COMPUTER MODELING OF A SUBLINING BREEDING SYSTEM

F.E. Bridgwater, W.C. Woodbridge [1] and M.F. Mahalovich [2]

Abstract.-- Many tree breeding programs have adopted some form of multiple population strategy (sublining) to manage inbreeding. Many questions arise about how these populations should be constructed and managed. Among these are the questions of how individuals should be assigned to sublines and how selection should be done within and among sublines. Computer simulation of a sublining breeding strategy suggests that: (1) selecting the best individual from the four best families (self or outcross) ranked on expected breeding values is an alternative that will give good genetic gains and result in relatively moderate rates of increase in coancestry within sublines; (2) assigning parents to sublines at random or disassortatively rather than by positive assortment will increase within-subline genetic variance and result in greater expected genetic gains.

Keywords: Positive assortative mating, disassortative mating, random mating, elite populations.

INTRODUCTION

The North Carolina State University-Industry Cooperative Tree Improvement Program adopted a breeding strategy in 1992 for the third cycle of selection and breeding. The strategy aims to provide maximum genetic gains in the short term as well as to maintain genetic diversity to ensure the viability of breeding populations in the long term. The details of the strategy may be found elsewhere (McKeand and Bridgwater 1992), but the fundamental population structure is a hierarchy of three populations. A **mainline breeding population** will have about 160 parents available for each Cooperative member and will be maintained in small sublines (size 4, in the plan) primarily to provide for long term genetic gains. The most intensively selected and managed level in the hierarchy will be elite **populations** of about 40 parents. The elite populations will be bred as rapidly as possible to provide maximum short-term genetic gains. A third level in the population hierarchy will be extreme genotypes maintained as a genetic diversity archive.

The goal is to cycle elite populations as rapidly as possible. Since production populations will be derived from these, genetic gains will be realized from plantations sooner than if larger populations, requiring more time and effort, were used. Elite populations may be managed in a variety of ways and will be structured to meet the needs of individual or groups of cooperative members. One option under evaluation is to subdivide elite populations into sublines as small as 4 parents each and to mate these in diallels which will include self-fertilization. This action will result in inbreeding and an increase in homozygosity at a rate that depends on the method of selection employed. When inbreeding reaches levels that require reduction, the plan is to enrich the elite populations from the much larger mainline population which will have been bred at a slower rate. The greater selection intensity possible in the larger mainline populations is expected to provide material suitable for inclusion in the elite populations. Introductions from the mainline populations will reduce coancestry in elite populations.

---

[1] Research Geneticist, and Research Forester, USDA-Forest Service, Southeastern Forest Experiment Station, Raleigh, NC

[2] Selective Breeding Specialist, USDA-Forest Service, Intermountain Research Station, Moscow, ID

Regular systems of breeding in small populations have been examined in great detail elsewhere (Falconer 1989). However, in reality, breeding populations are managed in a much more eclectic manner than theoretical breeding schemes. We have used computer modeling to simulate alternative selection and mating schemes that breeders might use to manage multiple populations in sub lines.

MODEL CHARACTERISTICS

Genetic Model

The genetic model permits variable numbers of loci up to 128 in multiples of 8 loci. At present, all loci act independently and have only two alternative alleles (0 or 1) at a locus. The genetic value assigned to an allele is a variable specified when the population is generated. At present, all alleles of the same type have the same value.

Mainline Population Generation

Populations of up to 500 individuals can be generated (Figure 1). Each individual in the population is generated by randomly assigning an allele to each locus. Gene frequencies may be varied when the population is generated, but were made 0.5 in the base population (Generation 0) for this simulation. A genotypic value was calculated for each individual by summing gene values over loci. Total genetic variance (V_i) was assumed to **be** the sum of additive (VA and dominance (V_d) variances, that is, V, = V, Vd. Variances due to epistasis, linkage, and linkage disequilibrium were assumed to be zero. Since populations were at equilibrium only at generation 0, genetic variances at each locus were calculated from genotype frequencies, rather than gene frequencies:

Genotype	A_1/A_1	A_1/A_2	A_2/A_2
Value	G_{11}	G_{12}	G_{22}
Frequency	P_{11}	P_{12}	P_{22}

Fitting the model $G_{ij} = u + a_1 + a_2$ (where u = the overall mean, a_1 and a_2 are the average effects of A_1 and A_2) by least squares minimizes:

$$Q = P_{11}[G_{11}-u-2a2_1]^2 + P_{12}[G_{12}-u-a_1-a_2]^2 + P_{22}[G_{22}-u-2a_2]^2$$

and the genetic variances are:

$$V_a = P_{11}(2a_1)^2 + P_{12}(a_1+a_2)^2 + P_{22}(2a_2)^2 \text{ and,}$$

$$V_d = P_{11}(G_{11}-u-2a_1)^2 + P_{12}(G_{12}-u-a_1-a_2)^2 + P_{22}(G_{22}-u-2a_2)^2.$$

The expressions for the a1 are:

$$a_1 = -a_2(P_{12}+P_{22})/(2P_{11}+P_{12}) \text{ and } a_2 = -a_1(2P_{11}+P_{12})/(P_{12}+2P_{22}).$$

Phenotypic values may be assigned to individuals in one of two ways. A narrow-sense heritability (h^2) can be specified and the environmental variance (V_s) is calculated from:

$$V_p = V_a/h^2 \text{ and } V_e = V_p - V_a$$

or V_e can be specified, in which case:

$$V_p = V_e + V_g \text{ and } h^2 = V_a/V_p.$$

Individuals generated in the mainline population may be ranked by breeding value, genotype, phenotype, or at random and are saved to a file that is used as input to the breeding strategy model.

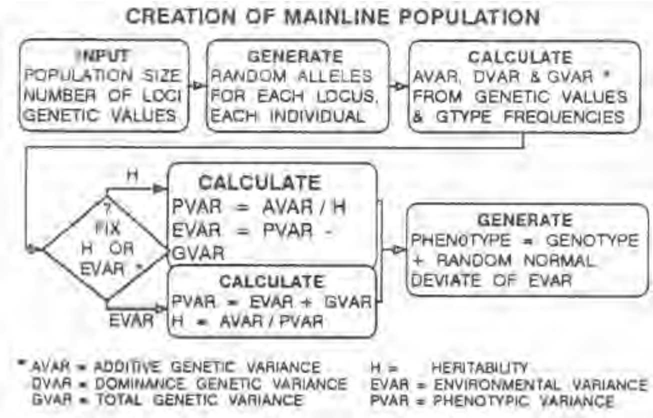

Figure 1. Flow Diagram for the generation of a mainline population.

Breeding Strategy Model

Elite populations may be generated in different ways by selecting a subset of the mainline population. That is, if the top quartile of the mainline population based on breeding values were to be included in an elite population, it would be necessary to generate a mainline population ranked by breeding values. Then the top quartile should be selected for the breeding strategy model.

After a mainline population is generated, the breeding strategy model (Figure 2) permits the assignment of individuals to sublines in different ways. Three possibilities are: (1) Positive assortative mating (PAM) (1,2,3,4), (5,6,7,8), ... (n-3,n-2,n-1,n); (2) Dissasortative mating (DAM) (1,2,n-1,n), (3,4,n-3,n-2), ... (n/2-2,n/2-1,n/2+1,n/2+2); and (3) Random mating (RAM) (Assigned to 4-parent diallels at random), where 1,2,3,4, ... n imply ranks based on expected mid-parent values.

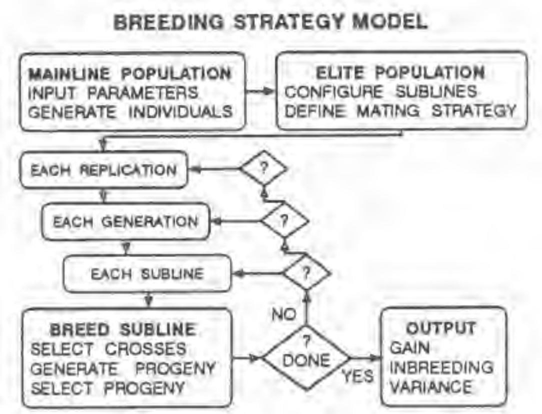

Figure 2. Flow Diagram for the breeding strategy model.

Selection within sublines is done in each generation by ranking each parent and crosses among them based on expectations from their breeding values (Figure 3). Individuals within progeny groups are generated by random segregation at each parental locus and are selected based on phenotypes. An allele identifier is assigned to each allele at each locus to facilitate the calculation of inbreeding coefficients ("F").

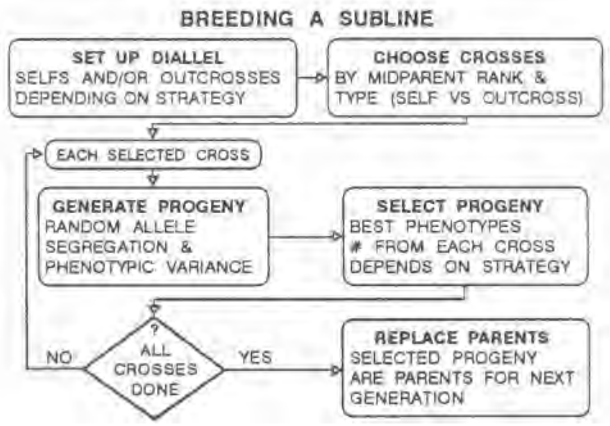

Figure 3. Mating a Subline

Different schemes for mating the selected individuals may be designated. To date we have evaluated nine selection and mating schemes that represent a broad array (Figure 4). Reports may be generated for each replication, generation, or subline, or all three. Statistics generated by the model within and among sublines and for the population total include: (1) Additive, dominance, and total genetic variances, (2) Mean genotypic value, (3) Mean breeding value (Calculated by mating each individual to a standard check population with equal gene frequencies), (4) Genetic gain in breeding value (BV in generation n - BV in generation n-1),(5) Inbreeding coefficient ("F") (Correlation among alleles "identical by descent") , and (6) Variances for numbers 1-5 above over replicate runs.

RESULTS AND DISCUSSION

Comparison of Selection Schemes

The 9 selection schemes in Table 1 were compared based on their mean breeding values (Figure 4) and mean inbreeding coefficients (Figure 5) for 20 generations. A mainline population of 160 entries was generated using an h^2 = 0.2 and a degree of dominance = 0.5. Parents with the 40 best breeding values were included in the elite population. For the purposes of this comparison, parents were assigned at random to 4-parent sublines in generation O. The 4 selected parents in each generation were mated in 4-parent disconnected diallels with self-matings and 100 progeny were generated per mating.

Mating schemes 1, 2, and 3 assigned 4, 2, or 1 S1 progeny from 1, 2, or 4 parents, respectively, to a 4-parent subline. Breeding values plateaued after 6 to 8 generations (Figure 4) and inbreeding coefficients increased very rapidly to over 0.9 in 4 generations (Figure 5). The most extreme of these (Method 3) is effectively maintaining 40 selfed lines through time. Mean breeding values were calculated by mating each individual in each subline to an equilibrium population and averaging the individual breeding values for each subline and over sublines. Thus, mean breeding values illustrate the potential for gain in a production population formed in such a way that individuals from different sublines were not allowed to produce progenies. That is, mean breeding values do not reflect the reduced vigor and seed yields that will arise with increased coancestry in the breeding populations. It may, therefore, be wise not to use such extreme methods of selection and mating.

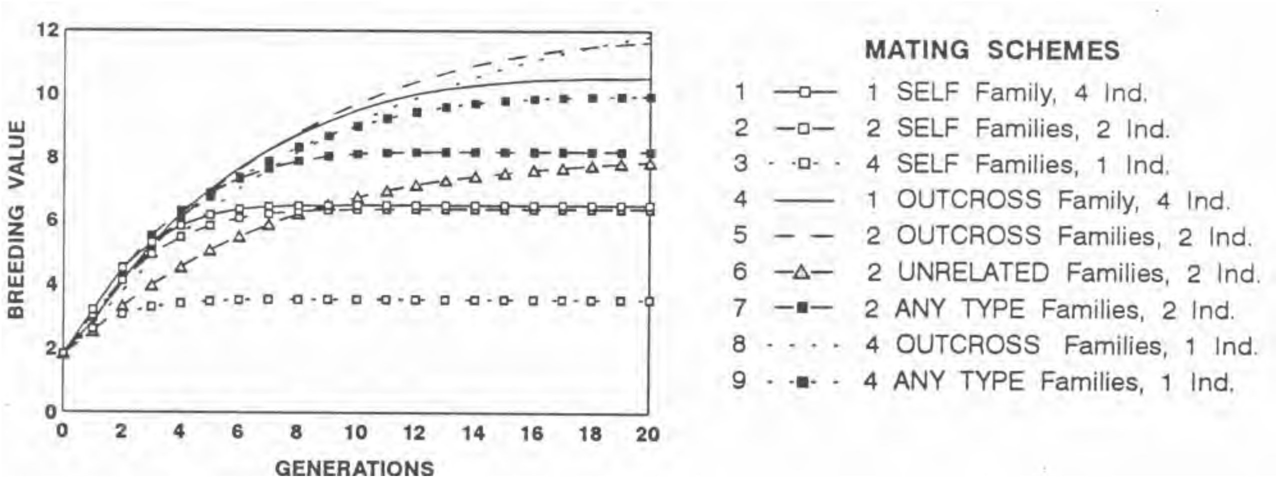

Figure 4. Mean breeding values for an elite population of 40 selected from a mainline of 160, selected by nine different methods.

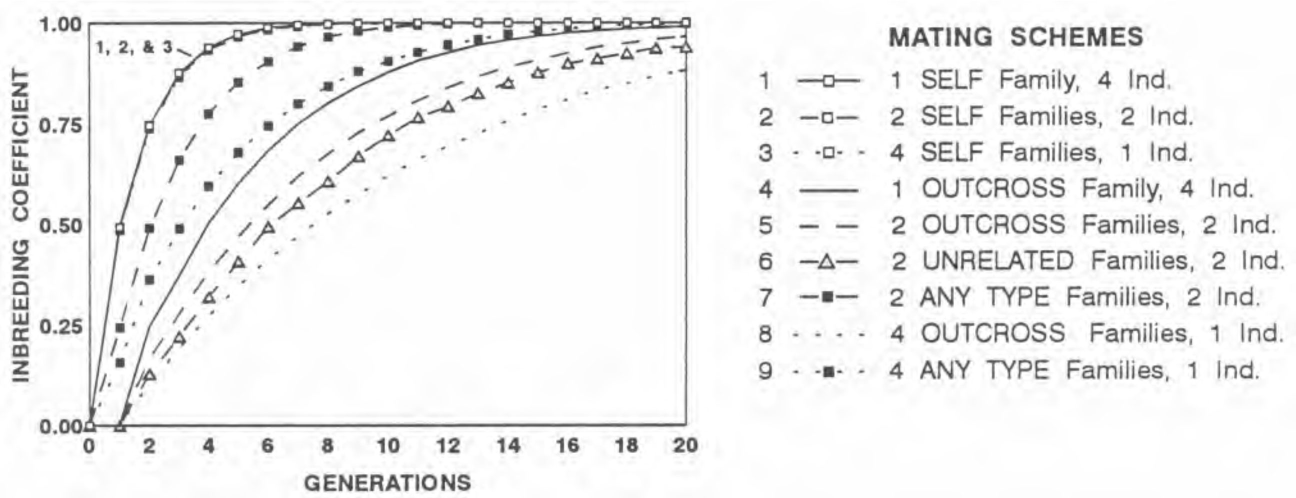

Figure 5. Mean inbreeding coefficients ("F") for an elite population of 40 selected from a mainline of 160, selected by **nine** different methods.

Methods 4, 5, and 8 represent 1, 2, and 4 of 6 full-sib families saved, excluding only the selfs. The full-sibs may also be related as half-sibs. Method 6 specifies unrelated families, implying that 2 individuals were chosen from each of crosses 1x2 and 3x4, where the parent number is the rank based on breeding values. Methods 7 and 9 are the simplest in that they permit selection of the 2 or 1 best individuals from each of the 2 or 4 best families of any type. Relationships are disregarded and selection from selfed lines is permitted.

Methods 7 and 9 have higher mean breeding values than 4, 5, and 8 through at least 3 generations, but plateau and reach higher values of "F" sooner. Since the Cooperative's plan is to enrich elite populations from the mainline populations, method 7 or 9 may be the best for selecting and mating in elite populations. The more conservative choice of these two methods is 9 which reached high values of "F" later (Figure 5). However, both the rate of increase in mean breeding value and "F" depend upon being able to produce selfed progenies for each entry. Since it is known that all parents do not produce selfed progenies with equal ease, we examined the more realistic assumption that only 50% of the parents in each generation would produce an adequate number of selfed progenies to permit selection among them. We assumed no correlation between breeding values and the ability to produce selfed progenies, and simply assigned a probability of 0.5 that a selfed progeny could be selected from a parent regardless of its breeding value. The differences in mean breeding values were insignificant through 10 generations. Furthermore, "F" increased more slowly when selfed progenies could be selected for only half the parents.

Comparison of Methods for Subline Assignment

We next examined the impact of method of assigning parents to sublines in generation 0. For the purposes of this examination, we used the same mainline populations of 160 and 40 generated as above. The methods of assignment to sublines were those described in the section on "Breeding Strategy Model", above. We used only mating method 9 with 50% selfing success in this analysis.

Production populations will be formed from the elite population in a way that prevents individuals from different sublines from producing progeny. Thus, selection will be largely, if not exclusively, within rather than among sublines. Given that, it should be desirable to favor assignment schemes that promote increased additive genetic variance within sublines. A comparison of mean breeding values for PAM, DAM, and RAM shows that mean breeding values increase more from generation 0 to 1 for DAM and RAM than for PAM (Figure 6). This difference arose because the within-subline additive genetic variances was greater for DAM and RAM in generation 0. This difference occurred in both population sizes, but was much smaller for the population size 40, which is a selected subset of the larger population of 160. Thus, positive assortment of parents to sublines can be expected to give smaller genetic gains than either dissasortative or random assignment.

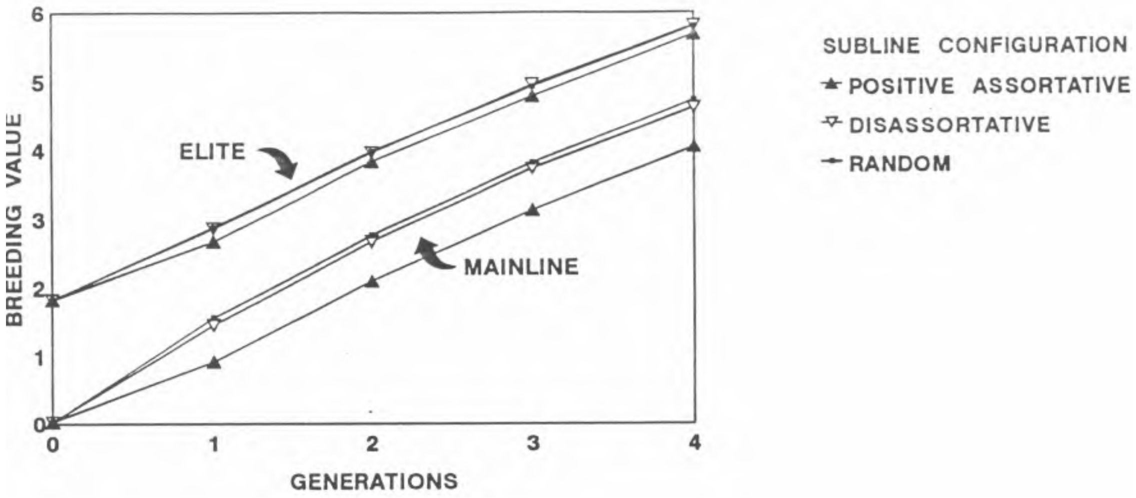

Figure 6. Mean breeding values for a mainline (160) and an elite (40) population

SUMMARY AND CONCLUSIONS

Computer modeling different aspects of a sublining breeding strategy suggests that: (1) selecting the best individual from the four best families (self or outcross) ranked on expected breeding values is an alternative that will give good genetic gains and result in relatively moderate rates of increase in coancestry within sublines; (2) assigning parents to sublines at random or disassortatively rather than by positive assortment will increase within-subline genetic variance and result in greater increases in breeding value for the first generation of selection and mating.

LITERATURE CITED

Falconer, D.S., 1989. *Introduction* to *quantitative genetics.* John Wiley & Sons, Inc., New York. 438 pp.

McKeand, S.E. and F.E. Bridgwater. 1992. Third-generation breeding strategy for the North Carolina State University-Industry Cooperative Tree Improvement Program. In: Proc. IUFRO Conf. S2.02-08, Breeding Tropical Trees. Solving tropical forest resource concerns through tree improvement, gene conservation and domestication of new species. Oct. 8-18, 1992. Cartagena and Cali, Colombia. 7pp.