Stochastic simulation of forest tree breeding populations: A users guide to POPSIM v2.0
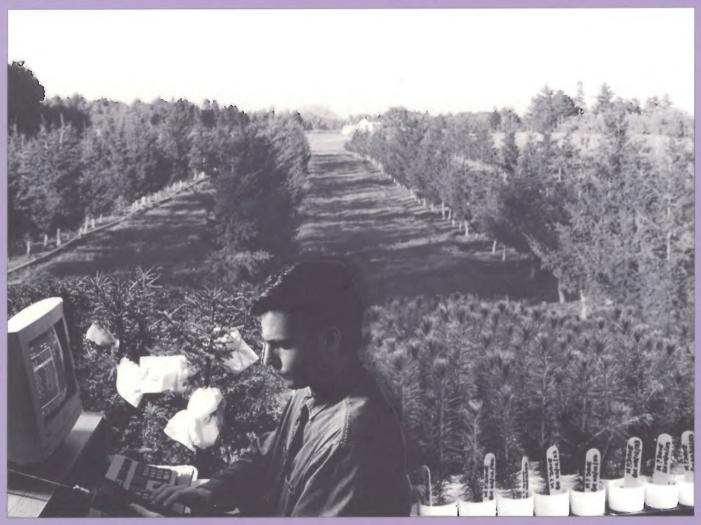
by
T.J. Mullin
and
Y.S. Park

Natural Resources
Canada

Ressources naturelles
Canada

Canadä

STOCHASTIC SIMULATION OF FOREST TREE

BREEDING POPULATIONS:

A Users Guide for POPSIM version 2.0

by

T.J. Mullin[1]

and

Y.S. Park

Information Report M - X - 195E

[1]Genesis Forest Science Canada Inc.
361 Valleydale Road
PO Box 1321
Truro, N.S. Canada B2N 5N2

Abstract

This report describes the design and application of a software, called POPSIM, for the simulation and analysis of tree breeding population managment strategies on modern personal computers. The purpose of the software is to provide breeders with a generalized decision-support tool to critically examine multi-generation breeding plans using a stochastic simulation approach. The theoretical basis for the simulation is given, together with a description of the options currently available to the user. Genetic effects are simulated according to an additive-dominance-epistasis model for a base population of trees that are mated according to a user-prescribed breeding plan. Individual progeny from the breeding plan are generated, simulating the recombination of genetic effects and the random environmental effects during field testing. Selection is carried out among the test progeny to assemble various kinds of production populations for deployment of improved stock, *e.g.*, seed orchard and clonal mixture, and to advance the breeding populations to the next cycle of breeding. Genetic variances are calculated directly from the simulated data and reported for each generation, along with the average level of inbreeding, effective population size, and the cumulative genetic gain realized by each type of production population. Examples are provided that illustrate the statistics generated by the program, and demonstrate the flexibility of the software to handle various approaches to management of genetic gain and diversity.

Résumé

Ce rapport décrit la structure et l'utilisation du programme POPSIM mis au point afin de simuler et d'analyser les stratégies de gestion de populations génétiques d'arbres sur les ordinateurs personnels d'aujourd'hui. Le but était de doter les sélectionneurs d'un outil global d'aide à la décision qui facilite l'examen critique des programmes de sélection sur plusieurs générations, à partir d'un modèle stochastique. Les fondements théoriques de la simulation sont exposés, avec les options offertes à l'utilisateur. Le programme reproduit les effets génétiques, selon un modèle addition-dominance-épistasie, sur une population d'arbres croisés en fonction d'un programme de sélection déterminé par l'utilisateur. Les lignées produites sont obtenues en simulant la recombinaison des effets génétiques et en intégrant les effets environnementaux aléatoires observés lors d'un test en milieu réel. La sélection s'opère au sein de la lignée soumise à ce test, de manière à composer divers types de populations de production pour le déploiement des sujets améliorés, p. ex. verger à graines et combinaison clonale, et dans le but de faire avancer les populations génétiques vers le cycle suivant de sélection. Les variances génétiques, calculées directement à partir des données simulées, sont fournies pour chaque génération, tout comme le niveau moyen d'autofécondation, la taille utile de la population et le gain génétique cumulatif présenté par chaque type de population de production. Le lecteur trouvera également des exemples des données statistiques produites par le logiciel et pourra apprécier la capacité qu'offre cet outil d'aborder de différentes façons la gestion du gain et de la diversité génétiques.

3

Table of Contents

Introduction

The prediction of genetic gain from one round of selection in a tree improvement program is a relatively straightforward calculation, requiring only appropriate estimates of heritability, selection intensity and the phenotypic variance of the selection units. Several assumptions are usually made regarding the normal distribution of variances and the absence of inbreeding. There are several papers in the literature demonstrating how this basic calculation can be applied to a wide variety of situations found in applied breeding programs (*e.g.*, Cotterill 1986; Cotterill and Jackson 1989; Matheson and Lindgren 1985; Namkoong *et al.* 1966; Shelbourne 1969, 1991).

Although such gain calculations are useful, and indeed essential for the evaluation of the economic returns from breeding, comparisons of breeding strategies based only on gain estimates from one round of selection are generally inadequate. More often the objective of an improvement program will be to maximize gain over time after several cycles of breeding. This objective will often be further constrained by the loss of genetic variance, accumulation of inbreeding, effective population size, and ultimately cost (Mullin 1994). While predicting the gain from a single round of selection may be relatively straightforward, resulting changes in population structure are less readily described, yet have a profound impact on the outcome of future breeding and selection. Algebraic expressions of such sequences of breeding and selection are, if not impossible, extremely complex.

Using computers to simulate genetic systems is not really new; the technique has been reported in the literature for at least 35 years (Fraser 1957). Even so, the application of simulation to investigate the outcome of breeding efforts with forest trees and other plant crops has been very limited, although recently several authors have incorporated simulation techniques in their comparisons of specific breeding strategies. In her analysis of the effects of positive assortative mating on cumulative gains and inbreeding, Mahalovich (1990) used a gene effects model, including additive and partial dominance effects at 50 loci. King and Johnson (1991, 1993) demonstrated the flexibility of computer simulation methods when they used a parameter-based stochastic model for additive genetic effects to consider gains and effective population size for five mating schemes for advanced-generation breeding of *Pinus radiata* D. Don in New Zealand. To date, the most significant progress in the simulation of artificial selection systems has been made by domestic animal breeders (*e.g.*, de Roo 1987; de Vries *et al.* 1990; Meyer and Smith 1990; Ruane and Thompson 1991).

Simulation models may be classified into two broad categories: (i) "stochastic" models, often referred to as Monte Carlo simulation, where random processes are mimicked by the generation of pseudo-random numbers, and (ii) "deterministic" models, where the outcome of processes is predicted through an algebraic mechanism. In genetic systems, the random recombination of alleles through sexual reproduction is a process that lends itself to simulation by stochastic methods. By repeating the simulation through several iterations, one can predict the mean outcome of stochastic events and describe the variance of outcomes about this mean.

Using computers to simulate stochastic processes, the investigator can set very broad limits on the complexity of model parameters and interactions, without resorting to highly theoretical or abstract mathematics (Levin 1969). Provided that genetic effects are simulated in a way that makes biological sense, the simulation of complex breeding plans requires elementary mathematics and is primarily an exercise in logic and efficient programming. A deterministic model of the same system will, at best, require extremely complex mathematics and will likely require simplification to the point that the model is no longer a realistic analogue (Kempthorne 1988).

A typical breeding plan describes numerous decision steps and employs a variety of limits, many of them somewhat arbitrary, to maintain diversity in the population while achieving as much gain as possible in each generation. Comparisons of breeding strategies based on gain estimates from one round of selection are generally inadequate, when the objective is usually the optimization of gain over time for several breeding cycles. Criteria for evaluation of breeding strategies are rarely limited to genetic gain alone, but will most likely include limits for reduction of genetic variance, accumulation of inbreeding, effective population size, and ultimately cost (Mullin 1994). Simulation techniques can make such complex comparisons relatively straightforward.

Previous simulation tools developed for tree improvement applications have been limited to specific problems and were designed for use on large, main-frame computers. This paper describes the design and application of a second-generation

software tool, called POPSIM, for the simulation and analysis of tree breeding population management strategies on modern personal computers. Our objective was to provide breeders with a generalized decision-support tool to critically examine multi-generation breeding plans. The theoretical basis for the simulation is given, together with a description of the options currently available to the user. Examples are provided that illustrate the statistics generated by the program, and demonstrate the flexibility of the tool to handle diverse approaches to management of genetic gain and diversity.

While this documentation is intended to be complete, we have assumed that the user is familiar with the terminology and procedures commonly used in the design and application of tree breeding strategies. Some users, particularly those without formal training in quantitative genetics, may find it useful to review concepts in an introductory textbook on forest tree breeding or crop improvement. Similarly, the guide assumes a general familiarity with basic procedures and configuration of DOS-compatible personal computers. Details on these procedures can be found in the original hardware and operating system documentation, and in a host of other references available in any computer store or library.

General Description of the POPSIM Simulation Software

The general approach used by the POPSIM simulator is to generate genetic and environmental effects for a base population of trees that are mated according to a user-defined breeding plan. Individual progeny from the breeding plan are generated, simulating the recombination of genetic effects and the random environmental effects during field testing. Selection is carried out among the test progeny to assemble various kinds of production populations for deployment of improved stock, and to advance the breeding population to the next cycle of breeding. Genetic variances are calculated directly from the simulated data and reported for each generation, together with the average level of inbreeding and effective population size. The cumulative genetic gain realized by each type of production population is presented as a percentage of the trait mean in the base breeding population.

Defining the phenotype of simulated trees

In this simulation, each tree's phenotype is considered to be the sum of independent genetic and environmental effects:

$$P = A + D + I + E \qquad [1]$$

where P is the phenotypic, or observed value for a given trait or an index of traits; A is the additive genetic effect of substituting one allele at a locus for another; D is the dominance genetic effect of within-locus allele effects which remain after subtracting additive effects, i.e., the intra-locus interaction between alleles; I is the epistatic genetic effect that arises from interactions among all alleles that affect expression of the trait; and E is the environmental effect.

The total phenotypic variation in a population is thus described as the sum of independent variances for each of these effects:

$$\sigma_P^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2 + \sigma_E^2 \qquad [2]$$

Simulating a base population of N trees is a matter of describing the mean and variance of each of these components in the population, and sampling from each distribution to produce the effects that together define the phenotype of each individual. If we assume that each of these effects is normally distributed, the random sampling from the distribution takes the form:

$$x_{ij} = \mu_j + r\sqrt{\sigma_{x_j}^2} \qquad [3]$$

where x_{ij} is a normal variate for the jth effect of the ith tree, with mean of μ_j and variance $\sigma_{x_j}^2$, and r is a normal deviate drawn from a distribution of random numbers with mean of 0 and variance of 1. From the generated data for N trees in the population, we can readily calculate estimates of the population mean $(\overline{X}_j)$ and variance $(\hat{\sigma}_X^2)$ for each of the genetic and environmental effects, using standard formulae:

$$\overline{X}_j = \frac{\sum_i x_{ij}}{N} \qquad [4]$$

and

$$\hat{\sigma}_{X_j}^2 = \frac{\left(\sum_i x_{ij}^2\right)^2 \Big/ N}{N-1} \qquad [5]$$

Simulating genetic recombination in offspring requires a description of the among-family and within-family inheritance for each genetic effect, as detailed below.

Additive effects

For additive genetic effects, the among-family component is the average additive effect of the female (f) and male (m) parents, $i.e.$, the mid-parent value, and the additive effect for each offspring (A_i') may be generated around this mean with a variance equivalent to the within-family portion of additive variance, $\hat{\sigma}_A^2$, adjusted by the expected reduction due to inbreeding of the parents (Dempfle 1990):

$$A_i' = \frac{A_f + A_m}{2} + r\sqrt{\frac{1 - \frac{1}{2}(F_f + F_m)}{2}\,\hat{\sigma}_A^2} \qquad [6]$$

where F_f and F_m are the coefficients of inbreeding for the female and male parents, respectively.

Dominance effects

The among-family dominance effect for a given female-male parent combination is drawn randomly from a normal distribution with mean equivalent to the dominance effect in the parent population $\overline{D}$ and a variance of $\frac{1}{4}\hat{\sigma}_D^2$. The dominance effects for individuals within this family (D_i') are then drawn randomly from a distribution with mean equal to the within-family dominance effect and variance $\frac{3}{4}\hat{\sigma}_D^2$. The dominance effect is corrected to account for inbreeding depression, using the regression of phenotypic value on the inbreeding coefficient (de Boer and van Arendonk 1992; Borralho 1994):

$$D_i' = \overline{D} + r_{fm}\sqrt{\frac{1}{4}\hat{\sigma}_D^2} + r_i\sqrt{\frac{3}{4}\hat{\sigma}_D^2} + bF_{fm}\hat{\sigma}_P \qquad [7]$$

where r_{fm} is a random normal deviate drawn for each fm family combination, r_i is a random normal deviate for each individual in the family, b is a regression coefficient expressing the reduction in phenotypic values in units of phenotypic standard deviation ($\hat{\sigma}_P$), and F_{fm} is the inbreeding coefficient of the family members. The regression coefficient may be calculated from experimental data by using a linear regression of inbreeding coefficients on corresponding standard deviation units of phenotypic values. For example, inbreeding depression on the mean of selfed progeny is often 15 to 30% lower than outcrossed progeny (Griffin and Cotterill

1988; Park and Fowler 1984). If the reduction in the mean of selfed progeny ($F = 0.5$) is one standard deviation below the phenotypic mean, the regression coefficient $b = -2.0$.

Epistasis effects

All progeny epistatic effects (I_i') are considered to be within-family, with mean and variance equal to that of the parent population. While the generation of the effect resembles the method used to assign random environmental effects, the epistasis effect remains constant for all individuals cloned from the same genotype:

$$I_i' = \overline{I} + r\sqrt{\hat{\sigma}_I^2} \qquad [8]$$

Environmental effects

The simulator produces all tree populations in the same "environment," so that environmental effect (E_{ik}') for the kth ramet of the ith genotype is drawn from a distribution with the same mean and variance as that specified for the original population:

$$E_{ik}' = \overline{E} + r\sqrt{\hat{\sigma}_E^2} \qquad [9]$$

Inbreeding and effective population size

Changes in genetic diversity of populations are represented by the average inbreeding and estimates of effective population size. The inbreeding coefficient is the probability that two genes in an individual are identical by descent and describes the proportion of homozygotic loci in excess of that found in an infinitely large, random-mating population. The inbreeding coefficient for each tree in the population is determined by analyzing the pedigree of parent trees to determine the degree of common ancestry. The calculation is fairly simple in small pedigrees, but quickly becomes very cumbersome in large, complex pedigrees. Fortunately, techniques for pedigree analysis have made tremendous advances and efficient algorithms have become available in just the past few years that make the calculation of inbreeding coefficients straightforward ($e.g.$, Meuwissen and Luo 1992; Tier 1990). The average inbreeding for the population is calculated simply as the average inbreeding coefficient for all trees.

In POPSIM, two approaches are used to determine the effective size of breeding and seed orchard populations, both based on the accumulation of gene correlations. The "inbreeding effective size" is defined by Falconer (1981) as the number of individuals that would give rise to the observed rate of inbreeding, if they bred in the manner of the idealized population. In POPSIM, we estimate the inbreeding effective population size, N_e, from the rate at which inbreeding accumulates, ΔF. Falconer (1981: equation 3.12) gives the inbreeding coefficient in any generation t, referring to the base population at t_o:

$$F_t = 1 - (1 - \Delta F)^t \qquad [10]$$

By rearranging this formula, we can calculate the average rate of change in inbreeding at a given generation, when the inbreeding coefficient is available from the pedigree record:

$$\Delta F = 1 - (1 - F_t)^{-t} \qquad [11]$$

Falconer (1981: equation 4.1) gives the rate of inbreeding as a function of inbreeding effective size as:

$$\Delta F = \frac{1}{2N_e} \qquad [12]$$

By rearrangement of this equation and substituting the previous equation for ΔF the calculation of inbreeding effective size is:

$$N_e = \frac{1}{2\left(1 - (1 - F_t)^{-t}\right)} \qquad [13]$$

Note that the inbreeding effective size of the base population is undefined and, if no inbreeding has accumulated, remains infinitely large in subsequent generations.

The concept of inbreeding effective size was developed originally with random-mating, unstructured populations in mind and is not readily expanded to subdivided populations due to differences in the accumulation of gene correlations among and within breeding groups. When the subdivision is complete with no genetic exchange, the correlation of genes from random individuals from different groups is zero. In this case, the N_e of the population is calculated as the N_e within groups, multiplied by the total number of groups, although this probably overestimates the actual inbreeding effective size (Chesser et al. 1993).

A second measure of effective population size calculated by the program is based on the concept of coancestry, which is defined as the probability that genes sampled from parents will be identical by descent and equivalent to the coefficient of inbreeding that would occur after mating. The accumulation of gene correlations is altered when parents are established in a seed orchard and progeny result from matings among individuals from different breeding groups, so it makes sense to evaluate the size of the orchard based on the coancestry of the parents, rather than their accumulated inbreeding. The average inbreeding coefficient of the progeny (F_{t+1}) from all possible crosses between parents, including self matings, is equivalent to the average coefficient of coancestry for the parents (f_t). Using a similar derivation as used for inbreeding effective size (N_e) we can then calculate the number of individuals that would give rise to the calculated average coefficient of inbreeding in the next generation if bred in the manner of an idealized population (N_s):

$$N_s = \frac{1}{2f_t} \qquad [14]$$

This is similar to an effective size parameter proposed by Burrows (1984) based on average pair-wise coancestry; whereas Burrows excluded self mating, we have included all possible pair-wise matings, including selfs. Lindgren has suggested calling this parameter "status effective number," referring to the status of a population at a particular moment (Dr. Dag Lindgren, pers. comm.) and has recommended its suitability as a measure of diversity in any population (not just seed orchards) as its value is not affected by population structure, it is never undefined, and is always less-than or equal to the census number, N. In POPSIM, effective size of orchard populations is always expressed as N_s, although the user is given the option to calculate the effective size of breeding populations as either N_s or the more traditional N_e.

Breeding and selection plans

The remainder of the simulation logic is a matter of mimicking the mating designs and evaluation procedures used in a given breeding plan and describing the distribution of effects in the resulting

populations. In its current version, POPSIM can simulate several operations and approaches used in typical breeding plans (*e.g.*, Fowler 1986; Park *et al.* 1993), as outlined below.

Sublining of the breeding population

Breeding populations may be subdivided into "subline" groups, with full-sib mating for generation advancement, *i.e.*, to produce progeny for selection of the next-generation breeding population, restricted to crosses among parents within sublines (van Buijtenen and Lowe 1979). As the accumulation of inbreeding is restricted to within sublines, crosses among orchard parents selected from different sublines will always be totally outcrossed. Subdividing the population also lessens the impact of genetic drift, as sampling of gametes within each subline is independent of that in the remainder of the population and different alleles become fixed in the different groups (Caballero 1994; Lacy 1987).

When sublining is requested, parent trees are assigned to sublines in the base generation, after production of all seed orchard populations, but before any breeding is carried out for generation advancement. In order to ensure genetic equality, the assignment of trees to sublines is normally random; however, when positive assortative mating is requested, the user can elect to first rank the parents by breeding value and then assign them uniformly across all sublines. Once formed, the sublines are managed independently and maintained at the same size over subsequent generations.

Polycross testing for ranking of parents on breeding values

If a breeding strategy calls for selection of parents for a "tested" orchard, or if assortative mating is to be used for generation advancement, the program will simulate a polycross test of all parents in the breeding population. The user declares the number of male parents represented in the pollen "mix" and the number of progeny to be tested for each tree in the breeding population. The pollen mix remains constant through all generations of breeding and is drawn from the same population distribution as the base population. Male parents are drawn at random from the mix with equal probability for each offspring, and genetic and environmental effects are generated as described earlier.

The family means from the polycross test estimate the breeding value of each tree in the population. Trees in each subline are ranked for selection as a tested orchard population, or for assortative mating.

Full-sib mating for generation advancement

The simulation assumes that all breeding for generation advancement will be based on full-sib crosses, using the same mating design across all sublines and all generations. Considerable flexibility is permitted in the description of the mating design, as virtually any number and combination of parents may be used, subject only to the array storage limitations imposed in the user's copy of the program. The simulator uses a mating design "template" that resembles the grid representation used by many authors to describe proposed mating designs. The user will be warned if any parents are omitted from the proposed test design, although this may be desired in some strategies to exclude the lowest ranking parents in assortative mating schemes. Warnings are also issued if attempts are made to produce duplicate or reciprocal crosses, as the simulation will not generate correct genetic effects for such crosses.

Positive assortative mating for generation advancement

When the user requests assortative mating, the parents in each subline are first sorted in descending order based on breeding value estimated from the simulated polycross test. The mating design template can then be used to increase the number of crosses among the better parents, or to cross parents with similar or dissimilar breeding values, or both.

Field progeny testing with specified family size, and the option for clonal replication

The program assumes that all progeny testing is performed in the same environment, with the variance of environmental effects identical to that of the base population, and that all selection for generation advancement will be carried out in these test plantations. The user declares the number of test genotypes to be evaluated. The user may also elect to replicate test genotypes by cloning, in order to increase the efficiency of individual genotype selection (Libby 1964; Mullin

11

and Park 1992; Shaw and Hood 1985), and for selection of genotypes based on clone means for deployment of clonal mixtures. In this case, clonal replicates (ramets) will be generated with identical genotypes, but independent environmental effects.

Selection options

The program simulates three selection methods, and different methods may be applied to the selection of the next breeding population and each of the various production populations. The simplest of these is referred to as "mass selection," where selection is based solely on the individual tree's phenotype or, in the case where clonal replication is used in the progeny tests, on the clone mean performance. So-called "two-stage" selection is performed by identifying a fixed number of best phenotypes from each of the best families. "Combined-index selection" is based on an individual index value (CI) family and individual performance by their respective heritabilities (Falconer 1981):

$$CI = h_w^2(\overline{X}_{ifm} - \overline{X}_{fm}) + h_F^2(\overline{X}_{fm} - \overline{X}) \quad [15]$$

where $\overline{X}_{ifm}$ is the mean (or individual observation) for the *i*th genotype in the *fm*th family; $\overline{X}_{fm}$ is the mean performance of the *fm*th full-sib family of size n; $\overline{X}$ is the overall mean; h_F^2 is the heritability of full-sib family means, and calculated as

$$h_F^2 = h^2 \frac{1 + (n-1)r}{1 + (n-1)t} \quad [16]$$

h_w^2 is the heritability of within-family deviations, calculated as

$$h_w^2 = h^2 \frac{(1-r)}{(1-t)} \quad [17]$$

h^2 is the narrow-sense heritability of individual observations; r is the genotypic intraclass correlation, which is $\frac{1}{2}$ for the full-sib families; t is the phenotypic intraclass correlation; and n is the family size.

Baker (1986) showed that for families of large size, $t = rh^2$ so in the case of full-sib families the relevant heritability formulae become:

$$h_F^2 = h^2 \frac{1 + \frac{1}{2}(n-1)}{1 + \frac{1}{2}(n-1)h^2} \quad [18]$$

and

$$h_w^2 = h^2 \frac{\frac{1}{2}}{1 - \frac{1}{2}h^2} \quad [19]$$

Restrictions on relatives

Restrictions on the use of relatives may be imposed both during selection of the breeding population, and assembly of the various kinds of production populations. The user can specify the maximum number of selections to be made from a single cross. Candidates for selection may be also be compared with those previously selected and the number of half-sib relatives restricted. For this purpose, full-sib relatives are also regarded as members of the half-sib family. In addition, the user may stipulate the maximum number of progeny from a given parent that can enter the next assembled population. If it happens that the user's restrictions can not be satisfied at a particular point in the simulation, the restrictions are relaxed only as much as necessary to complete the selection process, and a warning is issued to the user. At the beginning of each selection operation, the restriction limits are returned to those originally specified by the user.

Seed orchard populations

Seed orchard populations for production of improved seeds are assembled by selecting trees from each subline; the total number of trees in the orchard must be divisible by the number of sublines. As breeding occurs within sublines, selections from different sublines will not be related. If more than one tree is selected from each subline, coancestry restrictions may be applied to control relatedness of trees in the orchard. The average inbreeding coefficient of the selected orchard parents is given. The effective population size and genetic gain from the orchard are also reported, assuming equal contributions to the gamete pool and random mating among the orchard parents.

Deployment of full-sib family mixtures

The program also allows the user to select a mixture of the best full-sib families from those established in the progeny test. Such a mixture could then be reproduced by repeating the cross, and perhaps by using vegetative propagation to "bulk-up" sufficient numbers for field deployment, as suggested by Mullin and Park (1992). As for

seed orchards, the selected mixture must contain an equal number of families from each subline. As all tested crosses are made within sublines, inbreeding will accumulate within the families, although the maximum number of related families selected can be specified by the user. The program does not, at this point, permit crosses made between sublines for the purpose of testing outcrossed families.

Deployment of clonal mixtures

The user may also choose to simulate the selection of a clonal mixture, consisting of an equal number of genotypes from each subline, of those included in the progeny test. If clonal replication is used as part of the progeny test, the selection will be carried out on clone means, otherwise selection is based on individual phenotypes. Again, as these clones are selected from among crosses made within sublines where inbreeding will accumulate, the user may specify the maximum number of related clones to be included in the mixture.

Using the Software

In this section, we cover the hardware requirements to run POPSIM, and give instructions for installing the software and configuring your machine to run POPSIM under Windows™. We assume that you are generally familiar with your Intel®-based computer and DOS. You should consult your documentation for DOS, Windows, or your hardware user's manual if you are uncertain about any of the procedures covered in this guide.

We have adopted some typographical conventions to make the text easier to follow. Anything written by the system on your screen or on printed output will be shown here in `Courier` font, whereas your responses will be displayed using `Courier bold`.

Computer hardware requirements

POPSIM is written in FORTRAN and compiled with the Microsoft® FORTRAN PowerStation (Microsoft 1993) for execution as a 32-bit DOS-extended program under DOS or Windows™ operating systems. This compiler uses a DOS memory extender to make RAM above 640 K and free disk space available for data storage, allowing the simulator to use array storage for processing of very large

pedigrees. Interim progeny test data are written to a peripheral storage device, which can be a regular hard-disk or, if sufficient RAM is installed, to a RAM disk. The computer must have a 32-bit processor, *i.e.*, an Intel 80386 or better, and a minimum of 4 Mb of RAM is recommended.

POPSIM can be run on any personal computer using a 32-bit Intel® or compatible microprocessor (*i.e.*, 386® or better). It can be run under any of the various flavors of DOS operating systems (*i.e.*, MS-DOS®, PC-DOS®, or DR-DOS® at version 5.0 or higher), or in a DOS window running under Windows™3.1 or Windows 95™. Although hardware requirements are not technically restrictive, there are some compelling reasons to carefully consider the selection of hardware and operating system. Stochastic simulation is CPU intensive, requiring the generation of many random numbers, in addition to iterative evaluation of data. Even on today's powerful PCs, the assessment of a given scenario will normally be measured in terms of minutes or even hours, as opposed to milliseconds.

Installing the program files and examples

Begin installation by creating a new directory on your hard drive. For example, if you want to install POPSIM on drive D, you would enter:

```
C:\> d:
D:\> md popsim
D:\> cd popsim
D:\POPSIM>
```

Put the distribution diskette into drive A, and copy the files to the new directory:

```
D:\POPSIM> copy a:*.*
```

Starting POPSIM

If your POPSIM directory is your current default directory, you can start the program simply by entering:

```
D:\POPSIM> popsim
```

If you want to have your data stored in a different directory, you must modify your AUTOEXEC.BAT file so that the directory containing POPSIM.EXE and DOSXMSF.EXE is listed in your PATH statement (see your DOS documentation for details on

the PATH statement). Change to your data directory (for example, \SIMDATA) and start the program by entering the program name:

```
D:\POPSIM> cd \simdata
D:\SIMDATA> popsim
```

Optimizing your machine's startup configuration

Users who are running POPSIM for the first time should be able to use the program without any changes to their standard machine configuration. As a minimum, you must have an extended memory device driver loaded in your CONFIG.SYS file. Experienced users will probably want to make some simple changes to their CONFIG.SYS file that will significantly improve execution speed, by optimizing the use of RAM and storing temporary data on a RAM disk, rather than on the slower hard drive.

Making extended RAM available to POPSIM

POPSIM will not run in the 640 K of conventional memory that is normally available to DOS programs, and requires access to a "DOS memory extender" (distributed with the program) to make "virtual memory" available in the extended memory region above 1 Mb and, if necessary, on the computer's hard drive. If POPSIM will not load, verify that you have completed these two important steps:

(a) The CONFIG.SYS file must load memory drivers, in order to make extended memory (above 1Mb) and upper memory (between 640 K and 1 Mb) available. Most DOS installations at version 5.0 or higher will normally include such a driver as HIMEM.SYS and EMM386.EXE, or the user can install a third-party driver such as QEMM386, 386Max, etc. Once the memory manager is loaded, you can make portions of extended memory available for program execution, while other portions are reserved to set up a software disk cache or a RAM disk drive (see below).

(b) You must also install DOSXMSF.EXE, which is the actual DOS extender that allows a 32-bit program to run under MS-DOS. The easiest way is simply to copy the DOSXMSF.EXE file into the same directory as POPSIM.EXE, or you can copy the file to any directory that is in your DOS PATH environment variable.

Running POPSIM under Windows™ 3.1

As Windows™ 3.1 is a 16-bit operating system, POPSIM must run within a DOS window to execute as a 32-bit DOS-extended program. DOSXNT.386 is a Windows DPMI device driver that is necessary if you want to run a 32-bit FORTRAN application in an MS-DOS session under Windows. To install DOSXNT.386, first copy the file from the distribution diskette to your Windows directory, using a command such as:

```
C:\> copy a:\dosxnt.386 c:\windows\*.*
```

Now use a text editor to add a "device" statement to the [386Enh] section of your SYSTEM.INI file. If Windows is in C:\WINDOWS, this line should be added to the end of the [386Enh] section:

```
[386Enh]
 .
 .
 .
device=c:\windows\dosxnt.386
```

You can run POPSIM from a DOS session started under Windows, or as a program item icon in Program Manager. When setting up a program item, we strongly recommend that you specify a command line to run POPSIM from a "PIF" file; this will help ensure that Windows allocates memory properly to the DOS memory extender. A sample PIF file (POPSIM.PIF) has been provided on the distribution diskette and can be edited as necessary to specify the appropriate program and startup directories on your system. Further, we recommend that you ensure that the "Close Window on Exit" box is NOT checked; this will leave the window open with any error messages visible should your session encounter a fatal error. See your Windows™ user's guide for instructions on how to set up a new Program Item and modify a PIF file to start POPSIM from an icon.

Using additional RAM to improve program performance

POPSIM stores interim progeny data and scenario summaries on a peripheral device such as a hard drive. The access time to these data will be an important factor in determining the speed at which the simulation is completed. You may already be using a portion of your extended memory as a software cache, to speed up disk access by reducing the number of disk accesses necessary to read and write data. However, POPSIM will often use extended memory to greater advantage if it is set up as a RAM disk.

For fastest execution, the user should specify a drive letter for a RAM disk of sufficient size during the user interface prompts or in the response file controlling the simulation. The size necessary to store the data can be estimated as 26 K per 1,000 trees in selection plantations, plus 19 K per generation for storage of accumulation of iteration summaries. Refer to your operating system user's guide for details on installing a device driver in your CONFIG.SYS file to set up a RAM disk. A sample device statement to set up a 2 Mb RAM disk in extended memory under DOS 5.0 is given below:

```
device=c:\dos\ramdrive.sys 2048 /e
```

TUTORIAL #1: Using interactive control interface to run POPSIM

If you are running POPSIM for the first time or are only running a single scenario, you will probably enter your simulation control parameters through the user interface prompts. You will be given an opportunity to confirm and edit your choices before the simulation is actually started. Inappropriate responses and those that exceed the capability of the program will normally be trapped and hints given on how to make a legal response. The selected parameters are stored in a so-called "response file", under a name the user provides during the interactive dialogue. Experienced users may prefer to edit their response file directly with a text editor or word processor, especially if multiple scenarios are to be run while the machine is unattended. Instructions and hints on using response files to control POPSIM execution are given in a later section.

Start the program as described in the preceding sections. Press the <Enter> key to clear the opening screen and copyright notice, and display the prompt:

```
SELECT METHOD FOR USER INPUT ...

I will prompt you to define your simulation
scenario(s), or you can direct me to a
Response File.
Do you want to use a Response File? (Y/N): N
```

Since you are using the interactive prompt system, press the N key, then <Enter>. POPSIM will accept all alpha-character input in either upper case or lower case, but note that lower case characters are not converted automatically to upper case. You will now be prompted to provide a file name to store your responses to the interactive prompts:

```
I'll put your answers in a response file that
you can edit later if you like.
I'd suggest using a filename extension of
.RSP to help identify this file.
Please enter filename for this response file
> TUTORIAL.RSP
```

You must enter a valid name for a DOS file, *i.e.*, up to 8 characters followed by an optional extension of up to 3 additional characters. If the file already exists, POPSIM will ask you to confirm that you want to overwrite it. POPSIM will only handle data files created on your default directory, the directory where you started POPSIM. Do NOT attempt to direct POPSIM to other drives or directory paths, as POPSIM will attempt to interpret this as a standard file name, with unpredictable results.

Now you'll be asked:

```
How many scenarios do you want to run?: 1
```

Enter the number of scenarios you wish to run. You will be prompted to enter the control parameters for each scenario, and your choices will be stored in the response file that you just named. The simulation will begin immediately after you finish the entry of the last scenario. For this tutorial we'll only set up the control data for one scenario.

The first series of prompts requests information about the general program settings that will be in effect during the run. The first step is to specify where POPSIM will create its temporary files:

```
FIRST, LET'S SPECIFY THE SETTINGS FOR SYSTEM
I/O ...

Progeny and summary data are written to
temporary files as they are generated.
Storage requirements are 26 K per 1,000 trees
in selection plantations per generation, plus
19 K for each generation simulated. The
simulation program will run quite a bit faster
if a RAM Disk is available with sufficient
space.
By default, these files will be written to the
root directory of drive C:
Do you want to write these files to a different
drive letter? (Y/N): N
```

If you have a RAM disk drive set up in your CONFIG.SYS file, you will probably want to store the interim results there, rather than on the slower hard disk. If you are using a RAM drive, or wish to store interim results on a different device, answer Y to this question and you will be prompted to enter the device drive letter. Otherwise, answer N and the interim data will be stored in the root directory of drive C. In either case, the temporary files will be deleted on normal termination of the program.

Next you will be prompted:

```
Do you want output directed to the printer?
(Y/N): N
```

Answering Y will direct output to the printer on completion of each scenario. The printer can still be accessed by other programs while POPSIM is working on its calculations. Now you will be prompted:

```
Do you want output directed to disk files?
(Y/N): Y
```

If you answer Y, POPSIM will explain its two types of file output, and you will be able to specify either one or both; we will ask for both in this tutorial. As illustrated in the following dialogue, you will be asked to provide a "root" file name, and we will use TUTORIAL for our example. The output file for the tab-delimited "series data file" will be written to a file with this root name and a .TXT extension. The "scenario summary" files are also generated using this root name with a 3-digit extension referring to the scenario number. Again, these files will be written to the default directory; do NOT attempt to direct POPSIM to a different drive or path. The "scenario summary" written to disk is identical to that sent to the printer. The "series data file" is intended for exporting POPSIM output to other applications, and its format is documented in Appendix "A". Interpretation of these files is explained in a later section.

```
POPSIM produces 2 types of file output: the
"scenario summary" is identical to the normal
printed output and is intended for use with a
word processor; the "series data file" is a
tab-delimited text file for importing into a
spreadsheet or database application.

Do you want scenario summaries written to disk
files? (Y/N): Y
Do you want a series data file written to a disk
file? (Y/N): Y

Output will be directed to separate files for
each scenario, with filenames taking the form
'ROOT.nnn', where 'nnn' refers to the scenario
number, or 'ROOT.TXT' for the series data
file.

Please enter ROOT name for output files (max.
8 characters): TUTORIAL
```

If a series data file ('ROOT.TXT') already exists with this name, POPSIM will offer to append the new data to the end of the existing data, which is useful if you are restarting the program midway through a number of scenarios. Otherwise, you'll be asked to confirm that you want to overwrite the existing file.

Now you will ask for detailed progress reports during execution:

```
Do you want detailed progress and timing
reports? (Y/N): Y
```

Answering Y to this question will direct POPSIM to direct detailed information on simulation progress to the screen, including current scenario, iteration, generation, and execution times for individual components of the breeding cycle.

POPSIM now presents a summary of your system setting responses for your approval:

```
SYSTEM SETTINGS FOR THIS JOB ...

A total of 1 scenarios will be processed.
The control data have been stored in file:
TUTORIAL.RSP
Progeny and summary data written to drive C:
Output will NOT be sent to the printer.
Scenario summaries WILL be written to file(s):
TUTORIAL.nnn
Series data WILL be written to file:
TUTORIAL.TXT
Detailed progress and times WILL be
displayed.

ARE THE ABOVE SETTINGS CORRECT? (Y/N): Y
```

You now have a chance to review your responses to the above questions and modify them if they are not correct. If you answer N, you will be taken again through each question. If everything looks OK, answer Y and you will start a series of questions concerning the scenario you wish to simulate.

```
PLEASE SPECIFY MAIN SIMULATOR SETTINGS ...

Enter a seed number for random number
generator: 12345

Enter number of generations for simulation:
5

Enter number of iterations to be completed:
25
```

The simulation uses pseudo-random number generators to drive the many stochastic events over uniform and normal distributions; ran2 and gasdev, respectively, by Press *et al.* (1992). A valid integer must be specified as a "seed" number to start the pseudo-random sequence. Repeating a simulation with the same seed number on this, or any other Intel-based personal computer, will generate identical results. The maximum number of generations that can be simulated will depend on the parameters set up for your particular copy of the compiled program, as this affects the amount of memory available for array storage. In this tutorial, we have specified parameters within the limits of demonstration copies of the software that will run on virtually any machine capable of running the Windows 3.1 operating system. We will request the simulator to complete five generations of breeding. The simulation must be repeated several times to produce reliable average results, but repeatability will not likely improve much beyond 25 iterations.

Now you will be prompted to describe the base population to be simulated:

```
PLEASE SPECIFY VARIANCE STRUCTURE OF ORIGINAL
BREEDING POPULATION ...

Enter number of trees in breeding population:
50

Enter number of sublines in breeding
population: 5
```

All breeding and selection activity will be applied to a breeding population, which might be assembled at the beginning of the program by selection of plus trees from wild stands. Again, the maximum number of trees that can be specified is determined by the parameters set in your copy of the program. The trees will be assigned to equal-sized sublines, or breeding groups (van Buijtenen and Lowe 1979). If a polycross is to be performed, this is completed in the first generation prior to the assignment of parents to sublines. From then on, all breeding and selection work is carried out *within* the sublines. The number of sublines must divide evenly into the total number of trees in the breeding population. If the population is not to be subdivided into breeding groups, enter the value **1**.

You will now be asked your preference for expression of effective population size, either as the inbreeding effective number (N_e) or as the status effective number (N_s). See the earlier discussion about effective population size.

```
Effective size of the breeding population can
be expressed either as Inbreeding effective
size (Ne), or Status effective number (Ns).
NOTE: size of the seed orchard populations is
ALWAYS reported as Ns.
Shall I also use Ns for breeding population
effective size? (Y/N): Y
```

Now enter the phenotypic mean of the trait in the population and the adjustment for inbreeding depression:

```
Enter population mean for quantitative trait:
100

Adjustments will be made for the effect of
inbreeding depression by means of a standard-
ized coefficient, expressed in units of
phenotypic standard deviation. This coeffi-
cient is the rate at which the mean value of
a trait is reduced per .1 units of F, and would
normally take a value between 0 and -4.

Please enter a standardized coefficient for
inbreeding depression: 0
```

The phenotypic trait mean can be any real number. Adjustments for inbreeding depression are made by means of a standardized coefficient, *i.e.*, expressed in units of phenotypic standard deviation, which describes the rate at which the mean value of the trait will be reduced per 0.1 units of inbreeding (F), as described earlier. This coefficient would normally take a value between 0 and -4. If you are unsure how to interpret this coefficient, or do not have relevant data for your species, you should probably set this parameter to zero.

You are then taken through a series of prompts to describe the variance structure of the base population by entering the variance of the quantitative trait for all genetic and environmental effects, as real numbers.

```
Enter variance of additive effects: 100

Enter the variance of dominance effects: 25

Enter the variance of epistatic effects: 0

Enter the variance of environmental effects:
375
```

The program will only consider one metric trait, although this could be an index value combining two or more traits. Similar prompts will request the variance for additive, dominance, epistasis, and environmental effects. Heritability of traits in the starting population will be calculated as:

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_D^2 + \sigma_I^2 + \sigma_E^2} \qquad [20]$$

and

$$H^2 = \frac{\sigma_A^2 + \sigma_D^2 + \sigma_I^2}{\sigma_A^2 + \sigma_D^2 + \sigma_I^2 + \sigma_E^2} \qquad [21]$$

where h^2 and H^2 are heritabilities in the narrow sense and broad sense, respectively.

POPSIM only conducts a polycross test of trees in the breeding population if you request either positive assortative mating for generation advancement, or establishment of a "tested" seed orchard. The program will request details on the type of polycross test to be conducted, regardless of whether or not the test is required according to the specifications given later in the dialogue.

```
Enter number of males represented in polymix:
20

Enter number of polycross progeny per tree: 50
```

The program uses a common pollen mix for testing of all parent trees in the starting and all subsequent breeding populations. The males represented in this mix are drawn from a population with a variance structure equivalent to that of the initial breeding population and their number is defined by the user. Specifying larger numbers of test progeny will improve the reliability of breeding value estimates, but represents a greater commitment of testing resources.

POPSIM then moves on to request information on the mating design and specifications for advanced-generation testing:

```
PLEASE SPECIFY DESIGN FOR ADVANCED-
GENERATION TESTS ...

Do you want to perform assortative mating? (Y/
N): Y
In that case, I can assign the parents RANDOMLY
to sublines, or UNIFORMLY based on breeding
value.
Do you want assignment to be UNIFORM? (Y/N):
Y

How many crosses will we make per subline? :
10
```

As stated in the earlier section covering the simulation theory, the assignment of trees to sublines is normally random; however, when positive assortative mating is requested, the user can elect to rank the parents by breeding value first and then assign them uniformly across all sublines. Either way, once formed, the sublines are managed independently and maintained at the same size over subsequent generations.

The mating design routine in POPSIM uses a "template" representing all of the crosses to be carried out within each subline. The trees in the design template are numbered 1 through however many trees are in each subline. When assortative mating is requested, the trees in each subline are sorted from best to worst, i.e., "positive" assortative mating, using the estimates of breeding value generated by the polycross. In this case, the parent numbers on the design template correspond to the ranking of their breeding values; otherwise, the trees are assigned to the mating template in random order. Examples of templates for both nonassortative (random order) and assortative mating are given in Figure 1. There is tremendous flexibility permitted in the declaration of mating design templates so that virtually any mating design can be accommodated, within the array storage limitations of your copy of the program, although the same template must be applied to all sublines in the breeding population.

In this tutorial, we will use the positive assortative mating template shown in Figure 1 to make ten crosses per subline. We have taken advantage of the ranking of the parents to ensure that at least one of the top five parents is represented in each of the crosses made, and our testing effort will be concen-

"Non-assortative" pair mating　　　　　　　"Positive assortative" pair mating

Figure 1. Sample templates for "non-assortative" and "assortative" mating designs

trated on crosses between the higher ranked parents. We are then taken through a series of prompts to enter the ten female-male combinations represented in the design template. Enter the grid references for each parent, separated by a space or comma:
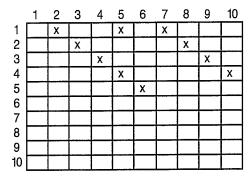
```
Please enter the number of the female
parent, followed by the male parent, for
each of the 10 crosses per subline ...

Female, male for cross #  1 = 1 2
Female, male for cross #  2 = 1 5
Female, male for cross #  3 = 1,7
Female, male for cross #  4 = 2,3
Female, male for cross #  5 = 2 8
Female, male for cross #  6 = 3 4
Female, male for cross #  7 = 3 9
Female, male for cross #  8 = 4,5
Female, male for cross #  9 = 4,10
Female, male for cross # 10 = 5 6
```

If self matings, duplicate or reciprocal crosses are requested, you will be warned and asked to confirm your selection. Note, however, that reciprocal effects are NOT simulated and that dominance effects will NOT be generated properly for duplicate crosses. When you have entered parent combinations for all of the crosses in the template, you will receive a warning if any of the parents have not been used at least once in the mating design. You will then be prompted:

```
Enter number of genotypes tested per cross: 100
```

This defines the family size for each full-sib family tested in the advanced-generation selection plantations. If clonal replication is to be used in the

progeny test, this number refers to the number of tested "ortets." The number of "ramets" (cloned individuals) to be tested of each genotype is then specified. If clonal replication is not desired, enter the value 1.

```
Enter number of ramets per genotype: 5
```

You are now asked to review your responses for each parameter:

```
MAIN PARAMETERS FOR SCENARIO # 1 of  1
 1. Seed for random number generator . . . . . . . . . 12345
 2. Number of generations for simulation . . . . . . . . . 5
 3. Number of iterations . . . . . . . . . . . . . . 25
 4. Number of trees in breeding population . . . . . . . 50
 5. Number of sublines in breeding population . . . . . . 5
 6. Effective size of breeding population expressed . . . Ns
 7. Mean of trait in base population . . . . . . . . . 100.0
 8. Rate of inbreeding depression . . . . . . . . . . .0000
 9. Additive variance . . . . . . . . . . . . . . . .100.0
10. Dominance variance . . . . . . . . . . . . . . .25.00
11. Epistasis variance . . . . . . . . . . . . . . . .0000
12. Environmental variance . . . . . . . . . . . . .375.0
       Heritabilities:   Narrow-sense:  .200  Broad-sense: .250
13. Number of males represented in polymix . . . . . . . 20
14. Number of polycross progeny tested per parent . . . . 50
15. PAM design (Uniform sublines), totaling    50 crosses.
16. Number of genotypes tested per cross . . . . . . . .100
17. Number of ramets cloned per genotype . . . . . . . . 5
          ARE THE ABOVE SETTINGS CORRECT? (Y/N) : Y
```

If you made any errors entering these parameters, now is your chance to correct them. To change anything, first answer N to the prompt; you will then be given an opportunity to specify which item you want to go back and change. If you return to item 15, you will have to re-enter the entire mating design template. When all of the responses are correct, answer Y to proceed to the definition of advanced-generation selection.

The next step is to specify the selection methods that will be applied to assemble the next breeding population (generation advancement), and each of the desired production populations.

```
NOW, PLEASE SPECIFY SELECTION METHODS ...

Selection methods available for generation
advancement:
        0. Random (no phenotypic selection)
        1. Mass selection (best trees)
        2. Two-stage selection (best trees in
           best families)
        3. Combined index selection

Enter number corresponding to desired method:
3

Do you want to restrict the number of genotypes
selected from a given cross? (Y/N):

What is the maximum number of genotypes
selected per cross?: 1

Do you want to restrict the number of
previously selected genotypes that can be
either full-sib or half-sib relatives of any
new selection? (Y/N): N

Do you want to restrict the number of progeny
that can be contributed by any single
parent? (Y/N): Y

What is the maximum number of progeny per
parent (self counted as 2)?: 3
```

Note that the method chosen will only be used for selection of the next generation of the breeding population; selection methods used for production populations are declared separately. The dialogue varies, depending on the selection method chosen. In this example, we have made ten crosses in each subline, and we have chosen to select exactly one tree from each cross to go into the next breeding population. The next prompts are used to restrict the number of full- and half-sib relatives that are permitted in the next generation, and how many progeny can be contributed by any given parent. For most selection methods, you can choose whether or not the number of relatives is controlled, although in the case of two-stage family selection, you are required to specify the number of full-sibs that are to be selected in each of the best families. A word of caution: the results of selection restrictions can be difficult to predict and it is easy to specify restrictions that cannot possibly be satisfied. In these situations, POPSIM will relax the restrictions as much as necessary to allow continuation of the simulation and will warn you that this has been done. We suggest that you not restrict both the numbers of half-sib relatives AND number of progeny contributed by a given parent, in the same scenario.

If you are only interested in following progress made in the breeding population, there is no need to have the simulation assemble production populations. However, POPSIM offers the opportunity to simulate the genetic and environmental effects represented by your choice of untested and polycross-tested seed orchards, repeat crossing and deployment of full-sib crosses, or clonal deployment of clones selected from the full-sib progeny tests. The prompts for selection methods are similar to those used for advanced-generation selection, but apply only to the specified production population. In all cases, the number of units selected for a production population must be a multiple of the number of sublines managed in the breeding population. You will be asked to specify the number of trees and any restrictions on relatedness among selected trees

```
PLEASE SPECIFY YOUR DESIRED PRODUCTION
POPULATIONS ...

Do you want to select for an UNTESTED seed
orchard? (Y/N): Y

Do you want to select for a PX-TESTED seed
orchard? (Y/N): Y

Selection methods available for untested
orchards are:
    1. Mass selection (best trees)
    3. Combined index selection
        (Note: In generation 0, we ALWAYS
         use mass selection)

Enter number corresponding to desired method:
3

Enter number of selected trees in seed
orchard: 10

Do you want to restrict the relatedness of
orchard trees selected from
the same subline? (Y/N): Y

Do you want to restrict the number of orchard
parents from a given cross?
(Y/N): Y

What is the maximum number of orchard
selections from the same cross?: 1

Do you want to restrict the number of previous
orchard selections that can be either FS or
HS relatives of any new selection? (Y/N): N

Do you want to restrict the number of
selections from a common parent?
(Y/N): Y

What is the maximum number from a common
parent?: 2
```

In our example, we have requested that both un-tested and tested orchards be selected with 10 trees. Combined index selection will be used as the basis for selection of "untested" orchards. The "PX-tested" orchard is, of course, selected using the breeding values determined from the polycross test. As in the situation with selection of the next breeding population, we can apply similar restrictions to the relatedness of trees permitted to enter the seed orchard. The same restrictions will be applied to both the "untested" and "PX-tested" orchard. Now we will proceed to request selection of mixtures for deployment of FS families and tested clones:

```
Do you want to select full-sib families from
among those tested for
generation advancement, to deploy as a family
mixture? (Y/N): Y

How many families are to be represented in
this mixture?: 10

Do you want to restrict the number of
previously selected families that can be
half-sib relatives of any new selection? (Y/
N): N

Do you want to restrict the number of families
from a common parent? (Y/N): Y

What is the maximum number of families from
a common parent?: 2

Do you want to select individual clones from
among those tested for
generation advancement, to deploy as a clonal
mixture? (Y/N): Y

Enter number of selected clones in mixture:
10

Do you want to restrict the relatedness of
clones in the mixture? (Y/N): Y

Do you want to restrict the number of clones
from a given cross? (Y/N): Y

What is the maximum number of clones from any
given cross?: 2

Do you want to restrict the number of previous
clone selections that can
be either full-sib OR half-sib relatives of
any new selection? (Y/N): N

Do you want to restrict number of selections
from a common parent? (Y/N): Y

What is the maximum number of selections from
a common parent? : 3
```

As we will only require one selection per subline, none of the selections will be related and POPSIM does not bother to ask for constraints on related-ness. We will also select the best full-sib family from the crosses made in each subline to be deployed as a family mixture. This deployment could be accomplished by repeat crossing in a control-pollinated seed orchard, or by means of vegetative multiplication. As with the seed orchards in our example, only one family must be selected from each subline, so POPSIM does not ask for con-straints on relatedness of selected families. We will also choose the best two unrelated clones from the crosses tested for each subline, for clonal deployment. Here we have a situation where related clones could be selected, so POPSIM prompts for constraints on the permitted related-ness.

As before, you are given an opportunity to review the selection methods specified for the advance selection and production populations and make corrections as necessary.

```
SELECTION PARAMETERS FOR SCENARIO # 1 of 1

1. Selection of breeding population:
   Combined index selection with up
   to    1 trees per FS family and with no
   restriction on numbers of HS      rela-
   tives, and up to  3 progeny per parent.
2. Selection of seed orchard populations
   UNTESTED orchards selected by
   combined index selection,
   and TESTED orchards based on PX-test,
   to select   10 trees, with up to    1 per
   cross and 32000 half-sib relatives, and
   up to    2 from a common parent.
3. A mixture of the best   10 families will
   be selected with no restriction on number
   of HS relatives, and up to     2 from a
   common parent.
4. A mixture of the best   10 clones will be
   selected with up to  2 per cross and 32000
   half-sib relatives, and up to    3 from
   a common parent.

ARE THE ABOVE SETTINGS CORRECT? (Y/N) : Y
```

Once you have completed the entry of control parameters for the requested number of sce-narios, the program will proceed with the simula-tion. Progress will be reported on the screen, with the level of detail requested earlier. When all simulations are completed, the output is directed to the file and/or printer specified, all temporary disk files are deleted, and the following message appears to prompt the user to terminate the pro-gram:

```
ALL DONE!  Press <ENTER> to exit ...
```

Interpretation and use of the output is discussed in detail after tutorial #2 which demonstrates how to run a scenario directly from a response file.

TUTORIAL #2: Controlling POPSIM with a response file

POPSIM can carry out the simulation of several different scenarios during the same run using a "response" file to store the control data. There are two ways to build a response file. The easiest method for a small number of scenarios, or as a way to generate a sample file according to the required format, is simply to use the user interface; you will be prompted to provide the name for a response file where POPSIM will store your parameters. After POPSIM finishes the interface dialogue, a file is written using the name provided by the user. You can use any legal file name, but we suggest using an extension of .RSP to distinguish the file as a response file. If you have completed Tutorial #1, there will be a response file called TUTORIAL.RSP in your default directory. As an alternative, you can create your own response file, or modify an existing one, using any DOS text editor.

A sample program-generated response file is included on the distribution diskette under the filename EXAMPLE.RSP, and a portion of the file is illustrated in Figure 2. This file shows the control parameters for the same scenario described in Tutorial #1, although the "root" name for the output files has been given as "EXAMPLE". The first block in the response file (to the line of asterisks) contains general program information, such as the number of scenarios to be run, and locations of temporary data storage and output devices. Then, for each scenario, a separate block describes the simulator settings, structure of the initial breeding population, testing and mating designs, selection methods for generation advancement and deployment options.

When the response file is generated by the program, POPSIM will annotate the data lines with descriptive information on variable names and values; this will be helpful if you want a guide to writing your own response files, without going through the tedious question-and-answer routine of the user interface. The portion to the left of the "equals" sign contains the data that are actually read and used by the program, while the portion to the right is descrip-

tive information. You can examine the file with your text editor to see how the response file should look, and modify it directly as required to change parameters or add additional scenarios.

Any time that a response file is used to control the program, POPSIM will first check the validity of the control data for all scenarios. If any errors are found, messages are generated explaining the problem and the simulation is terminated. Use your editor to correct the problems, then try running the program again. Except for the three file names specified in the first section describing job control, all data in the ten columns to the left of the "equals" sign must be numeric. Note also that the value "32000" has a special meaning to POPSIM, signifying that a selection restriction is to be ignored.

Interpreting the Output from POPSIM

Most often, users will direct POPSIM to "print" the output to an ASCII file. POPSIM writes output from multiple scenarios in separate output files; that way, if your system is halted for any reason, you can resume execution after the last scenario completed. POPSIM will ask you what scenario number to start at, so you do not have to modify the response file to resume execution if the initial run was interrupted. These output files are generated with a user-supplied "root" filename (up to 8 characters) and a program generated extension. So output files take the form ROOT.nnn, where nnn is a three-digit number corresponding to the scenario number.

Any word processor can be used to format an output file for printing. If the default document settings for your word processor are 8.5" x 11" paper with 1" margins, generally the only formatting that will be required is to specify a 10-point (12 cpi) mono-spaced font and to perform a global search-and-replace to change all [hard-return] to [hard-page]. Most word processors have a macro facility that makes it very easy to automate the required formatting prior to printing. If the output is to be made directly by POPSIM to the printer, be sure to set the default font on your printer to a mono-spaced typeface at 10 points (12 cpi) so that output will be properly formatted. If you only need to scan the output, this can be done using any text editor or "viewer."

```
        1 = NSCEN     - number of scenarios to be run
C:\PROGENY = PROGENY  - path\filename for temporary progeny data
C:\SUMMARY = SUMMARY  - path\filename for temporary summary data
        0 = IPRN      - printer output (0=No, 1=LPT1, 2=LPT2)
        3 = IFILE     - file output (0=No,1=Summaries,2=Series,3=Both)
EXAMPLE    = ROOT      - output file ROOT name
        1 = IS        - detailed timing reports (0=No, 1=Yes)
****************************************************************
        1 = KSCEN     - scenario number
    12345 = IISEED    - random seed number
        5 = NGEN      - number of generations
       25 = NITER     - number of simulation iterations
       50 = NT        - number of trees in breeding pop'n
        5 = NS        - number of sublines
        1 = ISIZE     - effective size: 0 = Ne, 1 = Ns
100.00000 = EFFBAR(4) - mean of trait in base pop'n
  .00000 = BID       - rate of inbreeding depression
100.00000 = EFFVAR(1) - additive variance
 25.00000 = EFFVAR(2) - dominance variance
  .00000 = EFFVAR(3) - epistasic variance
375.00000 = EFFVAR(4) - environmental variance
       20 = NP        - number of pollen parents
       50 = NX        - number of polycross progeny per tree
        2 = ISORT     - PAM: 0=none, 1=random, 2=uniform
       10 = NXS       - number of controlled crosses per subline
   1    2 = MATEF and MATEM for cross #  1
   1    5 = MATEF and MATEM for cross #  2
   1    7 = MATEF and MATEM for cross #  3
   2    3 = MATEF and MATEM for cross #  4
   2    8 = MATEF and MATEM for cross #  5
   3    4 = MATEF and MATEM for cross #  6
   3    9 = MATEF and MATEM for cross #  7
   4    5 = MATEF and MATEM for cross #  8
   4   10 = MATEF and MATEM for cross #  9
   5    6 = MATEF and MATEM for cross # 10
      100 = NC        - number of genotypes tested per cross
        5 = NR        - number of ramets tested per genotype
        3 = IBPSEL    - selection method for breeding pop'ns
        1 = NBPFS     - (maximum) number per FS cross for BP selection
    32000 = NBPHS     - maximum number HS relatives for BP selection
        3 = NBPPC     - maximum number progeny/parent for BP selection
        1 = IUNTST    - select untested orchard = 1, else 0
        1 = IPXTST    - select PX-tested orchard = 1, else 0
        3 = ISOSEL    - selection method for untested orchards
       10 = NSSO      - total number of selections in seed orchards
        1 = NSOFS     - maximum number per FS cross for SO selection
    32000 = NSOHS     - maximum number HS relatives for SO selection
        2 = NSOPC     - maximum number progeny/parent for SO selection
        1 = IFAMSEL   - select FS family mixture = 1, else 0
       10 = NFAM      - number of families in deployed mixture
    32000 = NFAMHS    - maximum number HS relatives in mixture
        2 = NFAMPC    - maximum number families with common parent
        1 = ICLNSEL   - select clonal mixture = 1, else 0
       10 = NCLN      - number of clones in deployed mixture
        2 = NCLNFS    - maximum number clones from any given cross
    32000 = NCLNHS    - maximum number HS relatives in mixture
        3 = NCLNPC    - maximum number clones with common parent
****************************************************************
```

Figure 2. Portion of POPSIM-generated response file EXAMPLE.RSP

The summary of scenario control settings

The first page of the output (Figure 3) gives a summary of the job and scenario settings. This should look familiar, as the presentation is similar to that used to summarize control data during the interactive input procedure. In our example, we ran a single scenario from a response file called EXAMPLE.RSP (shown in Figure 2), for a total of five generations, and repeated the simulation 25 times. The structure of the base population is given and the narrow- and broad-sense heritabilities are reported. Averaged over the 25 iterations, the mean and variance of each genetic and environmental effect for the breeding population in generation 0 will be very close to those specified by the user for the base population, but there will likely be a small difference due to the stochastic nature of the simulation. All of the control data for testing and selection are reproduced on this summary page, except for the actual mating design used for generation advancement. The 25 iterations in this example required about 17 minutes to complete.

The summary for each generation

The subsequent pages give the average results and standard deviations for all iterations simulated, one page for each generation. Although in our example the simulation was continued for five generations, we present and discuss here only the results for generation 5, illustrated in Figure 4.

All scenarios simulated by POPSIM will report a summary of genetic parameters in the breeding population. By the end of five generations in our example, the mean of the additive, dominance, and environmental effects in the breeding population have increased to 46.57, 17.80, and 118.0, yielding an average phenotypic effect, i.e., sum of all effects, of 182.4. On the other hand, the variances of these effects have been substantially reduced. The additive variance, that started at 100 in the base population, has decreased to only 46.57. No epistasis was specified in this example, in effect specifying an additive-dominance genetic model, and this effect continues to be reported with a mean and variance of zero. Also, in generation 5, selected progeny now originate from related parents in the sublines, so that inbreeding has accumulated within the sublines with an average F=0.1723 and the effective population size (status number) is reduced to 10.36.

When requested by the user, the output will also include descriptions of the various types of production populations: untested and polycross-tested seed orchards, full-sib family mixtures, and clonal mixtures. Statistics similar to those reported for the breeding population are given, i.e., the mean and variances of effects, inbreeding and, in the case of seed orchards, effective population size. Additionally, when describing production populations, POPSIM also reports the "cumulative gain percent" expected from deployment. For seed orchards, this represents the cumulative increase in the mean of additive effects, as progeny raised from orchard seeds will capture gain only from additive effects, divided by the trait mean in the base population and multiplied by 100. Note that when the trait mean is set initially to zero, the simulation will proceed normally, but the reported genetic gain will be forced to zero.

A similar procedure is used to calculate gain from full-sib family mixtures. Here, the deployment of seedlings or clonal propagules from selected full-sib families will capture gain from selection for both general and specific combining ability effects, and the cumulative genetic gain will include the mean increase in both additive and dominance effects. Similarly, for deployment of a mixture of selected clones, the cumulative gain includes the mean increase in all genetic effects, i.e., the sum of the additive, dominance, and epistasis effects.

Some Limitations and Areas for Future Development

POPSIM can be the basis for investigation of many aspects of population management, but the program is considered a "work-in-progress." Although considerable flexibility exists in the current version of the software, it is expected that ideas from the user base will result in the introduction of additional features as the software is applied to evaluate real-life breeding proposals. POPSIM was designed to give insight into the effectiveness of various population-management procedures and to guide the design of operational breeding programs that must balance the quest for genetic gain against the requirement to maintain diversity in plantations of improved stock. However, all simulation models have limitations and the desire to produce a flexible tool is hampered by some important practical problems.

```
                 FOREST TREE BREEDING POPULATION SIMULATOR
                              Version 2.0.15

          Copyright (C) 1994-1995, Genesis Forest Science Canada Inc.

CONTROL DATA FOR SCENARIO # 1 OF  1
   The user has supplied control data in file: EXAMPLE.RSP
   Progeny and summary data written to drive C:
   Output will NOT be sent to the printer.
   Scenario summaries WILL be written to file EXAMPLE.001
   Series data WILL be written to file EXAMPLE.TXT
   Detailed progress and times WILL be displayed.
MAIN SIMULATOR SETTINGS ...
   1. Seed for random number generator . . . . . . . . . . .  12345
   2. Number of generations for simulation . . . . . . . . . .    5
   3. Number of iterations . . . . . . . . . . . . . . . . .     25
STRUCTURE OF BASE POPULATION (Generation 0) ...
   4. Number of trees in breeding population . . . . . . . .     50
   5. Number of sublines in breeding population . . . . . . .     5
   6. Effective size of breeding populations expressed as  . .   Ns
   7. Mean of trait in base population . . . . . . . . . . .  100.0
   8. Standardized rate of inbreeding depression . . . . . .  .0000
   9. Additive variance in base population . . . . . . . . .  100.0
  10. Dominance variance in base population  . . . . . . . .  25.00
  11. Epistatic variance in base population  . . . . . . . .  .0000
  12. Environmental variance . . . . . . . . . . . . . . . .  375.0
      Heritabilities:    Narrow-sense:   .200  Broad-sense:   .250
DESIGN FOR POLYCROSS TESTING ...
  13. Number of males represented in polymix . . . . . . . .     20
  14. Number of polycross progeny tested per parent  . . . . .   50
DESIGN FOR GENERATION ADVANCEMENT PLANTATIONS . ..
  15. PAM design (Uniform sublines) totalling   50 crosses.
  16. Number of genotypes tested per cross . . . . . . . . . .  100
  17. Number of ramets cloned per genotype . . . . . . . . . .    5
SELECTION METHODS ...
  18. Selection of breeding population:
         Combined index selection
         with up to    1 trees per FS family and32000 HS relatives,
         and up to    3 progeny per parent.
  19. Selection of seed orchard populations:
         UNTESTED orchards selected by combined index selection,
         and TESTED orchards based on PX-test,
         to select  10 trees,
         with up to    1 per cross and 32000 HS relatives,
         and up to    2 from a common parent.
  20. Deployment of full-sib family mixture:
         The best  10 families will be selected
         with up to 32000 half-sib relatives,
         and up to    2 from a common parent.
  21. Deployment of clonal mixture:
         The best  10 clones will be selected
         with up to    2 per cross and 32000 half-sib relatives,
         and up to    2 from a common parent.
TIMING:
   First iteration started:  1995.02.22 at 13:15:36.24
   Last iteration completed: 1995.02.22 at 13:32:39.44
```

Figure 3. Sample printout of simulation control parameters

```
RESULTS FOR SCENARIO #  1 AFTER 25 ITERATION(S)

GENERATION  5
   BREEDING POPULATION of   50 trees:

           Effect           Mean    (Std. Dev.)   Variance  (Std. Dev.)
      Additive effects      46.57   (11.37    )    43.93    (22.90    )
      Dominance effects     17.80   (5.105    )    15.90    (8.317    )
      Epistasis effects     .0000   (.0000    )    .0000    (.0000    )
      Environmental effects 118.0   (1.212    )    30.48    (7.940    )
      Total phenotype       182.4   (11.42    )    56.45    (22.30    )
      Average Inbreeding (F) .1723  (.8713E-02)    .3323E-02(.4924E-03)
      Effective pop'n size  10.36   (.2379    )

   UNTESTED Seed Orchard Population (  10 best untested phenotypes):

           Effect           Mean    (Std. Dev.)   Variance  (Std. Dev.)
      Additive effects      50.16   (12.11    )    42.08    (25.25    )
      Dominance effects     19.52   (5.995    )    13.28    (9.318    )
      Epistasis effects     .0000   (.0000    )    .0000    (.0000    )
      Environmental effects 118.7   (1.643    )    30.29    (15.66    )
      Total phenotype       188.3   (12.82    )    40.69    (24.91    )
      Average Inbreeding (F) .1873  (.1683E-01)    .3382E-02(.1116E-02)
      Effective pop'n size  5.929   (.2240    )
      Cumulative gain %     50.16   (12.11    )

   POLYCROSS-TESTED Seed Orchard Population (  10 best polyX-tested trees):

           Effect           Mean    (Std. Dev.)   Variance  (Std. Dev.)
      Additive effects      51.12   (12.05    )    39.41    (24.86    )
      Dominance effects     17.99   (5.565    )    18.22    (20.71    )
      Epistasis effects     .0000   (.0000    )    .0000    (.0000    )
      Environmental effects 116.3   (2.158    )    30.92    (12.99    )
      Total phenotype       185.5   (12.09    )    55.11    (31.38    )
      Average Inbreeding (F) .1836  (.1701E-01)    .3311E-02(.1164E-02)
      Effective pop'n size  6.177   (.2319    )
      Cumulative gain %     51.12   (12.05    )

   Deployment of Full-sib FAMILY MIXTURE (     10 best families):

           Effect           Mean    (Std. Dev.)   Variance  (Std. Dev.)
      Additive effects      51.09   (12.18    )    41.53    (25.77    )
      Dominance effects     19.30   (5.654    )    14.30    (7.381    )
      Epistasis effects     .0000   (.0000    )    .0000    (.0000    )
      Environmental effects 100.2   (.2789    )    76.05    (3.084    )
      Total phenotype       170.6   (13.08    )    126.6    (26.89    )
      Average Inbreeding (F) .2164  (.2056E-01)    .3176E-02(.9048E-03)
      Cumulative gain %     70.39   (13.12    )

   Deployment of CLONAL MIXTURE (  10 best clones):

           Effect           Mean    (Std. Dev.)   Variance  (Std. Dev.)
      Additive effects      54.92   (13.97    )    39.57    (27.70    )
      Dominance effects     21.84   (6.949    )    14.92    (9.354    )
      Epistasis effects     .0000   (.0000    )    .0000    (.0000    )
      Environmental effects 122.0   (2.303    )    27.74    (14.47    )
      Total phenotype       198.7   (14.44    )    30.44    (21.37    )
      Average Inbreeding (F) .2101  (.1886E-01)    .3367E-02(.1574E-02)
      Cumulative gain %     76.76   (15.97    )
```

Figure 4. Sample printout of population structure for generation 5

POPSIM is, by definition, limited by current quantitative theory describing the recombination of genetic effects. This is particularly true for the simulation of nonadditive effects that must be considered when strategies incorporating deployment of full-sib families or clones are to be evaluated. Many users will only be interested in the simulation of additive effects in strategies that deploy material produced in seed orchards, and the algorithms presented here for inheritance of additive effects are considered somewhat more reliable than those for nonadditive effects. POPSIM calculates family variance components directly and performs an adjustment to within-family additive variance to account for the effects of inbreeding. Even so, the rapid erosion of additive variance observed over several generations under some selection schemes raises some doubt about the validity of current quantitative theory when extended over several breeding cycles, in the absence of mutation or other mechanisms that might generate "new" genetic variance. Users should interpret simulations of such long-term scenarios with caution.

The breeding population in the POPSIM model is closed and generations are discrete. Furthermore, all sublines are managed in an identical fashion. In a real-life program, sublines may be out-of-phase with each other and are often managed differently under a two-tiered "nucleus" breeding scheme (Cotterill *et al.* 1989) with some identified as "elite" breeding groups. Some genetic exchange may occur among sublines and new material may be introduced. Selections established in production orchards will usually be the known best parents and may come from different generations. While some of these features could be incorporated in a model, the simplifications used by POPSIM are thought to be reasonable and still permit the comparison of management strategies in a meaningful way. Similarly, POPSIM is limited to only a few of the many approaches that could be employed in selection and only one trait (or multiple-trait index) is considered. Alternative selection schemes could be added to the software, as required.

Particularly when compared to deterministic models, POPSIM is quite slow and generating reliable results requires several iterations. The program must process realistic quantities of simulated test data, performing all of the evaluation and selection calculations that would be carried out in a real-life breeding program. Furthermore, the simulation of stochastic processes requires the generation of random numbers and POPSIM spends most of its execution time performing this procedure. Random number generators vary tremendously in both speed and randomness, and we have opted for a very reliable, but rather slow algorithm.

A Final Note on Distribution and Support

POPSIM is protected under Canadian copyright legislation. Although a demonstration version with limited capacity is available from the Canadian Forest Service - Maritimes Region, and may be distributed freely, licenses for customized versions and technical support are available only from the developer, Genesis Forest Science Canada Incorporated. Interested parties should contact the senior author for details on software distribution, technical support, capabilities of the current version, and arrangements for modifications to meet specific needs:

Dr. Tim Mullin
Genesis Forest Science Canada Inc.
PO Box 1321
Truro, NS B2N 5N2
CANADA

Telephone/fax: 902-897-0339
E-mail: *genesis@fox.nstn.ca*

Acknowledgements

Literature Cited

Baker, R.J. 1986. Selection indices in plant breeding. CRC Press, Boca Raton, FL.

Borralho, N.M.G. 1994. Heterogeneous selfing rates and dominance effects in estimating heritabilities from open pollinated progeny. Can. J. For. Res. 24: 1079-1082.

Burrows, P.M. 1984. Inbreeding under selection from unrelated families. Biometrics 40: 357-366.

Caballero, A. 1994. Developments in the prediction of effective population size. Heredity 73: 657-679.

Chesser, R.K., Rhodes, O.E., Jr., Sugg, D.W., and Schnabel, A. 1993. Effective sizes for subdivided populations. Genetics 135: 1221-1232.

Cotterill, P.P. 1986. Genetic gains expected from alternative breeding strategies including simple low cost options. Silvae Genet. 35: 212-223.

Cotterill, P.P., and Jackson, N. 1989. Gains expected from clonal orchards under alternative breeding strategies. For. Sci. 35: 183-196.

Cotterill, P., Dean, C., Cameron, J., and Brindbergs, M. 1989. Nucleus breeding: a new strategy for rapid improvement under clonal forestry. *In* Breeding tropical trees: population structure and genetic improvement strategies in clonal and seedling forestry: Proceedings of a conference held in Pattaya, Thailand 28 November - 3 December 1988, by IUFRO Working Parties S2.02-08 Tropical species provenances and breeding, and S2.02-09 Eucalypt provenances and breeding. *Edited by* G.L. Gibson, A.R. Griffin, and A.C. Matheson. Oxford Forestry Institute, Oxford, UK and Winrock International, Arlington, VA. pp. 39-51.

de Boer, I.J.M., and van Arendonk, J.A.M. 1992. Prediction of additive and dominance effects in selected or unselected populations with inbreeding. Theor. Appl. Genet. 84: 451-459.

Dempfle, L. 1990. Problems in the use of the relationship matrix in animal breeding. *In* Advances in Statistical Methods for Genetic Improvement of Livestock. *Edited by* D. Gianola and K. Hammond. Springer Verlag, Berlin, Germany. pp. 454-473.

de Roo, G. 1987. A stochastic model to study breeding schemes in a small pig population. Agric. Syst. 25: 1-25.

de Vries, A.G., van der Steen, H.A.M., and de Roo, G. 1990. Multi-stage selection in a closed dam line of pigs. Livest. Prod. Sci. 24: 161-180.

Falconer, D.S. 1981. Introduction to quantitative genetics. 2nd ed. Longman, London and New York.

Fowler, D.P. 1986. Strategies for the genetic improvement of important tree species in the Maritimes. Inf. Rep. M-X-156, Can. For. Serv. - Maritimes, Fredericton, NB.

Fraser, A.S. 1957. Simulation of genetic systems by automatic digital computers. I. Introduction. Aust. J. Biol. Sci. 19: 484-491.

Griffin, A.R., and Cotterill, P.P. 1988. Genetic variation in growth of outcrossed, selfed and open-pollinated progenies of Eucalyptus regnans and some implications for breeding strategy. Silvae Genet. 37: 124-131.

Kempthorne, O. 1988. An overview of the field of quantitative genetics. *In* Proceedings of the Second International Conference on Quantitative Genetics, NC State University, Raleigh, NC, 1987. *Edited by* B.S. Weir, E.J. Eisen, M.M. Goodman, and G. Namkoong. Sinauer Assoc., Sunderland, MA. pp. 47-56.

King, J.N., and Johnson, G.R. 1991. Computer simulation of clonal selection gains. In Proceedings: FRI/NZFP Forests Ltd clonal forestry workshop, 1-2 May 1989, Rotorua, NZ. Edited by J.T. Miller. FRI Bulletin No. 160, Forest Research Institute and NZFP Forests Ltd., Rotorua, NZ. pp. 51-52.

King, J.N., and Johnson, G.R. 1993. Monte Carlo simulation models of breeding-population advancement. Silvae Genet. 42: 68-78.

Lacy, R.C. 1987. Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. Conserv. Biol. 1: 143-158.

Levin, B.R. 1969. Simulation of genetic systems. *In* Computer Applications in Genetics. *Edited by* N.E. Morton. Proceedings of the International Conference on Computer Applications in Genetics, Univ. Hawaii Press, Honolulu, HW. pp. 38-46.

Libby, W.J. 1964. Clonal selection, and an alternative seed orchard scheme. Silvae Genet. 13: 32-40.

Mahalovich, M.F. 1990. Modeling positive assortative mating and elite populations in recurrent selection programs for general combining ability. Ph.D. thesis, NC State Univ., 129 p.

Matheson, A.C., and Lindgren, D. 1985. Gains from the clonal and the clonal seed-orchard options compared for tree breeding programs. Theor. Appl. Genet. 71: 242-249.

Meuwissen, T.H.E., and Luo, Z. 1992. Computing inbreeding coefficients in large populations. Genet. Sel. Evol. 24: 305-313.

Meyer, K., and Smith, C. 1990. Comparison of theoretical and simulated equilibrium genetic response rates with progeny testing in dairy cattle. Anim. Prod. 50: 207-212.

Microsoft. 1993. Microsoft® FORTRAN PowerStation Professional Development System Version 1.0 for MS-DOS® and Windows™ Operating Systems. Microsoft Corporation, Redmond, WA.

Mullin, T.J. 1994. Evaluating the economics of alternative breeding and deployment strategies for northeastern conifers. In Proceedings of the 24th Meeting of the Canadian Tree Improvement Association, Part 2, 9-12 August 1993, Fredericton, NB. Edited by J. Lavereau. Natural Resources Canada, Fredericton, NB. pp. 82-105.

Mullin, T.J., and Park, Y.S. 1992. Estimating genetic gains from alternative breeding strategies for clonal forestry. Can. J. For. Res. 22: 14-23.

Mullin, T.J., and Park, Y.S. 1995. Stochastic simulation of population management strategies for tree breeding: a new decision-support tool for personal computers. Silvae Genet. [*in press*]

Namkoong, G., Snyder, E.B., and Stonecypher, R.W. 1966. Heritability and gain concepts for evaluating breeding systems such as seedling orchards. Silvae Genet. 15: 76-84.

Park, Y.S., and Fowler, D.P. 1984. Inbreeding in black spruce (Picea mariana (Mill.) B.S.P.): self-fertility, genetic load, and performance. Can. J. For. Res. 14: 17-21.

Park, Y.S., Simpson, J.D., Adams, G.W., Morgenstern, E.K., and Mullin, T.J. 1993. An updated breeding strategy for black spruce (*Picea mariana* (Mill.) B.S.P.) in New Brunswick. In Papers presented at the Workshop on Breeding Strategies of Important Tree Species in Canada, Fredericton, New Brunswick, August 18, 1993. Edited by Y.S. Park and G.W. Adams. Can. For. Serv. - Marit. Reg. Inform. Rep. M-X-186E, Fredericton, NB. pp. 41-54.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. 1992. Numerical recipes in FORTRAN: the art of scientific computing. (2nd ed.) Cambridge University Press, Cambridge, UK.

Ruane, J., and Thompson, R. 1991. Comparison of simulated and theoretical results in adult MOET nucleus schemes for dairy cattle. Livest. Prod. Sci. 28: 1-20.

Shaw, D.V., and Hood, J.V. 1985. Maximizing gain per effort by using clonal replicates in genetic tests. Theor. Appl. Genet. 71: 392-399.

29

Shelbourne, C.J.A. 1969. Tree breeding methods - A review of tree breeding strategies in relation to classical plant breeding methods with quantitative genetic expectations of gain, and predictions using data from Pinus radiata. Technical Report No. 55, New Zealand Forest Service, Wellington.

Shelbourne, C.J.A. 1992. Genetic gains from different kinds of breeding population and seed or plant production population. Paper presented at the IUFRO symposium "Intensive forestry: the role of eucalypts", held in Durban, South Africa, in September 1991. South African Forestry Journal 160: 49-65.

Tier, B. 1990. Computing inbreeding coefficients quickly. Genet. Sel. Evol. 22: 419-430.

van Buijtenen, J.P., and Lowe, W.J. 1979. The use of breeding groups in advanced generation breeding. In Proceedings of the 15th Southern Forest Tree Improvement Conference. pp. 56-65.

Weir, R.J., and Todd, D. 1994. Third-cycle breeding strategy: a description and economic appraisal for the North Carolina University - Industry Cooperative Tree Improvement Program. In Proceedings of the 24th Canadian Tree Improvement Association, Part 2, 9-12 Aug. 1993, Fredericton, NB. Edited by J. Lavereau. Natural Resources Canada, Fredericton, NB. pp. 41-51.

APPENDIX "A"
Definitions of fields in tab-delimited series data file root.TXT

Column	Variable name	Definition
A	ROOT	Root name for data file
B	KSCEN	Scenario number
C	IGEN	Generation number (0, ... NGEN)
D	NGEN	Total number of generations simulated
E	NITER	Number of iterations
F	NT	Census size of breeding population
G	NT/NS	Number of trees in subline
H	NS	Number of sublines in breeding population
I	ISIZE	Effective size of breeding population expressed: 0 = Ne, 1 = Ns
J	TRAITBAR	Mean value of trait in base population
K	BID	Standardized rate of inbreeding depression
L	AVAR	Additive genetic variance of trait in base population
M	DVAR	Dominance genetic variance of trait in base population
N	IVAR	Epistasic genetic variance of trait in base population
O	EVAR	Environmental variance in selection environments
P	NSH2	Narrow-sense heritability in base population
Q	BSH2	Broad-sense heritability in base population
R	NP	Number of males represented in polycross pollen mix
S	NX	Number of progeny tested per polycross (if required)
T	ISORT	Positive assortative mating: 0 = none, 1 = random, 2 = uniform
U	NXS	Number of crosses made per subline for generation advancement
V	NXTOTAL	Total number of crosses made for generation advancement
W	NC	Number of genotypes tested per cross = family size
X	NR	Number of ramets tested per genotype
Y	IBPSEL	Selection method used to advance breeding population:
Z	NBPFS	(Maximum) number per full-sib cross to enter next breeding population
AA	NBPHS	Maximum number of half-sib relatives to enter next breeding population
AB	BPHSRLX	NBPHS restriction relaxed ? (0=No, 1=Yes)
AC	NBPPC	Maximum number of progeny per parent to enter next breeding population
AD	BPPCRLX	NBPPC restriction relaxed ? (0=No, 1=Yes)
AE	IUNTST	Select untested orchard ? (0=No, 1=Yes)
AF	IPXTST	Select tested orchard based on polycross assessment ? (0=No, 1=Yes)
AG	ISOSEL	Selection method used for untested orchard:
AH	NSSO	Number of selections required for seed orchard
AI	NSOFS	Maximum number per full-sib cross permitted to enter seed orchard
AJ	NSOHS	Maximum number of half-sib relatives permitted to enter seed orchard
AK	NSOPC	Maximum number of progeny per parent permitted to enter seed orchard
AL	SORELAXED	NSOHS or NSOPC restrictions relaxed for seed orchards ? (0=No, 1=Yes)
AM	IFAMSEL	Select full-sib family mixture ? (0=No, 1=Yes)
AN	NFAM	Number of families represented in full-sib family mixture
AO	NFAMHS	Maximum number of half-sib relatives permitted in full-sib family mixture
AP	NFAMPC	Maximum number crosses per parent permitted in full-sib family mixture
AQ	FARELAX	NFAMHS or NFAMPC restrictions relaxed for family mixture ? (0=No, 1=Yes)
AR	ICLNSEL	Select tested clonal mixture ? (0=No, 1=Yes)
AS	NCLN	Number of genotypes represented in clonal mixture

Column	Variable name	Definition
AT	NCLNFS	Maximum number of full-sibs permitted to enter clonal mixture
AU	NCLNHS	Maximum number of half-sib relatives permitted in clonal mixture
AV	NCLNPC	Maximum number progeny per parent permitted in clonal mixture
AW	CLRELAX	NCLNHS or NCLNPC restrictions relaxed for clonal mixture ? (0=No, 1=Yes)
AX	BPADDE	Mean additive genetic effect in breeding population
AY	BPADDESD	Standard deviation of mean additive genetic effect in breeding population
AZ	BPADDV	Additive genetic variance in breeding population
BA	BPADDVSD	Standard deviation of additive genetic variance in breeding population
BB	BPDOME	Mean dominance genetic effect in breeding population
BC	BPDOMESD	Standard deviation of mean dominance genetic effect in breeding population
BD	BPDOMV	Dominance genetic variance in breeding population
BE	BPDOMVSD	Standard deviation of dominance genetic variance in breeding population
BF	BPEPIE	Mean epistasis genetic effect in breeding population
BG	BPEPIE SD	Standard deviation of mean epistasis genetic effect in breeding population
BH	BPEPIV	Epistasis genetic variance in breeding population
BI	BPEPIVSD	Standard deviation of epistasis genetic variance in breeding population
BJ	BPENVE	Mean environmental effect in breeding population
BK	BPENVESD	Standard deviation of mean environmental effect in breeding population
BL	BPENVV	Environmental variance in breeding population
BM	BPENVVSD	Standard deviation of environmental variance in breeding population
BN	BPPHEE	Mean phenotypic effect in breeding population
BO	BPPHEESD	Standard deviation of mean phenotypic effect in breeding population
BP	BPPHEV	Phenotypic variance in breeding population
BQ	BPPHEVSD	Standard deviation of phenotypic variance in breeding population
BR	BPINB	Mean inbreeding coefficient in breeding population
BS	BPINBSD	Standard deviation of mean inbreeding coefficient in breeding population
BT	BPINBV	Variance of inbreeding coefficient in breeding population
BU	BPINBVSD	Standard deviation of inbreeding coefficient variance in breeding population
BV	BPEPS	Mean effective population size of breeding population
BW	BPEPSSD	Standard deviation of mean effective population size of breeding population
BX	USADDE	Mean additive genetic effect in untested orchard
BY	USADDESD	Standard deviation of mean additive genetic effect in untested orchard
BZ	USADDV	Additive genetic variance in untested orchard
CA	USADDVSD	Standard deviation of additive genetic variance in untested orchard
CB	USDOME	Mean dominance genetic effect in untested orchard
CC	USDOMESD	Standard deviation of mean dominance genetic effect in untested orchard
CD	USDOMV	Dominance genetic variance in untested orchard
CE	USDOMVSD	Standard deviation of dominance genetic variance in untested orchard
CF	USEPIE	Mean epistasis genetic effect in untested orchard
CG	USEPIESD	Standard deviation of mean epistasis genetic effect in untested orchard
CH	USEPIV	Epistasis genetic variance in untested orchard
CI	USEPIVSD	Standard deviation of epistasis genetic variance in untested orchard
CJ	USENVE	Mean environmental effect in untested orchard
CK	USENVESD	Standard deviation of mean environmental effect in untested orchard
CL	USENVV	Environmental variance in untested orchard
CM	USENVVSD	Standard deviation of environmental variance in untested orchard
CN	USPHEE	Mean phenotypic effect in untested orchard
CO	USPHEESD	Standard deviation of mean phenotypic effect in untested orchard
CP	USPHEV	Phenotypic variance in untested orchard
CQ	USPHEVSD	Standard deviation of phenotypic variance in untested orchard

Column	Variable name	Definition
CR	USINBE	Mean inbreeding coefficient in untested orchard
CS	USINBESD	Standard deviation of mean inbreeding coefficient in untested orchard
CT	USINBV	Variance of inbreeding coefficient in untested orchard
CU	USINBVSD	Standard deviation of inbreeding coefficient variance in untested orchard
CV	USEPS	Mean effective population size of untested orchard
CW	USEPSSD	Standard deviation of mean effective population size of untested orchard
CX	USCGN	Mean cumulative genetic gain from untested orchard
CY	USCGNSD	Standard deviation of mean cumulative genetic gain untested orchard
CZ	TSADDE	Mean additive genetic effect in PX-tested orchard
DA	TSADDESD	Standard deviation of mean additive genetic effect in PX-tested orchard
DB	TSADDV	Additive genetic variance in PX-tested orchard
DC	TSADDVSD	Standard deviation of additive genetic variance in PX-tested orchard
DD	TSDOME	Mean dominance genetic effect in PX-tested orchard
DE	TSDOMESD	Standard deviation of mean dominance genetic effect in PX-tested orchard
DF	TSDOMV	Dominance genetic variance in PX-tested orchard
DG	TSDOMVSD	Standard deviation of dominance genetic variance in PX-tested orchard
DH	TSEPIE	Mean epistasis genetic effect in PX-tested orchard
DI	TSEPIESD	Standard deviation of mean epistasis genetic effect in PX-tested orchard
DJ	TSEPIV	Epistasis genetic variance in PX-tested orchard
DK	TSEPIVSD	Standard deviation of epistasis genetic variance in PX-tested orchard
DL	TSENVE	Mean environmental effect in PX-tested orchard
DM	TSENVESD	Standard deviation of mean environmental effect in PX-tested orchard
DN	TSENVV	Environmental variance in PX-tested orchard
DO	TSENVVSD	Standard deviation of environmental variance in PX-tested orchard
DP	TSPHEE	Mean phenotypic effect in PX-tested orchard
DQ	TSPHEESD	Standard deviation of mean phenotypic effect in PX-tested orchard
DR	TSPHEV	Phenotypic variance in PX-tested orchard
DS	TSPHEVSD	Standard deviation of phenotypic variance in PX-tested orchard
DT	TSINBE	Mean inbreeding coefficient in PX-tested orchard
DU	TSINBESD	Standard deviation of mean inbreeding coefficient in PX-tested orchard
DV	TSINBV	Variance of inbreeding coefficient in PX-tested orchard
DW	TSINBVSD	Standard deviation of inbreeding coefficient variance in PX-tested orchard
DX	TSEPS	Mean effective population size of PX-tested orchard
DY	TSEPSSD	Standard deviation of mean effective population size of PX-tested orchard
DZ	TSCGN	Mean cumulative genetic gain from PX-tested orchard
EA	TSCGNSD	Standard deviation of mean cumulative genetic gain PX-tested orchard
EB	FMADDE	Mean additive genetic effect in FS-family mixture
EC	FMADDESD	Standard deviation of mean additive genetic effect in FS-family mixture
ED	FMADDV	Additive genetic variance in FS-family mixture
EE	FMADDVSD	Standard deviation of additive genetic variance in FS-family mixture
EF	FMDOME	Mean dominance genetic effect in FS-family mixture
EG	FMDOMESD	Standard deviation of mean dominance genetic effect in FS-family mixture
EH	FMDOMV	Dominance genetic variance in FS-family mixture
EI	FMDOMVSD	Standard deviation of dominance genetic variance in FS-family mixture
EJ	FMEPIE	Mean epistasis genetic effect in FS-family mixture
EK	FMEPIESD	Standard deviation of mean epistasis genetic effect in FS-family mixture
EL	FMEPIV	Epistasis genetic variance in FS-family mixture
EM	FMEPIVSD	Standard deviation of epistasis genetic variance in FS-family mixture
EN	FMENVE	Mean environmental effect in FS-family mixture
EO	FMENVESD	Standard deviation of mean environmental effect in FS-family mixture

Column	Variable name	Definition
EP	FMENVV	Environmental variance in FS-family mixture
EQ	FMENVVSD	Standard deviation of environmental variance in FS-family mixture
ER	FMPHEE	Mean phenotypic effect in FS-family mixture
ES	FMPHEESD	Standard deviation of mean phenotypic effect in FS-family mixture
ET	FMPHEV	Phenotypic variance in FS-family mixture
EU	FMPHEVSD	Standard deviation of phenotypic variance in FS-family mixture
EV	FMINBE	Mean inbreeding coefficient in FS-family mixture
EW	FMINBESD	Standard deviation of mean inbreeding coefficient in FS-family mixture
EX	FMINBV	Variance of inbreeding coefficient in FS-family mixture
EY	FMINBVSD	Standard deviation of inbreeding coefficient variance in FS-family mixture
EZ	FMCGN	Mean cumulative genetic gain from FS-family mixture
FA	FMCGNSD	Standard deviation of mean cumulative genetic gain FS-family mixture
FB	CMADDE	Mean additive genetic effect in clonal mixture
FC	CMADDESD	Standard deviation of mean additive genetic effect in clonal mixture
FD	CMADDV	Additive genetic variance in clonal mixture
FE	CMADDVSD	Standard deviation of additive genetic variance in clonal mixture
FF	CMDOME	Mean dominance genetic effect in clonal mixture
FG	CMDOMESD	Standard deviation of mean dominance genetic effect in clonal mixture
FH	CMDOMV	Dominance genetic variance in clonal mixture
FI	CMDOMVSD	Standard deviation of dominance genetic variance in clonal mixture
FJ	CMEPIE	Mean epistasis genetic effect in clonal mixture
FK	CMEPIESD	Standard deviation of mean epistasis genetic effect in clonal mixture
FL	CMEPIV	Epistasis genetic variance in clonal mixture
FM	CMEPIVSD	Standard deviation of epistasis genetic variance in clonal mixture
FN	CMENVE	Mean environmental effect in clonal mixture
FO	CMENVESD	Standard deviation of mean environmental effect in clonal mixture
FP	CMENVV	Environmental variance in clonal mixture
FQ	CMENVVSD	Standard deviation of environmental variance in clonal mixture
FR	CMPHEE	Mean phenotypic effect in clonal mixture
FS	CMPHEESD	Standard deviation of mean phenotypic effect in clonal mixture
FT	CMPHEV	Phenotypic variance in clonal mixture
FU	CMPHEVSD	Standard deviation of phenotypic variance in clonal mixture
FV	CMINBE	Mean inbreeding coefficient in clonal mixture
FW	CMINBESD	Standard deviation of mean inbreeding coefficient in clonal mixture
FX	CMINBV	Variance of inbreeding coefficient in clonal mixture
FY	CMINBVSD	Standard deviation of inbreeding coefficient variance in clonal mixture
FZ	CMCGN	Mean cumulative genetic gain from clonal mixture
GA	CMCGNSD	Standard deviation of mean cumulative genetic gain clonal mixture