ARBETSRAPPORT 965-2018

Assembling optimum breeding populations for the Swedish Scots pine breeding program

Sammansättning av optimala förädlingspopulationer inom det svenska förädlingsprogrammet för tall

skogforsk

Tim J. Mullin
Torgny Persson

Sammanfattning

De senaste förbättringarna i avelsvärderingsarbetet vid Skogforsk gör det nu möjligt att generera jämförbara avelsvärden för alla besläktade träd i en stamtavla för en given region. Även om skogsträdsförädlare nu enkelt kan producera listor över kandidatträd och deras avelsvärden kan de inte bara ta de bästa träden för att generera nästa testgeneration utan måste även ta hänsyn till graden av släktskap mellan utvalda träd. Den optimala lösningen är inte att helt undvika släktskap utan snarare att hitta den uppsättning träd som maximerar den genetiska vinsten givet en begränsning vad gäller högsta tillåtna genomsnittliga släktskap. Ett förbättrat hjälpmedel vid arbetet med optimeringen av balansen mellan vinst och diversitet vid ett urval har nyligen blivit tillgängligt genom implementering av nya optimeringsalgoritmer i urvalsverktyget OPSEL.

I den här studien undersökte vi hur OPSEL på bästa sätt kunde utnyttjas vid framställning av nästa testgeneration inom ett antal förädlingspopulationer av tall i norra Sverige. Vi sammanställde data på tillgängliga kandidatträd från de tre förädlingspopulationerna, Tpop T4, T7 och T11, samt utvärderade lämplig storlek på operativ korsningsinsats och vilken diversitetsbegränsning som skulle tillämpas för att uppnå uppsatta mål vad gäller genetisk bevarande. Vi optimerade sedan valet av föräldragenotyper till nästa testgeneration under olika urvalsscenarier, inklusive så kallat "bakåtval" (enbart ursprungliga plusträd), "framåtval" (endast avkomma från plusträd) och urval där både plusträd och deras avkommor fick ingå – där det sistnämnda scenariot konsekvent producerade det bästa resultatet. Vi utformade dessutom ett verktyg som underlättar planeringen av efterföljande korsningsarbete, med huvudsaklig uppgift att minimera släktskapet mellan utvalda föräldrar i varje enskild parkorsning.

skogforsk

Uppsala Science Park, 751 83 Uppsala
skogforsk@skogforsk.se
skogforsk.se

ISSN 1404-305X

Preface

This research has received funding from (1) Föreningen Skogsträdsförädling (The Swedish Tree Breeding Association), (2) the European Union Horizon 2020 Research and Innovation Program under grant agreement no. 676876 (Project GenTree), and (3) Skogforsk (The Swedish Forestry Research Institute). All trademark terms are the property of their respective owners.

Sävar 2017-09-22

Tim J. Mullin and Torgny Persson

Contents

Summary

Recent improvements in our breeding value estimation at Skogforsk makes it possible to generate comparable breeding values for all relatives in the entire pedigree for a given region. While breeders can now easily produce lists of candidate trees and their breeding values, they cannot simply take the best trees to generate the next cycle of test progeny, but must also take into account the degree of relatedness among selections. The optimal solution is not to completely avoid kinship, but rather to find the set of selections that maximizes gain under a relatedness constraint. This optimum balance of gain and diversity during selection has only recently been solved in a definitive way by the application of optimization algorithms in a selection tool called OPSEL.

In this study, we look at how best to utilize this new capability while forming the next-cycle breeding populations for Scots pine in northern Sweden. We compiled data on candidates from three sub-populations, Tpop T4, T7 and T11. We evaluated the size of the operational crossing effort that should be planned for each breeding region and the appropriate constraint on diversity that should be applied to achieve program goals for genetic conservation. We then optimized selection of the genotypes contributing to the next cycle under different scenarios, including so-called "backward selection" (only founders), "forward selection" (only progeny of founders) and selection from across all previous cycles – the latter consistently produced the best result. We also designed a tool to plan the crossing of these genotypes in a positive assortative fashion, respecting the optimum contributions identified for each, and avoiding excessive relatedness between parents in each cross.

Introduction

Much research has focussed on development of “optimum strategies” for managing tree breeding populations. In Sweden, breeding of pine and spruce is done in the context of a multi-population structure of independent subpopulations targeting a particular photoperiod and temperature climate (Danell, 1993). The challenge is to pursue gain in each subpopulation while conserving diversity for robust deployment and as the source of future gains. Exactly how this should be done has been the subject of many theoretical studies using deterministic approaches (e.g., Danusevičius and Lindgren, 2002a, b; Lindgren et al. 2008; 1993; 1996; Ruotsalainen and Lindgren, 1998; 2000; Wei, 1996; Wei and Lindgren, 1991; 1994; 2001; Wei et al., 1997). The development of computer tools, such as POPSIM (Mullin et al. 2010; Mullin and Park, 1995) has also allowed the topic to be studied by means of computer simulation (e.g., Andersson et al., 1999; Andersson et al., 1998a; Andersson et al., 1998b; Lindgren and Mullin, 1997; Lstibůrek et al., 2004a, b; Lstibůrek et al., 2005; Rosvall, 1999; Rosvall and Andersson, 1999; Rosvall et al., 1998; Rosvall and Mullin, 2003; Rosvall et al., 2003).

In Sweden, these studies have formed the basis of breeding strategy evaluations and recommendations for the primary species, Scots pine and Norway spruce (Rosvall et al. 2011; 2015; Rosvall and Mullin, 2013; Ruotsalainen and Persson, 2013; Westin and Haapanen, 2013). In general, there is agreement that the formerly recommended two-step breeding strategy involving backward selection of progeny tested candidate parents, followed by controlled crossing by double-pair mating (DPM) and balanced selection within families, is very time consuming and probably not as efficient as moving ahead more quickly with forward selection, particularly in combination with clonally replicated field tests and with some imbalance during testing and selection. The simulations suggest that reductions in cycle time from 33 to 21 years can be accompanied by 30 to 50 percent additional gain per year at the same annual investment.

Recent implementation of TREEPLAN® at Skogforsk makes it possible to generate comparable breeding values for all relatives in the entire pedigree for a given region. While the TREEPLAN® system produces ranked lists of candidate trees and their breeding values, breeders cannot simply take the best trees, but must also take into account the degree of relatedness among selections. Managing relatedness among selections is already complicated after the first cycle of breeding, when parents, siblings and other relatives have similar ranks. The optimal solution is not to completely avoid kinship, but rather to find the set of selections that maximises gain under a relatedness constraint. This optimum balancing of gain and diversity during selection among TREEPLAN® breeding values has only recently been solved in a definitive way by the application of optimization algorithms in a selection tool called OPSEL (Mullin 2014, 2017a; Mullin and Belotti, 2016; Mullin et al., 2016; Yamashita et al., 2018).

While much work has been done theoretically and through simulation, many of these ideas have yet to be implemented in actual sub-populations, each of which have their own peculiarities and unique history of selection, mating and testing. Populations considered in theoretical studies tend to be balanced and sampled from uniform distributions representing the “average” situation. Actual populations tend to somewhat fragmented, with different rates of progress within the population and different intensities of testing among the founders and their offspring. Add to this the complication of shifting from a two-step backward-selection breeding strategy, to a faster one-step strategy, the possible stumbling blocks and implementation issues are many.

Our goal in this study was to gain insight on the severity of these issues and the best ways to handle them, by implementing the newer strategies in the context of actual sub-populations. The Tpop T4, T7 and T11 Scots-pine breeding populations in northern Sweden are at the stage to move forward with the next cycle of breeding and testing. Our objective was to implement Best Linear Unbiased Prediction (BLUP) breeding value estimation for the Tpop T4, T7 and T11 populations as case study, and to apply optimum selection to assemble possible breeding populations to carry forward into crossing for the next recruitment population.

Methods and materials

COMPILATION OF CANDIDATE BREEDING VALUES

The project utilized selection, pedigree and phenotypic data from Tpop T4, T7 and T11. Using data already in the DATAPLAN system, we prepared breeding value estimates for the target deployment regions, as well as for deployment areas for adjacent breeding populations. The breeding values were transferred to candidate lists on which various selection options were applied.

The breeding progress within Tpop T4, T7 and T11 can be described as follows:

- Tpop T4: 68 Swedish and 143 Finnish founder trees were originally allocated to the breeding population. The average geographic origin of the founder trees is 66.9°N and the centre of the target area for the breeding population is set to 66.5°N on low elevation areas. Half-sib progenies of the founders (Swedish material: open-pollinated seed collected in the original stands, Finnish material: seed from controlled crosses using a common pollen mix of 19 Swedish founder trees) were planted in five field tests in 1986. Based on progeny-test results, 70 founders were selected and crossed in 1998, producing 82 full-sib F1 families that were planted in 2001 in an archive at the Sävar breeding station and in two field trials (Table 1). The archive and field trials were measured 2015.
- Tpop T7: 103 Swedish and 107 Finnish founder trees were originally allocated to the breeding population. The average geographic origin of the founder trees is 66.1°N and the centre of the target area for the breeding population is set to 65.0°N and 100 m in altitude. Half-sib progenies of the founders (Swedish material: open-pollinated seed collected in the original stands, Finnish material: seed from controlled crosses using a common pollen mix of 22 Swedish founder trees) were planted in five field tests in 1986. Based on progeny-test results, 69 founders were selected and crossed in 1998, producing 71 full-sib F1 families that were planted in 2001 in an archive at the Sävar breeding station and in three field trials (Table 1). The archive and field trials were measured 2015 and 2011, respectively.
- Tpop T11: 212 Swedish founder trees were originally allocated to the breeding population. The average geographic origin of the founder trees is 63.6°N and the centre of the target area for the breeding population is set to latitude 62.5°N and 325 m in altitude. A first batch of 130 founder trees were crossed over the period 1963–1979, producing 594 F1 families that were grown and planted in 12 field trials over the period 1970–1980 (Table 1). Forward selection was performed in 1986–1988, and 100 selected F1 individuals from six of the trials were crossed in 1988, producing 102 full-sib F2 families. The full-sib F2 progeny were grown and planted in three field trials in 1992 (Table 1). Six of the 12 F1 trials were measured in 2002, and the remaining F1trials and the three F2 trials were measured in 2012. In addition, 38 founder trees were crossed 1977 (polycross, using a common pollen mix from 26 founder trees) producing 38 half-sib families which were planted in four field trials 1980 and measured year 1993. Also, open-pollinated offspring from 37 trees (grown from seed collected in the original plus tree stands) were put out in progeny trials 1996 and measured in 2005.

Table 1. Description of the tests and the progeny material included in the study, based on information from the latest inventory.

	Test location		**Test plot**	**Narrow sense heritability**		**No Genotypes**	**No families**		**Average family size**
Test	Lat.°N	Long.°E	Type	Survival	Height		F1	F2	
T4 F1 trial 1	66.3	21.0	Single-tree	0.08±0.03	0.06±0.03	*2061*	214	–	9.6
T4 F1 trial 2	64.9	18.8	Single-tree	0.14±0.03	0.24±0.05	*2976*	198	–	15.0
T4 F1 trial 3	65.0	21.1	Single-tree	0.05±0.04	0.15±0.08	*1262*	190	–	6.6
T4 F1 archive	63.9	20.6	Row family	–	0.14±0.04	*1162*	82	–	14.2
T4 F1 trial 4	65.0	21.0	Multiple-tree	0.05±0.03	0.14±0.05	*5132*	32	–	160.4
T4 F1 trial 5	66.4	20.7	Single-tree	–	0.24±0.06	*1233*	72	–	17.1
T7 F1 trial 1	66.3	21.0	Single-tree	0.12±0.06	0.21±0.08	*1013*	171	–	5.9
T7 F1 trial 2	64.9	18.8	Single-tree	–	0.41±0.07	*2406*	189	–	12.7
T7 F1 trial 3	65.0	21.1	Single-tree	–	0.34±0.07	*1897*	209	–	9.1
T7 F1 trial 4	63.6	15.6	Single-tree	–	0.29±0.10	*1162*	180	–	6.5
T7 F1 trial 5	64.0	20.6	Single-tree	–	0.51±0.07	*2904*	204	–	14.2
T7 F1 archive	63.9	20.6	Row family	–	0.18±0.06	*848*	70	–	12.1
T7 F1 trial 6	64.0	19.9	Multiple-tree	–	0.58±0.24	*1296*	8	–	162.0
T7 F1 trial 7	64.4	19.1	Single-tree	0.05±0.03	0.19±0.06	*1074*	61	–	17.6
T7 F1 trial 8	64.1	20.1	Single-tree	0.12±0.04	0.62±0.11	*1018*	61	–	16.7
T11 F1 trial 1	63.6	16.8	Single-tree	0.13±0.06	0.29±0.08	*1620*	49	–	33.1
T11 F1 trial 2	63.2	17.6	Single-tree	0.30±0.10	0.22±0.08	*1034*	46	–	22.5
T11 F1 trial 3*	62.3	15.4	Single-tree	0.04±0.01	0.28±0.06	*5771*	172	–	33.6
T11 F1 trial 4	62.8	17.3	Single-tree	0.05±0.01	0.16±0.04	*2541*	181	–	14.0
T11 F1 trial 5	63.6	17.4	Single-tree	0.08±0.02	0.20±0.05	*1781*	178	–	10.0
T11 F1 trial 6	63.4	17.6	Single-tree	0.05±0.02	0.05±0.02	*1020*	71	–	14.4
T11 F1 trial 7	63.4	17.6	Single-tree	0.08±0.02	0.13±0.03	*1548*	201	–	7.7
T11 F1 trial 8	64.0	19.9	Single-tree	0.06±0.02	0.09±0.02	*3330*	199	–	16.7
T11 F1 trial 9*	62.8	17.1	Single-tree	0.03±0.01	0.11±0.03	*7053*	200	–	35.3
T11 F1 trial 10	65.0	20.4	Single-tree	0.23±0.05	0.19±0.04	*2934*	157	–	18.7
T11 F1 trial 11	64.0	19.3	Single-tree	0.04±0.01	0.06±0.02	*1691*	150	–	11.3
T11 F1 trial 12	63.9	20.6	Single-tree	0.02±0.01	0.17±0.04	*1794*	150	–	12.0
T11 F2 trial 13	63.0	16.7	Single-tree	0.03±0.03	0.11±0.04	*1570*	–	102+	10.6
T11 F2 trial 14	62.8	14.9	Single-tree	0.04±0.03	0.04±0.03	*1255*	–	99+4	8.7
T11 F2 trial 15	62.8	14.4	Single-tree	–	0.16±0.05	*1770*	–	101+	12.0
T11 F1 trial 16	63.4	17.6	Single-tree	0.1±0.03	0.08±0.03	*534*	38	–	14.1
T11 F1 trial 17	63.1	17.0	Single-tree	0.05±0.02	0.25±0.06	*2334*	38	–	61.4
T11 F1 trial 18	62.7	17.1	Single-tree	0.03±0.03	0.16±0.05	*1513*	38	–	39.8
T11 F1 trial 19	62.5	15.4	Single-tree	0.03±0.02	0.22±0.06	*1939*	38	–	51.0
T11 F1 trial 20	63.5	19.3	Single-tree	–	0.24±0.05	777	37+2	–	19.7
T11 F1 trial 21	63.4	16.7	Single-tree	0.06±0.02	0.31±0.04	715	32+2	–	20.8
T11 F1 trial 22	63.7	16.7	Single-tree	0.06±0.03	0.32±0.05	397	28+2	–	12.9
T11 F1 trial 23	64.5	17.2	Single-tree	0.19±0.04	0.24±0.05	375	28+2	–	12.2

During the field visit, it appeared that most trees in the trial had been stormed.

For each of the Tpop T4, T7 and T11 populations, the work began with the assembly of all available founder trees and all surviving genotypes in associated F1 and F2 field trials into three separate candidate lists. Not only the founder trees originally allocated to Tpop 4, 7 and 11, but all northern Swedish founder trees (Tpop T1-T12), regardless of original breeding population assignment, were nominated as potential candidates for each of the three candidate lists.

Target deployment breeding values for Tpop T1-T12, latitude of origin for the founder tree and information on exact location in archives and test sites were added to all candidates in the lists. The breeding values used for ranking were calculated as the sum of the breeding values of the breeding objectives for quality and per-hectare volume production (in their separate deployment regions) weighted by economic relevance. The candidate lists were then checked if a given founder tree had been involved in some crossing activity in another breeding population in the past and if so excluded from the list. Thereafter, the candidate lists were checked that no ancestor were missing, i.e. checking that the pedigree was complete.

HOW MANY TO SELECT? WHAT DIVERSITY TARGET?

Optimizing selection involves finding a group of genotypes and their contribution frequencies that maximizes genetic gain, while still satisfying the requirement for operational crossing under a specified diversity constraint. Basic parameters defining the selection task are how many genotypes should be selected and the appropriate target level of diversity.

In the strategy originally proposed by Danell (1993), he suggested that totally balanced double-pair mating (two crosses per selected parent) and selection of 50 breeding parents per cycle would maintain a variance effective population size $N_{eV} = 100$. This could be achieved by selecting one individual from each of 50 families, then mating each of these exactly twice to create a recruitment population of 50 families. Indeed, this has been accepted practice in Sweden for many years. Danell felt that this would be a rather conservative approach in the Swedish context, which could perhaps be safely relaxed to $N_{eV} = 50$.

In the simulation work reported by Rosvall et al. (2015), it was found that “expanding” the breeding population by selecting and breeding a greater number of parents could result in additional gain, even under circumstances where the group coancestry of the selected population was kept constant. After consulting with Skogforsk breeders, we determined that doubling the number of crosses per breeding cycle from 50 to 100 was operationally feasible and, in many instances, similar to the number of crosses currently implemented, as additional crosses are often made as a “safety buffer” against cross failure. We adopted 100 crosses as the operational specification to be met by our optimization.

The diversity of our selected breeding population is not defined by its census size, but rather the average relatedness among the selections, their “Group Coancestry” (GC) (Cockerham, 1967). In selecting 50 unrelated, progeny-tested founders, the sub-population starts with Status Number $N_S = 50$, or GC = 0.01. Table 2 illustrates the subsequent reduction of N_S and accumulated GC with each cycle of selection. To maintain $N_{eV} = 100$, we accumulate GC at the rate of 0.005 per cycle, whereas to maintain $N_{eV} = 50$ the rate of GC increase is doubled. It follows from the study by that an intermediate constraint on diversity would be appropriate to maintain $N_{eV} = 75$. Our target for selecting the next-cycle breeding parents was thus $N_S \geq 20$, or GC ≤ 0.025.

Table 2. Target GC and Ns for breeding populations that maintain inbreeding effective population size of 50 to 100. The target constraint on diversity used in this study is in large bold italic font.

	N_{ev} = 100		N_{ev} = 75		N_{ev} = 50	
Cycle	**N_s**	**GC**	**N_s**	**GC**	**N_s**	**GC**
0	50	0.0100	50	0.0100	50	0.0100
1	33	0.0150	28.6	0.0175	25	0.0200
2	25	0.0200	***20***	***0.0250***	17	0.0300

IMPROVEMENTS TO THE OPTIMAL SELECTION ALGORITHM

Our optimization approach initially adopted the concept of using 100 genotypes equally to generate the 100 families. This is the optimization approach described by Rosvall et al. (2015) and utilized a mixed-integer programming solution implemented in OPSEL (Mullin, 2014; Mullin and Belotti, 2016). While this produces an optimum that is truly exact, the restriction to use each selected genotype equally was considered unnecessary, as we expected greater gain could certainly be achieved by optimizing the contribution of each selected genotype (Woolliams et al., 2015). Discussions with Skogforsk breeders confirmed that there was no operational reason to enforce equal crossing of each selection.

We thus looked to the “unequal contributions” solution within OPSEL, which was recently upgraded with a Second-Order Cone Programming (SOCP) solver that runs much faster than the former Semi-Definite Programming (SDP) approach (Mullin 2017a; Mullin et al., 2016; Yamashita et al., 2018). While the object function was originally intended to optimize unequal contributions to a seed orchard (Ahlinder et al., 2014; Mullin, 2014), we realized that, with some minor modifications, the SOCP solution could also be appropriate to optimize a specified total number of crosses.

The total number of contributions required is twice the number of desired crosses, as two contributions are required per cross. The SOCP solution gives the exact optimum proportion, which must then be multiplied by the twice the number crosses to obtain the express the solution as integer contributions. Some minor modifications to the integer conversion within OPSEL ensured that the contributions sum exactly to the required total. While the modified solution is, technically, no longer an “exact” integer solution, it will be very close and considered adequate for most operational applications. For our purpose here, we determined that 100 crosses per breeding population was a realistic level of effort. OPSEL was therefore asked to optimize a total of 2 × 100 = 200 contributions.

Initial optimization and fine-tuning of breeding population assignment

With optimization parameters determined, we were able to create an input file for OPSEL and perform an initial optimized selection. After this first optimization attempt, an investigation was performed of the performance in surrounding breeding populations of nominated founder trees. If a selected founder was ranked significantly higher in another breeding population and whose latitude of origin was acceptable in that other breeding population within established transfer rules, the founder tree was marked as unavailable for optimization and a second OPSEL run was performed with the updated candidate list.

Generation of crossing plan

Having generated an integer solution to optimize genotype contributions, it then remains to plan the field crossing among the selected parents, while avoiding excessive relatedness between mates that could result in inbreeding depression in the progeny. Numerous studies have demonstrated the effectiveness of positive assortative mating (PAM) for expanding genetic variance and creating opportunities to select superior production populations such as seed orchards (Lstibůrek et al., 2004b; Rosvall et al., 2003; 2015; Rosvall and Mullin, 2003;). We thus developed a new computer application, XDesign, to automate the generation of mating designs for the resulting selection lists, ensuring that positive assortative mating was applied as strictly as possible, while avoiding a user-specified level of coancestry between mates (Mullin, 2017b). XDesign is applied as a separate step, using the selection output file from OPSEL as its main input.

For positive-assortative mating, XDesign first prepares a ranked list of selected genotypes and their optimum contributions. The algorithm then starts at the top of the list, assigning selections as parents of potential crosses, working its way down the list while avoiding self and reciprocal crossing, until all contributions are exhausted, to complete a provisional list of crosses. If the mating is to be positive assortative, the ranking is based on decreasing EBV. When random assignment of mates is desired, XDesign simply applies a randomization procedure to shuffle the ranking.

Especially when the ranking of parents is positive assortative, we would expect that parents of some crosses in this provisional list may be closely related. Starting at the top of the provisional list, XDesign evaluates the coancestry between the parents of each cross. When this coancestry exceeds a user-specified value, XDesign searches down the list to find a mate whose coancestry is satisfactory and then “swaps” the two parents. The process is repeated until XDesign reaches the bottom of the list. It is possible that Designs will run out of potential swaps near the bottom of the list, but this occurs rather infrequently under positive assortative mating, as the number of contributions allocated to low-ranking genotypes tends to be quite small.

Comparison of selection options

In our case study populations, we had at least two generations of material that could be considered as selection candidates, the F0 founders, and various types of descendants including progeny of open-pollinated, poly-crossed and single-pair control-crossed families. In Swedish breeding, selection has not usually considered founders when producing crosses in advanced generations. While the founder genes are represented among their progeny, the solution to a selection optimization with a constraint on diversity might find greater average gain if some of the founders are used again in the next cycle of breeding.

We explored this possibility by comparing the results of strict “forward” selection, ignoring founders as potential breeding parents in the next round, as well as “forward-backward” selection, where genotypes across all generations with legitimate breeding values were considered as breeding candidates. The estimation of BLUP breeding values in TREEPLAN® allows us to generate breeding values that are comparable across generations.

Results and Discussion

IDENTIFYING SELECTION CANDIDATES

Even after the small number of population assignment changes following the initial optimizations, the three breeding populations represented rather similar numbers of founder plus-trees, each having about 2 850 individuals (Table 3). While similar in terms of numbers of founders, the genetic structure of the candidates from F1 and F2 generations varied greatly. The F1s in T4 and T7 had no paternal ID in the pedigree, as they arose from open-pollination or from polycrosses. In the case of T4, there were slightly more polycrosses (PX) than open-pollinated (OP), while the T7 population had over twice as many open-pollinated crosses. The constraint on group coancestry applied to optimal selection in these two populations was based on the female side of the pedigree; the unknown male parents were regarded as unique, non-inbred genotypes.

The advanced-generation candidates from T11 were very different in that the majority of male parentage was known. Only about 9 percent of the F1 candidates were progeny from OP, while the remainder of the F1s and all the F2 candidates were progeny from single-pair crosses. The T11 breeding population was the only one of the three that had F2 candidates, and in total had double the number of candidates available, compared with T4 and T7.

Table 3. Number of candidates available to OPSEL, after a pedigree check, showing the crossing of origin of candidates available for each breeding cycle in each breeding population as founders, open-pollinated (OP), polycross (PX) or controlled single-pair crosses (pair cross). The numbers of candidates shown are after the fine-tuning of breeding population assignment described in the text.

	Population		
Pedigree generation	**T4**	**T7**	**T11**
F0 (founders)	2 855	2 841	2 860
F1 OP	6 299	9 382	2 263
F1 PX	7 527	4 236	–
F1 pair cross	–	–	25 622
F2 pair cross	–	–	4 595
Total	**16 681**	**16 459**	**35 340**

Optimization under the various selection scenarios

OPSEL was able to complete the optimization of all selection scenarios for the three test populations without problem. There were some cases where the group coancestry constraint was violated to a very small degree, due to rounding of contributions to whole numbers summing to 200. The results are summarized in Table 4 and the various selection scenarios discussed below.

Backward selection scenario: Beginning with similar numbers of founder candidates, but with varying amounts and quality of progeny test data, the optimized "backward" selection utilized similar numbers of selected founder genotypes, between 30 and 33. The average estimated breeding value in each population varied considerably from 280 to 344. While breeding values from TREEPLAN are comparable, they are of course reported with reference to the target environment of the sub-population.

Forward selection scenario: Optimizing the 200 cross contributions by selecting only from F1 or F2 candidates ("forward" selection) always use contributions from more genotypes, up to twice as many as selection backward. For population T11, where the selection candidates were almost all progeny from pair crosses, the average breeding value of the selected group compared with the backward selection scenario was 19 percent higher, with over 90 percent of contributions coming from the pair-cross progeny. Selection in the T4 and T7 subpopulations produced over 60 percent of contributions from the OP progeny, but differences in average breeding value from backward selection were inconsistent, with forward selection in T4 producing 3 percent less, but in T7 produced 16 percent more.

Selection across generations: Considering candidates from across all generations gives a combined "forward-backward" scenario. The total number of genotypes in the optimal solution was intermediate between the smaller number founders selected in the backward selection scenario, and the larger number when selecting only forward. Contributions from F0 founder genotypes varied widely from 17 to 50 percent in populations T11 and T4, respectively. In the T4 and T7 populations with no pair-cross progeny, a large proportion of selections were made in both the OP and PX progeny candidates. For T11, almost all of the forward selections were progeny of single-pair crosses.

Of particular note, for ALL subpopulations, the average breeding value of the selected group was ALWAYS greater when optimal selection considered ALL identities in the pedigree, across ALL generations, as candidates. The improvement varied from 4 to 21 percent compared with backward selection, and from 1.5 to 7 percent compared to forward selection.

Table 4. Population statistics for alternative optimization scenarios using Founders, F1/F2s and Founders plus F1/F2 candidates, in Scots pine breeding populations Tpop T4, T7 and T11.

Tpop	Statistics	Founders,%	F1/F2, %	Founders plus F1/F2, %
T4	No selected founder genotypes	31	–	13
	No selected F1 genotypes	–	63	35
	Founder contribution	100	–	50.5
	F1 OP	–	62	23.0
	F1 PX	–	38	26.5
	F1 single-pair crosses	–	–	–
	F2 single-pair crosses	–	–	–
	Group coancestry	0.0250	0.0251	0.0253
	Average breeding value (BLUP)	344.4	334.0	357.7
T7	No selected founder genotypes	33	–	6
	No selected F1 genotypes	–	57	40
	Founder contribution	100	–	21
	F1 OP	–	61.5	47
	F1 PX	–	38.5	32
	F1 single-pair crosses	–	–	–
	F2 single-pair crosses	–	–	–
	Group coancestry	0.0253	0.0245	0.0268
	Average breeding value (BLUP)	292.7	340.3	353.1
T11	No selected founder genotypes	30	–	9
	No selected F1/F2 genotypes	–	54	41
	Founder contribution	100	–	17
	F1 OP	–	4	2,5
	F1 PX	–	3.5	1
	F1 single-pair crosses	–	62,5	52
	F2 single-pair crosses	–	30	27.5
	Group coancestry	0.0255	0.0249	0.0250
	Average breeding value (BLUP)	279.6	332.3	336.7

Implementation and supporting field work

After the final optimization run considering candidates from across all generations, a document was created including the selected genotypes and their exact position in archives and field trials. Field personnel were sent to confirm the status of each selected candidate. If a nominated tree was discovered to be dead or otherwise inappropriate, the field workers were able to contact the office immediately by phone so that the OPSEL candidate file could be revised accordingly (setting maximum contribution to zero) and optimization rerun to create and updated solution. If a new founder tree popped up during the revised optimization, a check of the performance in surrounding breeding populations was performed, as described earlier.

Due to the rapid execution of the new SOCP optimization algorithm, the field staff had the updated document in their hands within an hour. The flexible organization and implementation of the field work really highlights the benefit of a fast optimization tool.

Planning the crossing work

Once the selected candidates were confirmed as healthy and available, we could then plan the 100 crosses required to implement optimum contributions to the production of the next recruitment population, using XDesign. As these pine breeding programs in northern Sweden are in their early stages, it was possible to set a very strict control over coancestry between mates, with a threshold value of zero. XDesign was successful in all cases in generating a crossing plan with this coancestry constraint. An example of the positive-assortative crossing plan is given in (Table 5).

Steps to ensure successful implementation

The test sites where selected F1 and F2 trees were located are spread over a wide geographic region (Table 1). We decided, to facilitate operational crossing activities, by collecting scions from all selected material for production of grafts established into crossing archives at the Sävar breeding station.

It has been shown that grafting scions from Scots pine into the crowns of reproductively mature trees (inter-stocks) can shorten the time for production of both male and female strobili, compared with conventional grafting on young rootstocks (Almqvist, 2013). We therefore decided to create two different archives: a primary archive with top-grafted material; and a second long-term conservation archive with conventional grafts. The goal will be to perform short-term crossing activities in the top-grafting archive, and thereby hopefully reduce the generation turnover time. If the top grafts do not develop as expected, we have the option to shift the crossing activities to the conventional grafts.

Future directions

Even as we update the documentation of our new optimization algorithm in OPSEL (Mullin, 2017a), research into optimization methodology has continued on the operations research front. A fast solution has recently been developed using a relaxation adjustment to the SOCP optimization to deploy selected contributions equally (Safarina et al., 2017) and work is now underway to use a similar adjustment to better optimize unequal contributions. At the same time, work is also underway to consider a constraint on gender, while producing an even better adjustment for an integer solution. These improvements will expand the utility of OPSEL to consider a wider variety of selection constraints, while still providing a convenient and fast solution.

XDesign has recently been put to operational use, but already we have ideas for improvement of the tool. In its current form, XDesign assumes that all genotypes are monoecious and available to contribute as both pollen or seed parent. While many of the species we use for breeding in Sweden are monoecious, including pine, this doesn't mean that all individuals in the crossing archive will produce both male and female strobili. A revision has been proposed to XDesign to apply an additional constraint on each candidate regarding its use as male or female. This is a required modification before XDesign can be used for dioecious species, and will improve operational flexibility for all applications.

Table 5. The crossing plan for Pop T11. XDesign applied positive-assortative mating of the optimized genetic contributions, with a threshold limit of coancestry between mates of zero.

	Parent 2																																												
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	45	46	47	49	50
Parent 1																																													
1				1	2	3	4	5	6		7	8	9	10	11	12	13																												
2				14	15	16	17	18	19		20	21	22	23	44																														
3				24	25																																								
4				28	29	26	27	30																																					
6						31																																							
7			33				32		34																																				
9	36		37					35					38																																
11											39																																		
13	41	42										40																																	
14										43																																			
15		46								45																																			
16			49						47							48		50	51																										
17																52	53		54																										
18																	55	56		57	58																								
19																		59	60	61	64	62	63																						
21																				65																									
22																		66				67																							
23																								68	69																				
25																								70																					
26																										71	72	73	74	75															
28																											76		77																
29																												78																	
30																							79						80		81														
33																																82	83	84	85										
36																																				86	87	88							
38																																					89	90	91						
40																																							92						
41																																								93	94				
44																																										95	96	97		
45																																												98		
48																																													99	100

References

Ahlinder, J., Mullin, T.J., and Yamashita, M. 2014. Using semidefinite programming to optimize unequal deployment of genotypes to a clonal seed orchard. Tree Genet. Genom. 10: 27-34. doi: DOI 10.1007/s11295-013-0659-z.

Almqvist, C. 2013. Metoder för tidig blomning hos tall och gran. – Slutrapport av projekt 40:4 finansierat av Föreningen skogsträdsförädling. [Early strobili induction in Scots pine and Norway spruce. – Final report of Project no. 40:4, funded by the Swedish Tree Breeding Association.] Arbetsrapport från Skogforsk Nr 800-2013. Skogforsk.

Andersson, E.W., Sánchez Rodríguez, L., and Andersson, B. 1999. Group coancestry-controlled selection in a Pinus sylvestris L. breeding program. Theor. Appl. Genet. 99(1): 73-80. doi: 10.1007/s001220051210.

Andersson, E.W., Spanos, K.A., Mullin, T.J., and Lindgren, D. 1998a. Phenotypic selection can be better than selection for breeding value. Scand. J. For. Res. 13(1): 7-11.

Andersson, E.W., Spanos, K.A., Mullin, T.J., and Lindgren, D. 1998b. Phenotypic selection compared to restricted combined index selection for many generations. Silva Fenn. 32(2): 111-120.

Cockerham, C.C. 1967. Group inbreeding and coancestry. Genetics, 56: 89-104.

Danell, Ö. 1993. Breeding programmes in Sweden. 1. General approach. In Progeny testing and breeding strategies: Proceedings of the Nordic Group of Tree Breeders, Edinburgh, 6-10 October 1993. Edited by S.J. Lee. Forestry Authority, Scotland. pp. 80-94.

Danusevičius, D., and Lindgren, D. 2002a. Efficiency of selection based on phenotype, clone and progeny testing in long-term breeding. Silvae Genet. 51(1): 19-26.

Danusevičius, D., and Lindgren, D. 2002b. Two-stage selection strategies in tree breeding considering gain, diversity, time and cost. For. Genet. 9(2): 145-157.

Lindgren, D., Danusevičius, D., and 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. doi: doi:10.1139/X08-114.

Lindgren, D., Gea, L., and Jefferson, P. 1996. Loss of genetic diversity monitored by status number. Silvae Genet. 45(1): 52-59.

Lindgren, D., and Mullin, T.J. 1997. Balancing gain and relatedness in selection. Silvae Genet. 46(2-3): 124-129.

Lindgren, D., Wei, R.-P., and Bondesson, L. 1993. Optimal selection from families. Heredity, 70(6): 619-621.

Lstibůrek, M., Mullin, T.J., Lindgren, D., and 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., and 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., and Li, B. 2005. Positive assortative mating with family size as a function of predicted parental breeding values. Genetics, 171(3): 1311-1320. doi: 10.1534/genetics.105.041723.

Mullin, T.J. 2014. OPSEL 1.0: A Computer Program for Optimal Selection in Forest Tree Breeding. Arbetsrapport från Skogforsk Nr 814-2014.

Mullin, T.J. 2017a. OPSEL 2.0: a computer program for optimal selection in tree breeding. Arbetsrapport från Skogforsk 954-2017, Uppsala, SE.

Mullin, T.J. 2017b. XDesign 1.0: a tool for generation of complex mating designs with optimal contributions. Arbetsrapport från Skogforsk 956-2017, Uppsala, SE.

Mullin, T.J., and Belotti, P. 2016. Using branch-and-bound algorithms to optimize selection of a fixed-size breeding population under a relatedness constraint. Tree Genet. Genom. 12(4).

Mullin, T.J., Hallander, J., Rosvall, O., and Andersson, B. 2010. Using Simulation to Optimise Tree Breeding Programmes in Europe: An Introduction to POPSIM. Arbetsrapport från Skogforsk Nr 711, Uppsala, Sweden.

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. 44(2-3): 132-141.

Mullin, T.J., Yamashita, M., and Belotti, P. 2016. Optimising selection in tree breeding with constraints on relatedness and operational flexibility. In IUFRO Forest Genetics for Productivity Conference, 14-18 March 2016, Rotorua, New Zealand. p. 40.

Rosvall, O. 1999. Enhancing gain from long-term forest tree breeding while conserving genetic diversity. Swedish University of Agricultural Sciences, Umeå, Sweden. p. 65.

Rosvall, O., Almqvist, C., Lindgren, D., and Mullin, T. 2011. Breeding strategies. In Review of the Swedish Tree Breeding Programme. Edited by O. Rosvall. Skogforsk, Uppsala, Sweden. pp. 55-66.

Rosvall, O., and Andersson, E.W. 1999. Group-merit selection compared to conventional restricted selection for various trade-offs between genetic gain and diversity. For. Genet. 6(1): 1-14.

Rosvall, O., Kroon, J., and Mullin, T.J. 2015. Optimized breeding strategies at equivalent levels of population diversity. Arbetsrapport från Skogforsk Nr. 882-2015.

Rosvall, O., Lindgren, D., and 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., and 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., and Mullin, T.J. 2013. Introduction to breeding strategies and evaluation of alternatives. In Best Practice for Tree Breeding in Europe. Edited by T.J. Mullin and S.J. Lee. pp. 7-27.

Rosvall, O., Mullin, T.J., and 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.

Ruotsalainen, S., and Lindgren, D. 1998. Predicting genetic gain of backward and forward selection in forest tree breeding. Silvae Genet. 47(1): 42-50.

Ruotsalainen, S., and Lindgren, D. 2000. Stratified sublining: a new option for structuring breeding populations. Can. J. For. Res. 30(4): 596-604.

Ruotsalainen, S., and Persson, T. 2013. Scots pine -- Pinus sylvestris. In Best practices for Tree Breeding in Europe. Edited by T.J. Mullin and S.J. Lee. Skogforsk (The Forestry Research Institute of Sweden), Uppsala, SE. pp. 49-63.

Safarina, S., Belotti, P., Fukuda, M., Moriguchi, S., Mullin, T.J., and Yamashita, M. 2017. Conic relaxation approaches for equal deployment problems. Discrete Applied Mathematics(Special Issue: Advances in Optimization): Submitted.

Wei, R.-P. 1996. Loss of genetic diversity following selection from populations with a family structure. Silvae Genet. 45(2-3): 153-159.

Wei, R.-P., and Lindgren, D. 1991. Selection effects on diversity and genetic gain. Silva Fenn. 25(4): 229-234.

Wei, R.-P., and Lindgren, D. 1994. Gain and effective population size following index selection with variable weights. For. Genet. 1(3): 147-155.

Wei, R.-P., and Lindgren, D. 2001. Optimum breeding generation interval considering buildup of relatedness. Can. J. For. Res. 31(4): 722-729.

Wei, R.-P., Lindgren, D., and Yeh, F.C. 1997. Expected gain and status number following restricted individual and combined-index selection. Genetics, 40: 1-8.

Westin, J., and Haapanen, M. 2013. Norway spruce -- Picea abies (L.) Karst. In Best practices for Tree Breeding in Europe. Edited by T.J. Mullin and S.J. Lee. Skogforsk (The Forestry Research Institute of Sweden), Uppsala, SE. pp. 29-47.

Woolliams, J.A., Berg, P., Dagnachew, B.S., and Meuwissen, T.H.E. 2015. Genetic contributions and their optimization. J. Anim. Breed. Gen. 132(2): 89-99. doi: 10.1111/jbg.12148.

Yamashita, M., Mullin, T.J. and Safarina, S. 2018. An efficient second-order cone programming approach for optimal selection in tree breeding. Optimization Letters (in press).