seedling trait J and a field trait F, compared to field-selection only, for two early-screening intensities: (a) when rG(JM) and rG(FM) are both 0.4 or both -0.4; and (b) when rG(JM) and rG(FM) are opposite in sign but equal in magnitude at |0.4|. (Assumptions and definitions are contained in Table 1)

such a scenario, unless J and F are strongly <u>negatively</u> correlated, two-stage selection yields greater expected gain than one-stage selection.

As expected, gain from two-stage selection varies in magnitude with the degree of selection-pressure placed on the early-screening trait J, while expected gain from the one-stage 'field-only' method remains unaffected. For every population scenario tested, rigorous early screening yields greater gain overall than does low-intensity early screening. This occurs despite the fact that greater early-screening gain sometimes results in reduced gain from field-test selection; the increase in gain resulting from the increased early-selection intensity even in these scenarios is greater than the decrease in gain from the field-selection step. Therefore, it appears that it is possible to allocate selection intensity between early-screening and field-test steps in a two-stage selection program so as to make two-stage selection at least as effective, and often much more effective, in overall expected gain than a one-stage field-test-only program.

Needless to say, it is not realistic to presume that the genetic

112

correlation between a chosen early-screening trait J and the 'mature-stand' selection goal M will always be the same as the genetic correlation between the chosen field-test trait (F) and M. Figures 3a-3d illustrate the effects of differences between J and F in the strength of their genetic relation- ships with M, ranging from a case where the early-screening trait J is less-strongly correlated with M than is the field-trait F (at rG = 0.2 and 0.4 respectively), to a case where J is twice as strongly correlated with M than is F (rG = 0.4 and 0.2 respectively). (Again, the attempt here is to utilize a range of theoretically-possible scenarios, rather than to pick a single empirical example, so that readers will have maximum latitude in finding their own 'most useful' case.) Of course, when rG(FM) is the absolute gain from field selection in both the one-stage and the two- stage examples is less than it is when rG(FM) is :0.4:. Relative to the 'conventional' one-stage approach, however, two-stage gain is affected in a predictable way. In situations where the early screening has a detrimental impact on gain from field-selection, a lower genetic correlation between J and M reduces expected gain from early-selection, and therefore decreases the detrimental impact; the two-stage approach 'looks better' relative to one-stage selection. In situations where early screening actually in- creases expected gain from field-selection, a stronger genetic correlation between J and M increases the magnitude of the positive effect, and two- stage selection becomes proportionally better relative to one-stage selec- tion. Even under these altered scenarios the greatest two-stage gain is produced when rigorous selection is applied in the early-screening step.

It is important to note that the relationships and results discussed here are dependent only upon underlying biological relationships, and are independent of scale. If the scale of J, F or M is reversed, all of the relationships of that variable with other variables in equations 3-5 will be affected, and the combined effects will cancel one another. It is assumed in this analysis, however, that selection will be carried out appropriately relative to scale: in other, words, that when rG(FM) or rG(FM) is negative, the 'lowest value' entries for J or F will be selected, so that expected progress in M will always be positive.

DISCUSSION, RECOMMENDATIONS

Quantitative methods are most appropriately used not in st aking de- cisions, but rather as tools to augment the knowledge and experience of the decisionmaker. In the case of two-stage screening, the quantitative tools used in this analysis provide four guidelines:

. Early family screening for a seedling trait J, when followed by selection for a trait F in a conventional field-test, may actually lead to reduced total expected gain in a 'mature-stand' selection goal (M) from the two stages combined, compared to what would be expected from field selec- tion alone. This will occur if the genetic correlations between J and M and F and M are of the same sign, and the early-screening trait is strongly positively correlated with the field-test trait (rG > 0.4), or alternative- ly, if rG(FM) and rG(FM) are opposite in sign and rG(JF) is strongly negative (rG < -0.4). In all other scenarios tested, two-stage selection provided at least as much and generally more total expected gain than one- staae 'field-only' selection.
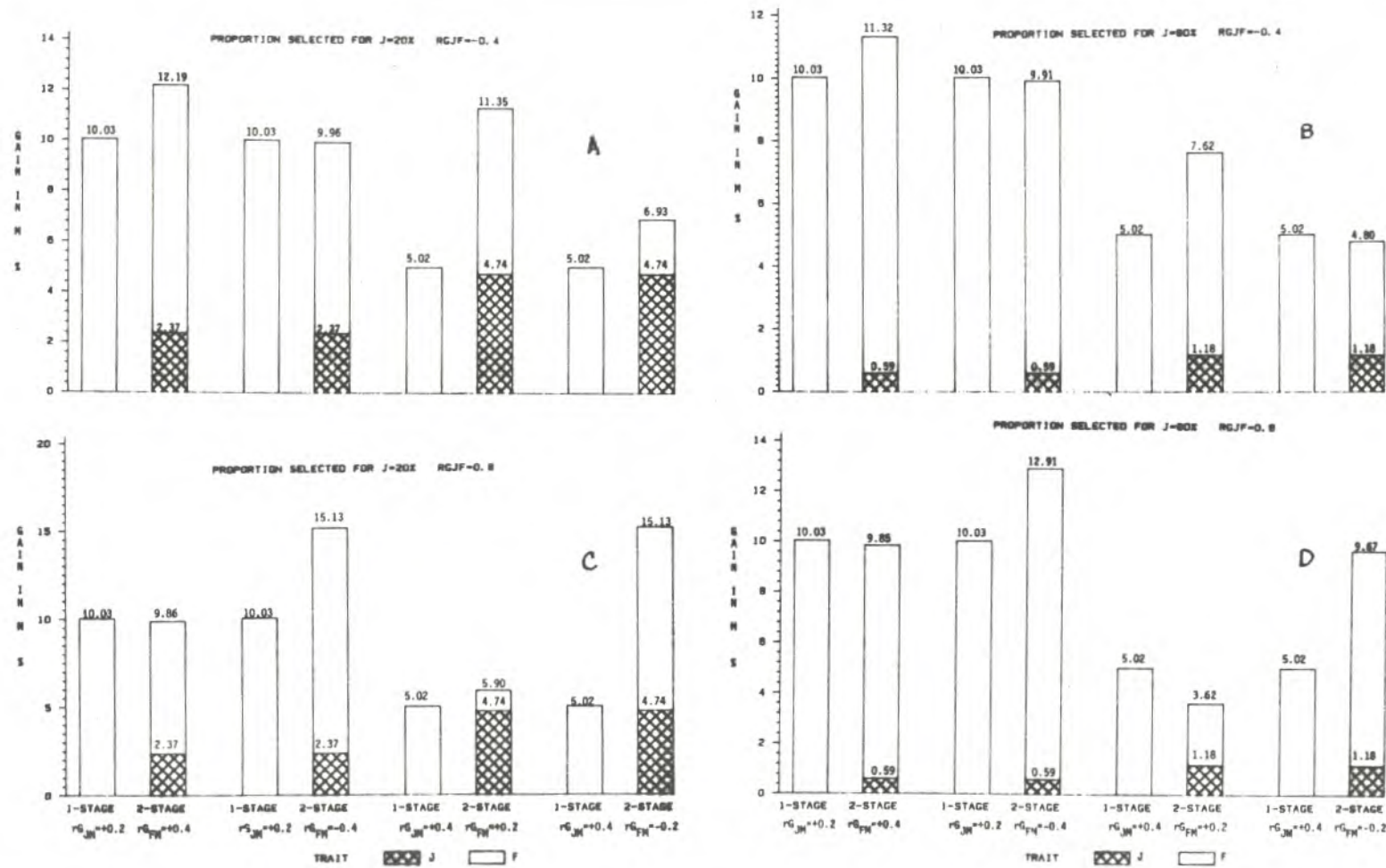
113

FIGURE 3. Predicted gain in a selection goal M resulting from two-stage selection, for an early-screening trait J followed by a field-test trait F: (a) when rG(JF) is negative and rigorous early selection is applied; (b) when rG(JF) is negative and weak early selection is applied; (c) when rG(JF) is strongly positive and rigorous early selection is applied; and (d) when rG(JF) is strongly negative and weak early selection is applied. (Assumptions and definitions are given in text and in Table 1.)

. If a two-stage selection program is assumed to require 1-2 more years to complete than a one-stage program, most two-stage scenarios tested yield less expected gain per year than one-stage selection, for the same rG(FM).

. The stronger the genetic correlation between J and M, the greater the gain from first-stage truncation, and the greater the resulting impact on second-stage gain. Under conditions where early screening is detrimental to field-test-selection progress, the detrimental impact is more severe when rG(JM) is increased, and less severe when rG(JM) is decreased in absolute value. On the other hand, when early screening increases expected field-test-selection progress, an increase in rG(JM) increases gain from both the first and the second-stage.

. Regardless of the scenario, if a two-stage selection approach is chosen, the greatest gain will result when strong selection pressure is placed on the early-screening traits - the resulting increase in gain from the early screening offsets the corresponding reduction in predicted gain from the later field-selection step.

Most tree-improvement decisionmakers operate with some 'feeling' for the general magnitude of interrelationships among selection-criteria, and between selection-criteria and selection-goals; in many cases, there are

actual published relationships. Although experience-based parameter-estimates will not always be unbiased, they are likely to be sufficiently accurate to allow for decisionmaking based upon the above results. To ignore these guidelines is to implicitly assume that the correlations among, early-screening and field traits are zero.

Because of today's grueling economic conditions, early-screening may appear very attractive as a means of reducing long-term field-testing costs. These results show, however, that two-stage selection may not always yield greater gain for the additional effort. The analytical methods discussed here allow the tree improver to choose the most appropriate relative magnitude of selection-emphasis to apply to different selection criteria throughout a given testing-cycle, to determine the maximum possible two-stage expected gain under the specific conditions suspected to be operating in the population of interest. It is critical that a careful economic analysis then be carried out to determine whether two-stage selection is a desirable course of action.

LITERATURE CITED

CUNNINGHAM, E.P. 1975. Multi-stage index selection. TAG 46: 55-61.

EISEN, E.J. 1983. Population genetics in animal improvement. Unpub. lecture notes, N.C. State University.

JAIN, J.P. and V.N. AMBLE. 1962. Improvement through selection at succes-sive stages. J. Indian Soc. Agric. Stat. 14: 88-109.

LAMBETH, C.C. 1983. Early testing - an overview with emphasis on loblolly pine. In: Proc. 17th S. For. Tree Imp. Conf., Athens, GA: 297-311.

LIN, C.Y. 1978. Index selection for genetic improvement of quantitative characters. TAG 52: 49-56.

NAMKOONG, G. 1970. Optimum allocation of selection intensity in two stages of truncation selection. Biometrics 26: 465-476.

TALPERT, C.B. and C.C. LAMBETH. 1984. Early testing and multiple-stage selection. In: Proc. S23 Meeting on Progeny Testing, Baton Rouge, LA (In Press).