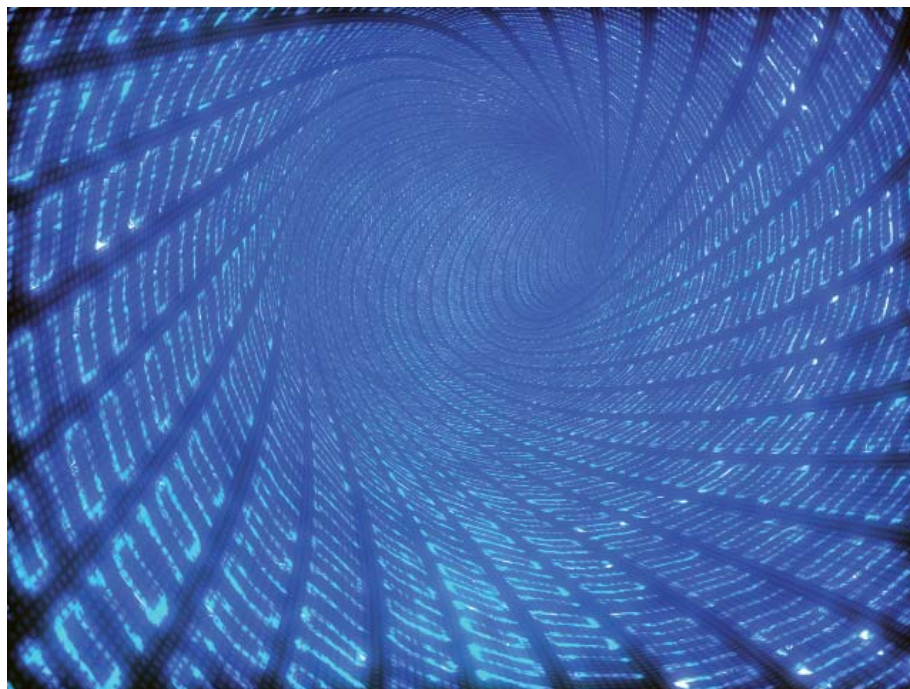


ARBETSRAPPORT

FRÅN SKOGFORSK NR 711 2010



Using simulation to optimise tree breeding programmes in Europe: An introduction to POPSIM™

T. J. Mullin, J. Hallander & O. Rosvall & B. Andersson

This research has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under the grant agreement no 211868 (Project Noveltree)

Keywords:

Förädlingsstrategi, Genetisk variation, Genetisk vinst, Populationsstorlek, Selektion, Simulering, Släktskap.

Breeding strategy, Genetic variation, Genetic gain, Population size, Selection, Simulation, Relatedness.

SKOGFORSK**– The Forestry Research Institute of Sweden**

The mission of Skogforsk is to pursue research and development for a sustainable, profitable and ecologically sound forestry. The Institute is supported by the entire forestry sector in Sweden through member fees and a charge levied on each cubic meter of wood harvested. Financial support is also given by the government according to a special agreement and by scientific funds for certain projects.

RESEARCH

- Forest production
- Wood supply

CONSULTANCY WORK

The Institute undertakes extensive consultancy work in its fields of special competence on behalf of forest enterprises, machine manufacturers and local and central government authorities. Such projects may involve specially commissioned research or the tailoring of new methods and routines to local conditions.

DISSEMINATION OF INFORMATION The Institute uses a variety of media and channels for disseminating the findings of its research. These include personal contact, Internet, courses and seminars, the trade press, films and its own series of publications.

Contents

| | |
|---|----|
| Sammanfattning | 2 |
| Summary | 3 |
| Introduction..... | 4 |
| Understanding parametric simulation models | 5 |
| Understanding the optimisation process | 7 |
| Hardware/software requirements for POPSIM™ | 8 |
| Becoming a registered user of POPSIM™ | 8 |
| A note on copyright, licensing and distribution | 8 |
| Installing POPSIM™ | 8 |
| Setting up scenarios – the input interface..... | 9 |
| Scenario identification..... | 10 |
| Simulation dimensions | 10 |
| Genetic structure of the managed population | 10 |
| Progeny testing specifications..... | 11 |
| Specifications for generating recruitment populations | 13 |
| Selection of the breeding population..... | 14 |
| Selection of seed orchard production population | 16 |
| Selection of a full-sib family production population..... | 17 |
| Selection of a mixture of selected clones | 18 |
| Local file settings | 19 |
| ASReml file specifications | 19 |
| Running a series of scenarios..... | 19 |
| Retrieving and viewing output..... | 20 |
| Saving scenario parameters and results | 20 |
| Analysing simulation results..... | 20 |
| When things go wrong!..... | 20 |
| Help to improve POPSIM™ | 21 |
| References..... | 21 |
| Appendix 1 How POPSIM™ calculates “combined index” | 25 |
| Appendix 2 How POPSIM™ calculates “group merit” | 27 |

Sammanfattning

Utvärdering och optimering av förädlingsstrategier för skogsträd kan med fördel hanteras med stokastisk simulering. Rekombination av alleler vid reproduktion (korsning) är en stokastisk effekt, och således även uttrycken av genetiska effekter i en population. Denna process kan efterliknas i en simulering, när en sluten population påverkas enligt en förutbestämd förädlingsstrategi. När en simulering upprepas många gånger kan slutsatser dras om sannolikhetsfördelningen av olika parametrar ifrån ett scenario, som hjälper till att bilda en uppfattning om ett ”typiskt” utfall. Få antaganden behövs för att prediktera resultatet av olika förädlingsmetoder med komplicerade släktskap över multipla generationer av förädling.

POPSIM™ är ett exempel på en parametrisk simulering. Med detta program kan första generationens plusträd (founders) förädlas enligt en given strategi. Programmet kan generera avkommor både för avkommeprövning och för den kommande förädlingspopulationen, utvärdera data ifrån fältförsök (BLUP avelsvärden), selektera fördelaktiga träd med specificerad selektionsmetod, generera produktionspopulationer (tex fröplantager, klonblandningar) och beräkna olika populationsegenskaper som exempelvis den genetiska vinsten, effektiv populationsstorlek och diversitetsförlust. Det finns många möjligheter och alternativ ifrån operativa förädlingsprogram inbyggda i POPSIM, som exempelvis avkommeprövning, flexibla korsningsdesigner som t.e.x. rangordnade och obalanserade, klonupprepning i fältförsök, olika selektionsmetoder inkluderat restriktion på släktskapet hos de utvalda träden.

POPSIM™ är ett projekt som pågått under 15 år och är under ständig utveckling. Nu har nya, viktiga verktyg lagts till för att anpassa programmet till Europeiska behov. Dessa verktyg inkluderar ett oberoende steg av avkommeprövning av plusträd, med fri avblomning eller kontrollerad pollenblandning, fullständig släktskapsinformation för BLUP-beräkningar av avelsvärden över samtliga generationer, utökade populationsstrukturer som exempelvis förselektion och avkommeprövning följt av en utökning av förädlingspopulationen med besläktade föräldrar och funktion för undvikande av att närbesläktade träd korsas även med rangordnade korsningar. Projektet har fått finansiering från EU:s sjunde ramprogram (FP7/2007-2013) under avtal nr 211868 (Projekt NovelTree).

Ett användarvänligt gränssnitt gör det enkelt att ställa upp jämförande scenarion, med hjälp av välkända funktioner i ett Excelark. Efter att simuleringen är genomförd kan användaren ladda in resultaten i Excelarket för vidare behandling av exempelvis medelvärden och standardavvikelser av alla parametrar för varje generation av förädling. Exempelvis kan förändringar i genetiska effekter och varianser, genetiska vinster i produktionspopulationerna samt inavelsökning och effektiva populationsstorlek följas. Användaren kan efteråt behandla resultaten och göra informativa och lättöverskådliga grafer. POPSIM™ kräver Excel 2007 (eller en senare version) men kan köras på både 32- eller 64-bits dator med Windows. BLUP beräkningarna kräver att mjukvaruprogrammet ASReml finns tillgängligt (finns hos VSN International). Dock är det inte nödvändigt att använda BLUP som metodik för avelsutvärdering.

POPSIM™ blev introducerat för Europeiska trädförädlare vid den gemensamma NovelTree & Treebreedex seminariet i Orleans (12–14 oktober, 2009). Skogforsk för register över användare i NovelTree som tillhandahåller uppdateringar och support.

Summary

Stochastic simulation is particularly well suited to the evaluation and optimisation of tree breeding strategies. The recombination of alleles during reproduction, and thus sampling of genetic effects, is a “stochastic” process. The simulation mimics this sampling when a closed population is managed in a specified way. When the simulation is repeated, the distribution of “actual” outcomes reflects the effect of random sampling on the likelihood of achieving the “average” outcome. Few assumptions are necessary, making it possible to predict outcomes for complicated pedigrees and combinations of breeding methods over several cycles of breeding.

POPSIM™ is an example of a parametric simulation. With it, we can breed founders according to a plan, generate offspring for progeny testing and for recruitment populations, evaluate field test data (BLUP breeding values), make selections according to specified criteria, generate deployment populations, and evaluate their genetic gain, effective size and diversity. Included in POPSIM™ are many features of “real” breeding programmes, including progeny testing, flexible mating designs, assortative and unbalanced mating, clonal replication of field tests, various approaches to selection, including restriction on relatedness of selected trees.

While POPSIM™ has its roots in 15 years of collaborative development, important new features have been incorporated to adapt the tool to European requirements. These include an independent progeny-testing cycle of plus-tree founders, progeny testing by open-pollination or polycross, maintenance of a complete pedigreed database for BLUP calculation of BVs across all cycles, additional population structures, including pre-selection + progeny testing and “expansion” of breeding populations with related parents, and avoidance of closely related matings, even under assortative mating. This research and development has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement n° 211868 (Project NovelTree).

A user-friendly interface makes it simple to set up scenarios for comparison, using familiar tools in an Excel spreadsheet environment. After the simulations are complete, the user can migrate the results into the Excel workbook and automatically generate a report of means and standard deviations for all parameters over all iterations, and for each cycle of breeding, changes in genetic effects and variances, genetic gain from various production and deployment scenarios, as well as inbreeding and effective population size. The user can then save all settings and tabulated results, and manipulate the data as required with Excel analysis and graphing functions. The POPSIM™ package requires Excel 2007 (or later), but will run under any 32- or 64-bit Windows OS environment. BLUP calculations are available, provided the user has access to a licensed version of ASReml (available from VSN International).

The POPSIM™ tool was introduced to European tree breeders at the joint NovelTree & Treebreedex workshop held in Orléans (12–14 October, 2009). A register of NovelTree users receiving updates and support is maintained by Skogforsk.

Introduction

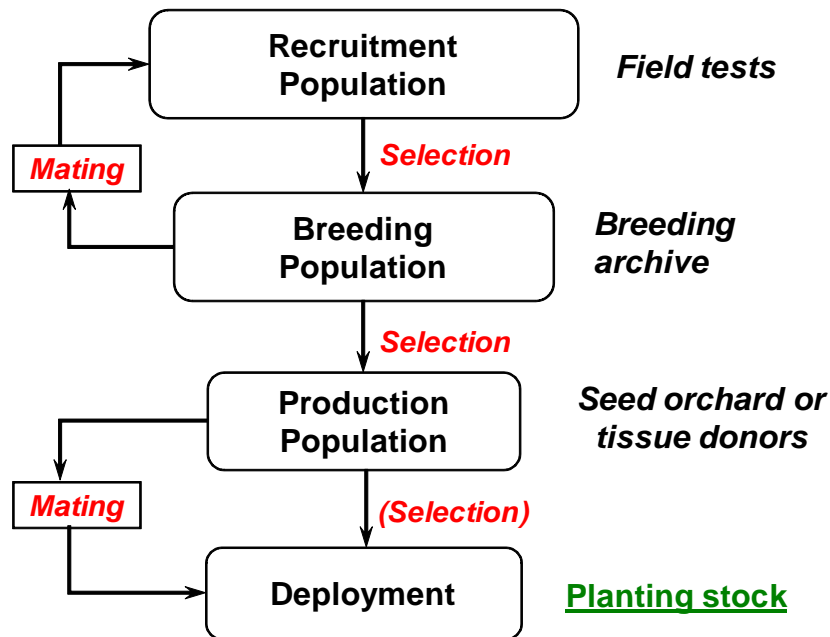


Figure 1.
Schematic representation of a typical tree breeding programme.

The breeding of forest trees, like most other higher crop plants, is generally carried out as repeated cycles of selection, mating, and testing, within managed populations. Figure 1 gives a schematic representation of such a programme. The process begins with a “recruitment population”, which in the first cycle is usually a wild forest of undomesticated trees. Individuals are selected in the recruitment population to form a “breeding population”, which is often archived by grafting to facilitate management. Individuals in the breeding population are mated and their progeny established in field tests, which become the recruitment population for the next cycle. Meanwhile selected individuals are identified to form “production populations”, such as seed orchards or tissue archives that will produce stock for forest planting. While many changes can occur over time in the genetic structure of the recruitment, breeding and production populations, the impact of a programme must ultimately be measured by the genetic quality and diversity of deployed planting stock.

The prediction of genetic gain from one round of selection in a tree improvement programme is a relatively straightforward calculation, requiring only appropriate estimates of heritability, selection intensity and the phenotypic variance of the selection units. There are many papers in the literature demonstrating how this basic calculation can be applied to a wide variety of situations

found in applied tree breeding programmes (e.g., Lindgren et al., 2008; Shelbourne 1969, 1992; Verry et al., 2000).

Although such gain calculations are useful, comparisons of breeding strategies based only on gain estimates from one round of selection are often inadequate. More often the objective of an improvement programme will be to maximise 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).

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). 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.

This paper describes the application of 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 was first documented by Mullin & Park (1995), and has been expanded periodically in other published work where the tool was applied (Andersson et al., 1998; Lindgren & Mullin 1997; 2006; Lstibůrek et al. 2004a, b; 2005; Rosvall et al. 1998; Rosvall & Mullin 2003; Rosvall et al., 2003; Weng et al., 2009).

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 (e.g., White et al., 2007). Similarly, the guide assumes a general familiarity with basic procedures and configuration of PC-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.

Understanding parametric simulation models

Deterministic models predict the outcome of a process through algebraic descriptions of the process and/or regression equations developed from observed sample data. Stochastic simulations attempt to “mimic” critical components of the process that are subject to random variation. The result of every simulation is different, as it is driven by stochastic sampling. To generalise about the outcome of a process, it is necessary to repeat the simulation many times, and observe the mean outcome and the variance around it.

Stochastic simulation models of genetic systems fall into two general categories: (1) allelic models, and (2) parametric models. Allelic models mimic the distribution and sampling of alleles and individual loci affecting a trait, while parametric models describe distribution of genetic effects, based on the

infinitesimal quantitative genetic model (Fisher, 1918), where variation in a trait (or index) is assumed to be “controlled by genes at an infinite number of loci, each with an infinitely small effect, in linkage equilibrium”.

POPSIM™ generally uses a parametric approach. In the model, each genotype is considered to be the sum of independent genetic and environmental effects, and the total phenotypic variation in a population is described as the sum of independent variances for each of these effects. The additive effect for an individual is generated from the mid-parent family mean and a within-family Mendelian sampling term:

$$A'_i = \frac{A_f + A_m}{2} + r \sqrt{\frac{1 - 0.5(F_f + F_m)}{2}} \sigma_A^2 \quad [1]$$

where A'_i is the additive effect for progeny i ; A_f and A_m are the additive effects of the female and male parents, respectively; F_f and F_m are the corresponding coefficients of inbreeding; σ_A^2 is the additive variance; and r is a normal deviate.

The dominance effect for an individual is generated as:

$$D'_i = r_{jm} \sqrt{\frac{1}{4} \sigma_D^2} + r_i \sqrt{\frac{3(1 - 0.5(F_f + F_m))}{4}} \sigma_D^2 \quad [2]$$

where σ_D^2 is the dominance variance; r_{jm} and r_i are random normal deviates for each family (jm) and individual (i), respectively.

Inbreeding depression is simulated by reducing the dominance effect by $bF_{jm}\sigma_p$, which describes the regression of inbreeding depression on the inbreeding coefficient of the family members F_{jm} , where the regression coefficient b expresses the reduction in phenotypic value in units of phenotypic standard deviation, σ_p , for the trait in the unselected base population.

Epistatic variance σ_I^2 and effects I'_i are considered to be within-family, so a simple sampling is made from a normal distribution of epistatic effects:

$$I'_i = r \sqrt{\sigma_I^2} \quad [3]$$

The environmental effect E'_{ik} for the k th ramet of the i th genotype assumes that the environmental variance σ_E^2 is normally distributed and equal in all generations:

$$E'_{ik} = r \sqrt{\sigma_E^2} \quad [4]$$

For each of the genetic and environmental effects, estimates of mean and variance are calculated from the generated data for the N trees in the breeding population. Group coancestry (Cockerham, 1967) and inbreeding are calculated from the pedigree. Status number N_s is an effective number calculated for each generation t from group coancestry Θ (Lindgren et al., 1996):

$$N_{s_t} = \frac{1}{2\Theta_t} \quad [5]$$

By repeating a simulation of a given scenario many times, both the stochastic variation in expected genetic effects and variances of a single scenario can be observed and described by simple statistics.

Understanding the optimisation process

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.

Of course, a real breeding programme is carried out only once, and while the expected outcome is the central tendency observed from many iterations, the actual outcome of any given iteration can differ. In addition to the mean values for each population parameter, POPSIM™ also reports the standard deviation, which describes the distribution of single iterations around the mean.

To carry out an optimisation, the user provides parameters that describe the type of population management and the dimensions applicable to the species and programme of interest. Experience or historical data are used to establish realistic genetic parameters and the expected timeline and costs are tabulated. Additional scenarios are described, where parameters are changed systematically, while keeping the overall investment per unit time constant. The results will give insight on the sensitivity of a particular breeding plan to actual genetic and management parameters. By rescaling the scenarios according to different budgets, the simulations can also investigate the sensitivity of breeding plans to available funding.

Hardware/software requirements for POPSIM™

POPSIM™ is written as a 32-bit program for PC-compatible computers, but will run on any computer under the Microsoft Windows Operating System from version “XP” to present, including 64-bit versions. While multitasking is possible when running POPSIM™, the processor requirements of simulation result in poor usability, so it is normally recommended to run the program with all other applications closed.

POPSIM™ requires that the user have a version of Microsoft Excel, no older than version 2007.

If the user wishes to simulate BLUP estimation of breeding values (other options are available), they will need run-time access to a licensed version of ASReml (Gilmour et al., 2009). POPSIM™ has been tested with all versions of ASReml up to version 3.

Becoming a registered user of POPSIM™

While the distribution of POPSIM™ is restricted, any person affiliated with a NovelTree beneficiary is eligible to use the package. Registered users will receive the package components in a zip file emailed directly to them. Periodic updates, bug fixes and lists of known issues will be circulated to all registered users as they are released. Support is available only to registered users, by contacting the same address.

A note on copyright, licensing and distribution

POPSIM™ is copyrighted material and its use is restricted. It has been made available to NovelTree beneficiaries by a special agreement with the copyright owner, BioSylve Forest Science NZ Limited. In accepting the notices on the program splash screens, the user agrees to conditions of use – please read these! Note also that the version of POPSIM™ distributed to NovelTree beneficiaries is NOT to be distributed to any other party without written consent from the copyright owner.

Installing POPSIM™

POPSIM™ is distributed as a “zip” file. The following procedure and folder locations are recommended for installing the program, although file locations could be changed, once successful operation of the program has been verified.

1. Extract the zip file to a convenient folder on your machine. The “POPSIM.exe” file will likely have had the “.exe” extension stripped, to enable the sending of the zip file through institution firewalls – if so, you will need to restore the name by adding “.exe” to the name. The “POPSIM.exe”, “help401.htm” and “help401_files” folder should be put together in a folder, such as C:\POPSIM\. The Excel workbook file (“POPSIM.xlsm”) should be put in a separate folder, such as C:\POPSIMFILES.

2. Check Excel settings. Click on the “Office” button (top left corner) and select “Excel options”. For the frontend to operate correctly, please check the following settings and adjust if required:
 - Excel Options → Trust Center → Trust Center Settings ... → Macro Settings "Enable all macros"
 - Excel Options → Save → delete any default file location, if one is specified.
3. Run the example scenarios supplied with POPSIM™, to verify installation. The “POPSIM.xlsm” file supplied with POPSIM™ is pre-loaded with examples that should run quickly on any machine, and do not require the use of ASRepl. Open the file by double clicking on the filename – DO NOT open the file by using the File → Open menu in Excel! Accept the information on the splash screen, then read the information on the “Instructions” sheet, where any changes to this user guide will be noted. Then select the “MainPage” sheet. If you installed POPSIM™ in folders other than those suggested above, you will need to modify the folder locations for input/output and executable files. Click on “Run simulations”, then “OK”. The program will check the scenario data and call the main POPSIM™ program. Accept the information on the POPSIM™ splash screen. You should now see simulation progress on the simulation control page (updating iterations and cycles). If you get into trouble, terminate the program by clicking “Ctrl + Alt + Del”, mark Excel in the task list and end task.

Setting up scenarios – the input interface

POPSIM™ uses Excel as a user interface to help the user prepare control data input, and manage the results output. The user will find it possible to use familiar “cut-and-paste”, formatting and file save features, as for any other Excel file. The “.xlsm” file supplied with POPSIM™ is preloaded with example data, but this workbook can be modified by the user and saved under a new name (but always as a “.xlsm” file type). Be aware that it is possible to “upset” the automated procedures built into the worksheet by adding rows or changing the macro programming. The user can always return to the example “.xlsm” workbook to restore functionality that may be lost by inadvertent changes to the workbook.

The process of specifying a set of simulation scenarios begins by clicking on the “Set up scenarios” button. You specify the number of scenarios to be run, and whether you want to delete all data in the sheet beyond this (this area of the sheet is often useful for making run notes, as any columns after the final scenario will be ignored by POPSIM™).

Column A on the MainPage gives an abbreviated description of each control parameter. Column B gives the short name for each parameter and indicates if the parameter is specified as an Integer (green background) or Real numeric input (orange background). POPSIM™ will ignore any parameter which is not required; for example, the value of RgOP is ignored if progeny testing uses

polycross families or if no progeny testing is performed. The control data are discussed here, line-by-line. Reminder notes have also been inserted as “comments” in the MainPage, which helps to select the correct options.

SCENARIO IDENTIFICATION

KSCEN Scenario number.

Each column is identified by a scenario sequence number, beginning at 1 and continuing to the number of scenarios specified by “Set up scenarios”.

SIMULATION DIMENSIONS

NGEN No. of generations.

Specify the number of breeding cycles to be completed. The simulation will report results for the generate the $F_0, F_1, \dots F_{\text{NGEN}}$ generations.

NITER No. of iterations.

Specify the number of iterations of the simulation. The simulation will be repeated **NITER** times, and the results reported as the mean outcomes and standard deviations for these iterations. Larger numbers of iterations will better estimate the expectation as the central tendency of the reported outcomes, but will of course add to the execution time. **NITER** should be set to at least 100 for reliable estimates of outcome expectations, while smaller numbers can be used for approximate results or preliminary experiments.

GENETIC STRUCTURE OF THE MANAGED POPULATION

NFndrs Number of founders.

This is the size of the original population of founders sampled from an “infinite” base or “wild” population. For tree breeding applications, it is usually convenient to consider this to be the original pool of “plus-trees” sampled from the natural forest. It is very common that this pool will be supply a smaller number of breeding parents, selected on the basis of OP or polycross progeny testing (see below).

NT Number of parents in the breeding population.

Specify the size of the population of parents bred to for a recruitment population. If $\text{NT} < \text{NFndrs}$, the original sample of founders will first be progeny tested to identify the best NT parents.

EFFBAR(4) Mean of trait in base population.

Specify the mean of the trait (or index of traits) in the original base population of **NFndrs**.

BID Rate of inbreeding depression.

Inbreeding depression is simulated by adjusting the dominance effect by the regression of phenotypic value on the inbreeding coefficient (Borrallho, 1994). **BID** is a regression coefficient expressing the reduction in phenotypic value per unit of inbreeding, in units of phenotypic standard deviation. 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 & Cotterill, 1988; Park & 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 **BID** = -2.0. Note that **BID** is ALWAYS a negative real number, and normally in the range of 0 to -4. Note also that there is no random within-family distribution of inbreeding depression; all full-sibs will have the same inbreeding coefficient and, thus, the same inbreeding depression.

EFFVAR(1) Additive genetic variance in the base population.

Specify the additive genetic variance of the trait (or index of traits) in the original base population of **NFndrs**.

EFFVAR(2) Dominance genetic variance in the base population

Specify the dominance genetic variance of the trait (or index of traits) in the original base population of **NFndrs**.

EFFVAR(3) Epistatic genetic variance in the base population

Specify the epistatic genetic variance of the trait (or index of traits) in the original base population of **NFndrs**.

EFFVAR(4) Environmental variance in base/recruitment population

Specify the environmental variance found in the base population. This variance is also applied to the generation of recruitment populations, as they are considered to be tested in a similar, undomesticated field environment.

Note that the user does not specify the trait heritability directly, as this is a function of the variances supplied, such that:

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_D^2 + \sigma_I^2 + \sigma_E^2}$$

To run scenarios at different heritability, the usual approach is to adjust the environmental variance, **EFFVAR(4)**.

PROGENY TESTING SPECIFICATIONS

PXF Progeny test option.

A progeny test may be required, either to perform selection of breeding parents from a pool of founders in the first cycle, or to support other breeding or seed orchard decisions across all cycles. The test can be performed as a “polycross”, where a controlled cross is made with a pollen mix, or by collecting “open-pollinated” seed (unknown pollen parents from the original “wild” population). The following options can be specified:

- 0 No progeny testing.
- 1 Polycross testing with assessment of PX family means.
- 2 Polycross testing with BLUP assessment (requires ASReml).
- 3 OP progeny testing with assessment of OP family means.
- 4 OP progeny testing with BLUP assessment (requires ASReml).

Note that when BLUP assessment is performed, the progeny test observations are carried forward and become part of the multi-generation dataset when BLUP is performed in later generations.

NP Number of PX pollen parents.

If PX testing is performed, a pollen mix will be generated as a sample of **NP** parents, drawn from the original base founder population.

NX Progeny test family size.

When progeny testing is performed, test families are generated with **NX** progeny per PX or OP family.

RgOP Pollen parent correlation.

When progeny testing is performed with OP families, the pollen parent may be genetically correlated with the seed parent (Squillace, 1974). **RgOP** describes this as phenomenon as a correlation coefficient, with a value between 0 and 1.

PXENV Environmental variance for progeny testing.

Specify the environmental variance to be applied to progeny testing. This allows the user to describe an environment different from that where the original founders or recruitment populations are assessed. If progeny testing is done as a short-term nursery test, this variance might typically be less than that found in field trial sites.

SPECIFICATIONS FOR GENERATING RECRUITMENT POPULATIONS

A recruitment population is generated each cycle by mating those individuals selected as parents, known as the “breeding population”.

ISORT Mate assortment option

Mates can be allocated at random, or they may be ranked and mated assortatively (positive assortative mating, or “PAM”). Options available are:

- 0 Random assortment of mates.
- 1 Assortment on progeny test family means (GCA values).
- 3 Assortment on phenotypic value (of seedling or mean of all ramets of a clone).
- 5 Assortment on the combined index of family mean and individual phenotype, weighted by their respective heritabilities (Baker, 1986; Falconer & Mackay, 1996).
- 7 Assortment on BLUP breeding values (requires ASReml).

RELFLAG Relative avoidance option.

Assortment may lead to mating of close relatives and often it is desirable to avoid this by making minor substitutions after assortment. This option is only available for balanced single-pair, double-pair, or triple-pair mating, in which case the **DESIGN** parameter **MUST** be set to 0 and the total number of crosses declared under **NXS**, below. Options available are:

- 0 No attempt to avoid mating close relatives.
- 1 Avoid mating half-sib relatives.
- 2 Avoid mating full-sib relatives.

DESIGN Mating design.

Generation of the recruitment population requires that individuals in the breeding population be mated by controlled pollination. Balanced single-pair, double-pair or triple-pair matings can be specified by setting **DESIGN** to 0, and declaring **NXS** equal to **NT**, $2 \times \text{NT}$, or $3 \times \text{NT}$, respectively. Otherwise, mating designs can be specified on the worksheets labelled MD1, MD2, ... MD10, and the sheet number (1, 2, ... 10) specified as the **DESIGN** parameter.

NXS Number of controlled crosses

If a balanced mating design is generated by **DESIGN** option 0, it will be necessary here to specify the total number of crosses among the breeding parents for single-pair, double-pair, or triple-pair matings, as described under **DESIGN**. If **DESIGN** has specified a user-declared design on one of the sheets

MD1, MD2, ..., MD10, this parameter will be generated automatically at execution time.

NC Number of genotypes tested per cross

This is the number of genotypes generated for a controlled-cross family in the recruitment population, or “family size”.

NR Number of ramets tested per genotype

This parameter describes the number of clonal replicates tested per genotype. If the recruitment population is propagated as seedlings, **NC** should be declared as 1.

SELECTION OF THE BREEDING POPULATION

IBPSEL Selection method for BP selection.

Numerous options are available for selection of breeding parent candidates from the recruitment population. These include standard approaches such as mass selection and combined-index (CI) selection (Baker 1986; Falconer & Mackay 1996), as well as combining gain and diversity targets with “Group-merit” selection (GMS) (Lindgren & Mullin, 1997)¹. The specific options available are:

- 0 Random (no selection is applied in the random sampling of **NT** parents).
- 1 Mass selection (selection is based on individual phenotypic value or clone mean).
- 3 CI selection with weights on family and individual values corresponding to their respective heritabilities, calculated from the test data.
- 4 CI selection with weights defined by a user-specified individual heritability given under **WEIGHT** (see below)
- 5 GMS on phenotype.
- 6 GMS on combined index.
- 7 GMS on BLUP breeding value (requires ASReml).
- 8 Mass selection on BLUP breeding value (requires ASReml).

WEIGHT User-specified individual heritability (CI selection) or weight on group coancestry (GMS).

If **IBPSEL** has the value of 4, the **WEIGHT** specified here is a user-provided individual heritability, which is used to calculate index weights, rather than using the individual heritability estimated from the test data.

¹ The user is referred to Appendices I and II for further information on combined-index selection and group-merit selection, respectively.

If **IBPSEL** has a value between 5 and 7, this **WEIGHT** is applied to the group coancestry in the calculation of group-merit value.

Otherwise, the value specified for **WEIGHT** is ignored.

NBPFS Maximum number of selections from a full-sib family.

NOTE: If the user has specified any **IBPSEL** method other than 5, 6, or 7 (all of which are forms of group-merit selection), restrictions can be imposed on the number of selections made from a given cross, the number of half-sib relatives for a given selection, and the number of selections related to a given parent: **NBPFS**, **NBPHS** and **NBPPC**, respectively. A value for any of these restrictions of 32 000 has the effect of ignoring the restriction. Otherwise, POPSIM™ will attempt to satisfy all other restrictions with values set at less than 32000. The user is warned that this can lead to unexpected results, as POPSIM™ may be forced to “relax” some restrictions to find selections that satisfy all restrictions. If the value specified for **IBPSEL** is 5, 6, or 7 (group-merit selection), **NBPFS**, **NBPHS** and **NBPPC** are ignored.

NBPFS declares the maximum number of selections that can be made from a given full-sib family.

NBPHS Maximum number of half-sib relatives.

POPSIM™ will restrict selection from the recruitment population such that no more than **NBPHS** previous selections are related to a selection candidate as half-sib relatives.

NBPPC Maximum number of progeny per parent.

POPSIM™ will restrict selection from the recruitment population such that any given parent of a selection candidate does not contribute any more than **NBPPC** selections.

PSType Preselection type.

Some breeding strategies require a form of pre-selection for one or more cycles. The following options are available:

- 0 No preselection imposed. The selection of NXS candidates is made directly from the recruitment population for all cycles
- 1 Preselection of **NT** × **PSF** candidates from the recruitment population is followed by estimation of candidate GCAs through progeny testing, in order to impose a final selection of **NT** breeding parents from among the preselected candidates.
- 2 This is a special option which allows **NT/PSF** initial founders to give rise to a breeding population of size **NT** after the first cycle. In other words, the F_1 recruitment population arises from **NT/PSF** founders, while the F_2 and subsequent recruitment populations arise from a larger breeding population of **NT** parents. This is a special option representing “Strategy 5” for the Scots pine programme in Sweden.

PreSelCyc Preselection cycles.

When **PSType** has been declared with a value of 1, the two-step preselection followed by progeny testing leading to final selection of the breeding population, is carried out for **PreSelCyc** cycles. After this number of cycles, POPSIM™ reverts back to **PSType** = 0 for all remaining cycles.

PSF Selection factor.

This parameter is used, as describe above under **PSType**, to determine the number of candidates per final selection to be progeny tested (**PSType**=1), or the expansion of original **NT/PSF** founders to form a breeding population of **NT** parents (**PSType**=2).

SELECTION OF SEED ORCHARD PRODUCTION POPULATION

All individuals selected into a breeding population are considered to have been archived and available for grafting into an open-pollinated seed orchard. POPSIM™ will report statistics for individuals selected into the orchard for each cycle, including the genetic gain expected under “ideal” orchard conditions, i.e., random panmictic mating, equal numbers of gametes per parent, and no impact of contaminating pollen from outside the orchard.

NSSO Number of selections for seed orchards.

NSSO individuals will be selected in each cycle to form a seed orchard production population.

ISOSEL Selection method advancing archived parents to seed orchards.

Selection to the orchard need not be based on the same criteria as applied to select the breeding population from the recruitment population, but only those individuals selected previously as members of a breeding population are eligible to advance to the orchard. The following options are available:

- 1 Forward selection on phenotype (or clone mean), i.e., without any additional progeny-test estimation of GCA.
- 2 Forward combined-index selection without any additional progeny-test estimation of GCA.
- 3 Forward selection on BLUP breeding value (estimated without additional progeny testing).
- 4 Backward selection on progeny-test GCA or BV estimates (as defined under **PXF** – see “Progeny testing specifications” above).

LGEN Selection across generations.

Candidates available for seed orchard establishment may be limited to those parents in the current cycle, or they may include breeding population members from previous cycles that have been maintained in archives. The options available are:

- 1 Orchard selections originate only from current cycle.
- 2 Orchard selections can originate from the current or any previous cycle (only available for **ISOSEL** = 3 or 4).

NSOFS Maximum number of orchard selections from a full-sib family.

NOTE: Restrictions can be imposed on the number of orchard selections made from a given cross, the number of half-sib relatives for a given selection, and the number of selections related to a given parent: **NSOFS**, **NSOHS** and **NSOPC**, respectively. A value for any of these restrictions of 32000 has the effect of ignoring the restriction. Otherwise, POPSIM™ will attempt to satisfy all other restrictions with values set at less than 32000. The user is warned that this can lead to unexpected results, as POPSIM™ may be forced to “relax” some restrictions to find selections that satisfy all restrictions.

NSOFS declares the maximum number of orchard selections that can come from a given full-sib family.

NSOHS Maximum number of half-sib relatives.

POPSIM™ will restrict selection to the orchard such that no more than **NSOHS** previous selections are related to a selection candidate as half-sib relatives.

NSOPC Maximum number of progeny per parent.

POPSIM™ will restrict selection to the orchard such that any given parent does not contribute any more than **NSOPC** selections.

SELECTION OF A FULL-SIB FAMILY PRODUCTION POPULATION

Tested families can also be used as a production population, either by repeating the cross with conventional controlled pollination or by mass control-pollinated orchards, and/or vegetative multiplication of seed (Bramlett, 1997; Bridgwater et al., 1998; Carson et al., 1992; Harbard et al., 1999; Sweet et al., 1992), commonly referred to as “family forestry” (Libby, 1992). POPSIM™ can assemble such a production population from the crosses tested in the current recruitment population.

NFAM Number of families selected for deployment.

This is the number of families selected for deployment among those tested in the current-cycle recruitment population. If a FS family production population is not required, set **NFAM** = 0.

NFAMHS Number of half-sib related families.

Once a family has been selected to the mixture, **NFAMHS** restricts how many families related at the half-sib level can be added to the mix.

NFAMPC Number of families with a common parent.

POPSIM™ will restrict selection such that any given parent does not contribute any more than **NFAMPC** families selected for deployment.

SELECTION OF A MIXTURE OF SELECTED CLONES

When the recruitment population has been deployed with clonal replicates, there is an opportunity to select the best clones and deploy these as a mixture. Regeneration of clones is problematic for many forest trees, as physiological ageing limits the feasibility of propagating stored material, unless the cloning has been by somatic embryogenesis and the cultures can be retrieved from cryo-storage (Park 2002). POPSIM™ can assemble such a production population from the genotypes clonally replicated in the current recruitment population.

NCLN Number of clones selected for deployment.

This is the number of clones selected for deployment among those genotypes clonally replicated in the current-cycle recruitment population. If a clonal production population is not required, set **NCLN** = 0.

ICLNSEL Method for selecting clones for the deployed mixture.

Selection to the deployed clone mix need not be based on the same criteria as applied to select the breeding population from the recruitment population, but only clones from the current-cycle's recruitment population are eligible. The following options are available:

- 1 Forward selection based on clone mean.
- 2 Forward selection on combined index value.
- 3 Forward selection on BLUP breeding value.

NCLNFS Maximum number of clones selected from a full-sib family.

NCLNFS declares the maximum number of clones that can be selected from a given full-sib family.

NCLNHS Maximum number of half-sib relatives.

In order for POPSIM™ to consider an clone for selection to the mixture, there must be no more than **NCLNHS** previous selections which are related to the candidate as half-sibs.

NCLNPC Maximum number of progeny per parent.

POPSIM™ will restrict selection of clones such that any given parent does not contribute any more than **NCLNPC** selections.

LOCAL FILE SETTINGS

The user must specify the location of the Excel workbook, where all temporary files will also be generated, and the “popsim.exe” file. If the user has installed the package with the recommended settings, these will be “C:\popsimfiles\” and “C:\popsim\”, respectively.

ASREML FILE SPECIFICATIONS

The use of ASReML with POPSIM™ is optional, but is required for any scenario where BLUP calculations are specified. The folder for work files will usually be the location of the Excel control workbook. The folder containing the ASReML.exe will depend on the ASReML installation, but will often be similar to “C:\Program Files\ASReML3\bin\”

Running a series of scenarios

Once the scenarios have been described on the Excel workbook MainPage, be sure to save as a macro-enabled workbook – a unique name can be assigned. To run POPSIM™, the user MUST open the workbook by double clicking on the file name – do NOT use the Excel Open → File command!

From the MainPage sheet, click on the “Run Simulations” macro button to launch the simulation. The user will be asked to specify which scenario to start the simulation – usually this would be “1”. The macro will conduct a general check of the data in the worksheet, to ensure that a readable control file called “input.txt” can be generated. The macro then calls “popsim.exe”, which is the main simulation engine. POPSIM™ will open a splash screen window which gives legal information, and also gives the “Build date” of the program. This date should be given in any correspondence requesting support. This window can be resized to allow other windows to be visible on the desktop, while the simulation is running. The Excel workbook can also be closed, as POPSIM™ does not require it to complete the simulation request.

POPSIM™ will conduct a thorough check of all control data for all scenarios. Any warnings or errors will be displayed, to guide in the correcting of scenario data. If the control data are error-free, the program will begin the simulation of the first scenario. The program provides an on-screen progress report, stating what step is being executed, in what iteration, generation and scenario, and also gives the total execution time since the program was launched.

Execution time can vary tremendously, depending on the dimensions, number of iterations requested, complexity and analysis procedures required. If BLUP calculations are required, the simulation will periodically make calls to ASReML, which will launch in a separate window. It is not unusual for scenarios to take several minutes per iteration, and several scenarios may require several hours, even days, to complete. If at all possible, avoid using the computer for other tasks while POPSIM™ is running. Multi-tasking response time will be poor, and you could inadvertently interrupt the simulation.

Once POPSIM™ has completed all iterations of all scenarios, the program will sound a “beep” and display a message on the screen that the “Program performed all tasks”. Hit ENTER to terminate the program and close the window.

Return to the Excel interface workbook to retrieve and view your results, as described below.

Retrieving and viewing output

Once POPSIM™ has reported that simulation scenarios are complete, the results are stored in a file called “output.txt”, but are best viewed through the Excel interface. Open the Excel macro-enabled control workbook used to launch the simulation, by double-clicking on the file name and accept the splash screen. Select the “Results” sheet, and click on the “Read Output” button to load the results.

POPSIM™ will read one line of results for each generation of each scenario. The results are arranged in columns describing the mean and standard deviation over all iterations, for each population parameter. The breeding population is listed first. Scroll to the right and you will find comparable data for the seed orchard and other production populations that may have been requested.

Saving scenario parameters and results

After completing a series of scenarios and loading the results into the Excel control data workbook, it is important to save a permanent record of the results and the control parameters that generated them. On the “Results” sheet, click on the “Save Results” macro button. Declare a suitable location and file name.

The resulting workbook contains 3 worksheets. The control data are stored in the first “Simulation Parameters” sheet, the results in the “Results” sheet, and a third “User Page” is provided as a place to conduct analyses of the results.

Analysing simulation results

Different population parameters will be of interest to different users, depending on the objective of their simulation. The usual analysis procedure involves graphing the development of different population parameters against time. The simulation results are ordered by generation, but the user may have different timelines on mind for different scenarios. Graphs can be prepared directly in the Results workbook, using the various tools provided in Excel. Macros can also be easily written in Excel to retrieve the parameters of interest and display these in custom formats or graphs, if many scenarios are to be analysed

When things go wrong!

The user should be aware that POPSIM™ is in a continuing state of development; new features are added, and others modified, in response to user-feedback and the needs of new research. Every user will ask different things of the simulation, and there will be times when development testing has not covered all possibilities. Unexpected things can happen!

POPSIM™ may provide messages from the Excel control data workbook or from the main program. Quite often, messages are generated when the user is attempting to declare an “illegal” combination of management parameters.

Reading the messages, combined with the parameter descriptions given in this manual, may help to identify the cause of the problem.

POPSIM™ has a large combination of parameters that can be controlled by the user, and not all combinations have been thoroughly tested during development. If POPSIM™ accepts all of the initial data as legal, there may still be times when the program crashes, with or without a message. When asking for help with such issues, it is very important to provide the following to help us reproduce and diagnose the problem:

1. Copy of the Excel control data workbook with the scenario(s) that caused the problem.
2. A full description and screen shot of any messages produced (the key combination Ctrl+PrtScn will put a screen shot in the clipboard, and this can then be pasted into a Word document or body of an email).
3. The “Build date” of the POPSIM main program (given on the main splash screen when the simulation starts), or the file date of your “popsim.exe” file.
4. Version of ASReml (if used).

Email this information to tim.mullin@skogforsk.se

Help to improve POPSIM™

POPSIM™ has developed with many ideas and contributions of tree breeders and geneticists in many countries. Users are encouraged to become engaged, to ask questions and to make suggestions for future capabilities. It is hoped that POPSIM™ will soon have a home on the internet, where European tree breeders can have access to future versions and obtain help from the POPSIM™ user community. Please visit periodically and contribute your ideas!

References

- Andersson, E.W., Spanos, K.A., Mullin, T.J. & Lindgren, D. 1998. Phenotypic selection can be better than selection for breeding value. *Scand. J. For. Res.* 13(1): 7–11.
- 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(5): 1079–1082.
- Bramlett, D.L. 1997. Genetic gain from mass controlled pollination and topworking. *J. For.* 95(3): 15–19.
- Bridgwater, F.E., Bramlett, D.L., Byram, T.D. & Lowe, W.J. 1998. Controlled mass pollination in loblolly pine to increase genetic gain. *For. Chron.* 74(2): 185–189.
- Carson, M.J., Vincent, T.G. & Firth, A. 1992. Control-pollinated and meadow seed orchards of radiata pine. *In Proceedings of IUFRO-AFOCEL Symposium on Mass Production Technology for Genetically Improved Fast Growing Forest Tree*

- Species, Bordeaux, France, 14–18 September 1992. AFOCEL, Nangis, France. pp. 100–109.
- Cockerham, C.C. 1967. Group inbreeding and coancestry. *Genetics*, 56: 89–104.
- Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to quantitative genetics*. 4th ed. Prentice Hall, Harlow, UK.
- Fisher, R.A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* 52: 399–433.
- Fraser, A.S. 1957. Simulation of genetic systems by automatic digital computers. I. Introduction. *Aust. J. Biol. Sci.* 19: 484–491.
- Gilmour, A.R., Gogel, B.J., Cullis, B.R. & Thompson, R. 2009. *ASReml User Guide*, Release 3.0. VSN International Ltd., Hemel Hempstead, UK.
- Griffin, A.R. & 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(3–4): 124–131.
- Harbard, J.L., Griffin, A.R. & Espejo, J. 1999. Mass controlled pollination of *Eucalyptus globulus*: a practical reality. *Can. J. For. Res.* 29(10): 1457–1463.
- Libby, W.J. 1992. Use of genetic variation for breeding forest trees. *In* *Plant breeding in the 1990s*. CAB International, UK, Wallingford. pp. 101–117.
- Lindgren, D., Danusevicius, D. & Rosvall, O. 2008. Balanced forest tree improvement can be enhanced by selecting among many parents but maintaining balance among grandparents. *Can. J. For. Res.* 38(11): 2797–2803.
- Lindgren, D., Gea, L. & Jefferson, P. 1996. Loss of genetic diversity monitored by status number. *Silvae Genet.* 45(1): 52–59.
- Lindgren, D. & Mullin, T.J. 1997. Balancing gain and relatedness in selection. *Silvae Genet.* 46(2–3): 124–129.
- Lstibůrek, M., Mullin, T.J. & El-Kassaby, Y.A. 2006. The impact of differential success of somatic embryogenesis on the outcome of clonal forestry programs. I. Initial comparison under multitrait selection. *Can. J. For. Res.* 36(6): 1376–1384.
- Lstibůrek, M., Mullin, T.J., Lindgren, D. & Rosvall, O. 2004a. Open-nucleus breeding strategies compared with population-wide positive assortative mating. I. Equal distribution of testing effort. *Theor. Appl. Genet.* 109(6): 1196–1203.
- Lstibůrek, M., Mullin, T.J., Lindgren, D. & Rosvall, O. 2004b. Open-nucleus breeding strategies compared with population-wide positive assortative mating. II. Unequal distribution of testing effort. *Theor. Appl. Genet.* 109(6): 1169–1177.
- Lstibůrek, M., Mullin, T.J., Mackay, T.F.C., Huber, D. & Li, B. 2005. Positive assortative mating with family size as a function of predicted parental breeding values. *Genetics*, 171(3): 1311–1320.
- 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. & Park, Y.S. 1995. Stochastic simulation of population management strategies for tree breeding: a new decision-support tool for personal computers. *Silvae Genet.* 44(2–3): 132–141.

- Park, Y. 2002. Implementation of conifer somatic embryogenesis in clonal forestry: technical requirements and deployment considerations. *Annals of Forest Science* 59(5/6): 651–656.
- Park, Y.S. & 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(1): 17–21.
- Rosvall, O., Lindgren, D. & Mullin, T.J. 1998. Sustainability, robustness and efficiency of a multi-generation breeding strategy based on within family clonal selection. *Silvae Genet.* 47(5–6): 307–321.
- Rosvall, O. & Mullin, T.J. 2003. Positive assortative mating with selection restrictions on group coancestry enhances gain while conserving genetic diversity in long-term forest tree breeding. *Theor. Appl. Genet.* 107(4): 629–642.
- Rosvall, O., Mullin, T.J. & Lindgren, D. 2003. Controlling parent contributions during positive assortative mating and selection increases gain in long-term forest tree breeding. *For. Genet.* 10(1): 35–53.
- 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*. 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.
- Squillace, A.E. 1974. Average genetic correlations among offspring from open-pollinated forest trees. *Silvae Genet.* 23(5): 149–156.
- Sweet, G.B., Dickson, R.L., Donaldson, B.D. & Litchwark, H. 1992. Controlled pollination without isolation - a new approach to the management of radiata pine seed orchards. *Silvae Genet.* 41(2): 95–99.
- Verryn, S.D., Snedden, C.L. & Parfitt, R.C. 2000. Program for the deterministic modelling of genetic gains of tree breeding and seed and clone production strategies. *Southern African Forestry Journal*(189): 3–9.
- Weng, Y.H., Park, Y.S., Simpson, D. & Mullin, T.J. 2009. Efficiencies of clonally replicated and seedling testing for spruce breeding and deployment strategies. *Silvae Genet.* 58(5/6): 292–299.
- White, T.L., Adams, W.T. & Neale, D.B. 2007. *Forest genetics*. CABI Publishing, Wallingford, Oxfordshire, UK.

Appendix 1

How POPSIM™ calculates “combined index”

“Combined-index selection” is based on an individual index value (CI) that weights family and individual performance by their respective heritabilities (Falconer & MacKay, 1996):

$$CI = h_w^2 (\bar{X}_{ifm} - \bar{X}_{fm}) + h_F^2 (\bar{X}_{fm} - \bar{X})$$

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

$$h_F^2 = h^2 \frac{1 + (n-1)r}{1 + (n-1)t} ;$$

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

$$h_w^2 = h^2 \frac{(1-r)}{(1-t)} ;$$

h^2 is the narrow-sense heritability of individual observations; r is the genotypic intraclass correlation, which is $\frac{1}{2}$ for 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:

$$h_F^2 = h^2 \frac{1 + \frac{1}{2}(n-1)}{1 + \frac{1}{2}(n-1)h^2} \quad \text{and} \quad h_w^2 = h^2 \frac{\frac{1}{2}}{1 - \frac{1}{2}h^2}$$

Appendix 2

How POPSIM™ calculates “group merit”

Group-merit selection is a technique first described by Lindgren & Mullin (1997), where the objective is to maximise gain while constraining relatedness so that genetic diversity is conserved. The approach is to define a selection criterion for a set of genotypes that combines both breeding value and diversity, and then to apply a procedure that searches for an optimal trade-off between them.

For the purpose of calculating group merit, breeding value (g) can be estimated in any of a number of ways. The diversity parameter used in the selection criterion is “group coancestry” (Θ), which is the probability that two genes taken at random from the gene pool, with replacement, are identical by descent (Cockerham, 1967).

The selection criterion, B_ω , for a set of genotypes, ω , can then be formulated as

$$B_\omega = \bar{g}_\omega - c\Theta_\omega$$

where, c is a weighting constant, \bar{g}_ω is the average of their breeding values, and Θ_ω is the average coancestry of the considered genotypes, which is a measure of their relatedness. Group-merit selection is simply a procedure to search for the set of ω genotypes that maximises B_ω .

A candidate for selection is penalised by the group coancestry that would result if it were included in the population together with the already selected members. The advantage a candidate offers is compared with other candidates for a particular stage of selection. Similarly, one can also compare individuals considered for substitution of already selected genotypes, and thus search for a set of members that maximises the group merit. The penalty weight (c) can be set to give the desired balance between gain and relatedness. A penalty of zero would obviously maximise expected gain, while an infinite penalty would minimise relatedness and an infinitely negative penalty would maximise relatedness.

The search procedure applied by POPSIM™ to select a set of N individuals is:

1. Breeding values are predicted for each candidate, by conventional methods such as combined-index, BLUP, etc.
2. The penalty weight c is chosen by the user and supplied to POPSIM™.

3. The subtraction term c^{Θ}_{ω} for each family is calculated, in turn, by considering the resulting group coancestry of the selected population ω , after the addition of another individual from this family. This term is subtracted from the predicted breeding values for each member of the family to estimate the group merit after the addition of that individual.
4. The candidate individuals are ranked by the selection criterion and the highest-ranking candidate is added to the selected population.
5. Steps 3 and 4 are repeated until N selections have been made, and the group merit of this set is recorded.
6. Steps 3 and 4 are repeated to add an $N+1$ th individual to the selected set.
7. Each of the $N+1$ selected genotypes is then removed, in turn, from the selected population and the group merit of the remaining N individuals calculated. The new set of N individuals giving the maximum group merit is identified.
8. If the group merit given by this new set of N individuals is less than before, the previous set of N individuals is retained and the selection algorithm is complete. Otherwise, this new set of N individuals is retained, and Steps 6 through 8 are repeated.

What this algorithm actually does is to first select N individuals consecutively, and then to substitute earlier selections by alternatively adding and discarding individuals until the process fails to find a set of N individuals with higher group merit. Note that a family will always be represented by its top-ranking phenotypes.

Arbetsrapporter från Skogforsk fr.o.m. 2009

| År 2009 | |
|---------|--|
| Nr 699 | Almqvist, C., Eriksson, M. & Gregorsson, B. 2009. Cost functions for variable costs of different Scots pine breeding strategies in Sweden. 12 s. |
| Nr 670 | Andersson, M. & Eriksson, B. 2009. HANDDATORER MED GPS. För användning vid röjningsplanläggning och röjning. 25 s. |
| Nr 671 | Stener, L.G. 2009. Study of survival, growth, external quality and phenology in a beech provenance trial in Rånna, Sweden. 12 s. |
| Nr 672 | Lindgren, D. 2009. Number of pollen in polycross mixtures and mating partners for full sibs for breeding value estimation. 15 s. |
| Nr 673 | Bergkvist, I. 2009. Integrerad avverkning av grotbuntar. 21 s. |
| Nr 674 | Rosvall, O. 2009. Kompletterande strategier för det svenska förädlingsprogrammet. 26 s. |
| Nr 675 | Arlinger, J., Barth, A. & Sonesson, J. 2009. Förstudie om informationsstandard för stående skog. 21 s. |
| Nr 676 | Nordström, M. & Möller J. J. 2009. Den skogliga digitala kedjan – Fas 1. 38 s. |
| Nr 677 | Möller J.J., Hannrup, B., Larsson, W., Barth, A. & Arlinger, J. 2009. Ett system för beräkning och geografisk visualisering av avverkade kvantiteter skogsbränsle baserat på skördardata. 36 s. |
| Nr 678 | Enström, J. & Winberg, P. 2009. Systemtransporter av skogsbränsle på järnväg. 27 s. |
| Nr 679 | Iwarsson Wide, M. & Belbo, H. 2009. Jämförande studie av olika tekniker för skogsbränsleuttag. – Skogsbränsleuttag med Naarva-Gripen 1500-40E, Bracke C16.A och LogMax 4000, Mellanskog, Färila. 43 s. |
| Nr 680 | Iwarsson Wide, M. 2009. Jämförande studie av olika metoder för skogsbränsleuttag. Metodstudie – uttag av massaved, helträd, kombinerat uttag samt knäckkvistning i talldominerat bestånd, Sveaskog, Askersund. 25 s. |
| Nr 681 | Iwarsson Wide, M. 2009. Teknik och metod Ponsse EH25. – Trädbränsleuttag med Ponsse EH25 i kraftledningsgata. 14. |
| Nr 682 | Iwarsson Wide, M. 2009. Skogsbränsleuttag med Bracke C16. – Bränsleuttag med Bracke C16 i tall respektive barrblandskog. 14 s. |
| Nr 683 | Thorsén, Å. & Tosterud, A. 2009. Mer effektiv implementering av FoU-resultat. – En intervjuundersökning bland Skogforsks intresenter. 58 s. |
| Nr 684 | Rytter, L., Hannerz, M., Ring, E., Högbom, L. & Weslien, J.-O. 2009 Ökad produktion i Svenska kyrkans skogar – Med hänsyn till miljö och sociala värden. 94 s. |
| Nr 685 | Bergkvist, I. 2009. Skördarstorlek och metod i förstagallring av tall och gran – studier av prestation och kvalitet i förstagallring. 29 s. |
| Nr 686 | Englund, M. 2009. Röststyrning av aggregatet på en engreppsskördare – En Wizard of Oz-studie. 32 s. |
| Nr 687 | Lindgren, D. 2009. Polymix breeding with selection forwards. 14 s. |
| Nr 688 | Eliasson, L., Nordén, B. 2009. Fyra olika studier med A-gripen. 31 s. |
| Nr 689 | Larsson, F. 2009. Skogsmaskinföretagarnas kundrelationer, lönsamhet och produktivitet. Under bearbetning. 44 s. |
| Nr 690 | Jönsson, P., Löfroth, C. & Englund, M. 2009. Förarstol för stående arbetsställning – en pilotstudie. 12 s. |
| Nr 691 | Brunberg, T., Lundström, H. & Thor, M. 2009. Gallringsstudier hos SCA vintern och sommaren 2009. 26 s. |
| Nr 692 | Eliasson, L. & Johannesson, T. 2009. Underväxtens påverkan på bränsleanpassad slutavverkning – Studie från avverkning hos Sca Skog AB. 11 s. |
| Nr 693 | Nordén, B. & Eliasson, L. 2009. En jämförelse av ett Hugglinksystem med en traktormonterad flishugg vid flisning på avlägg. 9 s. |
| Nr 694 | Hannrup, B. et al., 2009. Utvärdering av ett system för beräkning och geografisk visualisering av avverkade kvantiteter skogsbränsle. 42 s. |
| Nr 695 | Iwarsson Wide, M. 2009. Skogsbränsleuttag i vägkanter. Prestationsstudie – uttag av Skogsbränsle i vägkant med BRACKE C16. 14 s. |
| Nr 696 | Iwarsson Wide, M. 2009. Skogsbränsleuttag i vägkanter. Prestationsstudie – uttag av Skogsbränsle i vägkant med ponsse dual med EH 25. 15 s. |

| | |
|----------------|--|
| Nr 697 | Almqvist, C. & Wennström, U. 2009. Granfröplantageskötselresa 2009-08-31–200-09-03. Noter från besök i respektive plantage. 22 s. |
| Nr 698 | Wilhelmsson, L. m.fl. 2009. D3.1 Initial analysis of drivers and barriers. 41 s. |
| Nr 699 | Wilhelmsson, L. m.fl. 2009. D3.2 Existing models and model gap analyses for wood properties. 54 s. |
| År 2010 | |
| Nr 700 | Hannerz, M. & Cedergren, J. 2010. Attityder och kunskapsbehov – förädlad skogsodlingsmaterial. 56 s. |
| Nr 701 | Rytter, R.M. 2010. Detektion av röta i bokved – resultat av mätthöjd, riktning och tidpunkt. 10 s. |
| Nr 702 | Rosvall, O. & Lindström, A. 2010. Förädlings effekter i Sveriges skogar - kompletterande scenarier till SKA-VB 08. 31 s. |
| Nr 703 | von Hofsten, H. 2010. Skörd av stubbar – nuläge och utvecklingsbehov. 18 s. |
| Nr 704 | Karlsson, O. & Nisserud, F. 2010. Utveckling av en dynamisk helfordonsmodell för skotare. 73 s. |
| Nr 705 | Eliasson, L. & Johannesson, T. 2010. Förröjningens påverkan på grotskotning – En studie av produktivitet, ekonomi, grotkvalitet hos SCA skog. 9 s. |
| Nr 706 | Rytter, L. & Stener L.G. 2010. Uthållig produktion av hybridasp efter skörd – Slutrapport 2010 för Energimyndighetens projekt 30346. 23 s. |
| Nr 707 | Bergkvist, I. 2010. Utvärdering av radförbandsförsök anlagda mellan 1982-1984. 16 s. |
| Nr 708 | Hannrup, B. & Jönsson, P. 2010. Utvärdering av sågmotorn F11-iP med avseende på uppkomsten av kapsprickor – en jämförande studie. 28 s. |
| Nr 709 | Iwarsson Wide, M., Belbo, H. 2010. Jämförande studie av olika tekniker för skogsbränsleuttag i mycket klen skog Skogsbränsleuttag med Naarva-Gripen 1500-40E och Log Max 4000, Mellanskog, Simeå 28 s. |
| Nr 710 | Englund, M., Löfroth, C. & Jönsson, P. 2010. Inblandning av rött ljus i LED-lampor – Laboratoriestudier av hur människor uppfattar tre olika ljusblandningar. 7 s. |
| Nr 711 | Mullin, T.J., Hallander, J., Rosvall, O. & Andersson, B. 2010. Using simulation to optimise tree breeding programmes in Europe: an introduction to POPSIM™. 28 s. |
| Nr 712 | Jönsson, P. 2010. Hydrauliskt dämpad hytt – ett lyft för arbetsmiljön? 14 s. |
| Nr 713 | Eriksson, B. & Sonesson, J. 2010. Tredje generationen skogsbruksplaner – Slutrapport DElproj 4 – Arbetsgång vid planläggning. 23 s. |
| Nr 714 | Sonesson, J. 2010. Nya arbetssätt i skogsbruksplanläggning. |
| Nr 715 | Eliasson, L. 2010. Huggbilar med lastväxlarssystem. 13 s. |
| Nr 716 | Eliasson, L. & Granlund P. 2010. Krossning av skogsbränsle med en stor kross – En studie av CBI 8400 hos Skellefteå Kraft. 6 s. |
| Nr 717 | Stener, L.G. 2010. Tillväxt, vitalitet och densitet för kloner av hybridasp och poppel i sydsvenska försök. 46 s. |
| Nr 718 | Palmquist, C. & Sandberg, J. & Vibrationskomfort och ergonomi på förarstolar i skotare. |
| Nr 719 | Thor, M. 2010. Avverkning och hantering av virke och avverkningsrester vid angrepp av tallvedsnematoder i svensk skog. 42 s. |
| Nr 720 | Fogdestam, N. 2010. Studier av Biotassu Griptilt S35 i gallring. 11 s. |