and comparisons with seedling checks. For. Sci. **40**: 397–411 (1994). — Liang, K.-N., Zhou, W.-L., Zhong, C.-L. and Yang, Z.-J.: Eucalypt species/provenance trials at eastern Hainan Island. [Chinese]. Forest Research **7**: 399–407 (1994). — Liu, X., Yang, M., Wang, G., Pegg, R. E. and Arnold, R. J.: Development and genetic improvement of China's eucalypt resource. P. 92–93. In: M. J. Dieters, A. C. Matheson, D. G. Nikles, C. E. Harwood and S. M. Walker (eds.): Tree Improvement for Sustainable Tropical Forestry. Proc. QFRI-IUFRO conf. Caloundra, Queensland, Australia. 27 Oct. to 1 Nov. 1996 (1996). — Luz, H. F., Bierwagen, R., Ferreira, M. and Capitani, L. R.: Population studies and selection strategies of *Eucalyptus urophylla*. P. 204–205. In: M. J. Dieters, A. C. Matheson, D. G. Nikles, C. E. Harwood and S. M. Walker (eds.): Tree Improvement for Sustainable Tropical Forestry. Proc. QFRI-IUFRO Conf. Caloundra, Queensland, Australia. 27 Oct. to 1 Nov. 1996 (1996). — Mannion, E. G. and Wei, J.: An Introduction to the Dongmen Eucalypt Afforestation Project. Technical Communication No 40. The Dongmen State Forest Farm Eucalypt Afforestation Project, China (1989). — McCullagh, P. and Nelder, J. A.: Generalized Linear Models. London, Chapman and Hall (1989). — Mori, E. S., Kageyama, P. Y. and Ferreira, M.: Genetic variation and progeny x local interaction in *Eucalyptus urophylla*. IPEF International, Piracicaba **1**: 45–54 (1990). — Ngulube, M. R.: Provenance variation in *Eucalyptus urophylla* in Malawi. For. Ecol. Manag. **26**: 265–273 (1989). — Nirsatmanto, A., Seido, K., Kurinobu, S., Na'iem, M., Hardiyanto, E. B. and Suseno, H.: Analysis of provenance-progeny tests of *Eucalyptus urophylla* established at two locations in Indonesia. P. 206–207. In: M. J. Dieters, A. C. Matheson, D. G. Nikles, C. E. Harwood and S. M. Walker (eds.): Tree Improvement for Sustainable Tropical Forestry. Proc. QFRI-IUFRO Conf. Caloundra. Queensland, Australia. 27 Oct. to 1 Nov. 1996 (1996). — Pegg, R. E.: Selection of superior trees at Dongmen. Technical communication No 14. The Dongmen State Forest Farm Eucalypt Afforestation Project, China (1984). — Pegg, R. E., Gould, K. E. and Stevens, R. E.: Provenance trials at Dongmen. Proceedings of 3rd Technical Exchange Seminar. Nanning Guangxi, PRC (1987). — Pegg, R. E. and Li, H. J.: *Eucalyptus urophylla* potential at Dongman. In: Proceedings of 4th Technical Exchange Seminar. The Dongman State Forest Farm Eucalypt Afforestation Project, China (1989). — Qi, S.: Eucalypts in China. [Chinese]. Beijing, China Forestry Press (1989). —

Shelbourne, C. J. A.: Genotype-environment interaction: Its study and its implications in forest tree improvement. P. B-1(I), 1–28. In: IUFRO Genetics-SABRAO Joint Symposia. Tokyo, 1972 (1972). — Thurnbull, J. W.: Eucalypts in China. Aust. For. **44**: 222–234 (1981). — van Wyk, G.: Genetic improvement of timber yield and wood quality in *Eucalyptus grandis* (Hill.) Maiden. Part. I. Genetic parameters of growth characteristics. S. A. For. J. **153**: 1–11 (1990). — van Wyk, G. and Falkenhagen, E. R.: Genotype x Environment interaction in South African breeding material. P. 215–231. In: D. C. Grey, A. P. G. Schönau and C. J. Schutz (eds.): Site and Productivity of Fast-Growing Plantation, Proceedings of IUFRO Symposium. Pretoria, South Africa (1984). — Vercoe, T. and Clarke, B.: Trial growth performance of *Eucalyptus urophylla* S. T. Blake. Report to FAO Forestry Division on the growth *Eucalyptus urophylla* in international provenance trials and other growth trials. CSIRO Division of Forestry (1994). — Visscher, P. M., Thompson, R. and Hill, W. G.: Estimation of genetic and environmental variances for fat yield in individual herds and an investigation into heterogeneity of variance between herds. Livest. Prod. Sci. **28**: 273–290 (1991). — Volker, P. W., Dean, C. A., Tibbits, W. N. and Ravenwood, I. C.: Genetic parameters and gains expected from selection in *Eucalyptus globulus* in Tasmania. Silv. Genet. **39**: 18–21 (1990). — Wang, H.: A review of introduction, improvement, silviculture and utilization of eucalypts in China. P. 144–153. In: M.O.F. Scientific and Technological Information Centre, China (ed.). Development of Forestry Science and Technology in China, Beijing, China Science and Technology Press (1991). — Wang, H.-R.: Species and provenance trials of *Eucalyptus* in south east China. P. 357–359. In: G. I. Gibson, A. R. Griffin and A. C. Matheson (eds.). Breeding Tropical Trees: Population Structure and Genetic Improvement Strategies in Clonal and Seedling Forestry. [Proc. IUFRO Conference]. Pattaya, Thailand, November 1988 (1989). — Wei, X. and Borralho, N. M. G.: Use of individual tree mixed models to account for mortality and selective thinning when estimating base population genetic parameters. For. Sci. **44**: 246–253 (1998). — West, P. W. and Osler, G. H. R.: Growth response to thinning and its relation to site resources in *Eucalyptus regnans*. Can. J. For. Res. **25**: 69–80 (1995). — Zhou, W.-L. and Liang, K.-N.: *Eucalyptus urophylla* provenance trial. [Chinese]. Forest Research **4**: 172–177 (1991).

Maximising Genetic Response in Tree Breeding with Constraints on Group Coancestry

By R. J. Kerr[1][3], M. E. Goddard[2] and S. F. Jarvis[1]

Abstract

This study compares 4 methods to maximize response to selection while minimizing the inbreeding that accumulated over 6 generations of selection. A selection algorithm is described which maximizes the mean estimated breeding value of the selected trees while constraining the average coancestry among them and hence the status effective number of the population. This algorithm resulted in increased selection response, at a lower, target average coancestry level, compared to truncation selection. However the target average coancestry level for each generation is arbitrarily chosen. A variation of the algorithm is described which attempts to maximize response over 12 generations allowing for the decline in genetic variance caused by inbreeding. This leads to faster genetic gain and higher coancestry than typical tree breeding programs.

Once the selection has been carried out, a mate allocation algorithm arranges matings which give each parent the number of offspring required by the selection algorithm and which minimizes inbreeding in the offspring per mating. This delays the appearance of inbreeding by at least one generation but does not alter the long term rate of inbreeding. The combination of the selection and mating algorithms lead to breeding programs requiring less controlled pollinations than truncation selection with random mating and hence reduce the cost of the breeding program.

Introduction

Application of Cochran's optimal rule for selection (Cochran, 1951) in plant and animal improvement implies

[1] Southern Tree Breeding Association, P.O. Box 1811, Mount Gambier, S.A. 5290, Australia

[2] Animal Genetics and Breeding Unit, University of New England, Armidale, N.S.W. 2351, Australia

[3] Present address: Animal Genetics and Breeding Unit, University of New England, Armidale, N.S.W. 2351, Australia

parents are selected soley on the criterion of estimated breeding value (EBV). This rule is optimal for a single cycle of selection and will maximize genetic advance in the generation following selection. However repeated use of this rule will compromise long term genetic gain. This is because the accumulation of inbreeding in the population will reduce future genetic variance. If long term genetic gain is to be considered, what is the optimal rule for selection?

A growing area of research (recently summarised by WEI, 1995) is aimed at providing such an answer. A common technique is to use LAGRANGE multipliers to find an optimal set of genetic contributions from a set of available candidates. Many authors, applying this technique, formulate specific selection rules. WRAY and GODDARD (1994) provide a selection rule which optimally balanced genetic gain and inbreeding over a given time horizon. MEUWISSEN (1997) provides a selection rule which maximises genetic gain while constraining inbreeding to a predefined value. Implied in these rules is the assumption that any increase in coancestry equals a future increase in inbreeding. However these rules were formulated for use in animal breeding where hierarchical mating is the norm. In a series of tree breeding papers (LINDGREN and MATHESON, 1986; LINDGREN, 1991; LINDGREN et al., 1993) the selection rule maximized genetic gain at a given selected proportion and genetic diversity. This work is only relevant to a single cycle of selection and does not balance gain and diversity over many generations. More recently, LINDGREN and MULLIN (1997) have presented an iterative exchange algorithm to select a fixed number of individuals which will maximize genetic gain, given an imposed weighting on average coancestry. The method was demonstrated using computer simulated populations in which selection continued for 5 discrete generations. ZHENG et al. (1997) extended the method such that the number of selections could vary.

The techniques described above, while attempting to define genetic contributions from either a fixed or varying number of parents, do not specify to which individuals other individuals should be mated, or in other words, the mate allocations. Other techniques exist which are more suited to this problem. These are: minimum coancestry mating (TORO et al., 1988) which aims to avoid matings between close relatives; compensatory mating (GRUNDY et al., 1994) which aims to mate parents which are already well represented in other matings with parents which are poorly represented; and assortative mating (BAKER, 1973) which aims to mate best with best.

The aims of this study are twofold. Firstly, algorithms are presented which enable the breeder to make a more rational choice of the appropriate penalty factor in weighting average coancestry against expected genetic gain. The algorithms are similar to those presented by WRAY and GODDARD (1994) and MEUWISSEN (1997). Methods which maximize total response over 12 generations or increase coancestry by a predefined amount each generation are compared with truncation selection. Secondly, a technique based on iterative exchange and improvement is explained which can translate the set of genetic contributions into an optimum mating arrangement. The use of the algorithms to define genetic contributions and the technique to convert them to mate allocations is illustrated using computer simulation.

Method

Selection is defined by a 1 x n_c contribution vector c where n_c is the number of all available candidates prior to selection. A contribution is valued between 0 and 1 and expresses the proportion of the next generation's gene pool that is contribut-

ed by an individual candidate. The sum of contributions is one ($\sum_i c_i = 1$) and a contribution of zero means the candidate has not been selected.

Allocation is the process of assigning individuals to particular matings in accordance with c. These two processes have been separated deliberately, though they can be considered jointly. If considered jointly the problem increases in complexity and it is doubtful whether treated together achieves any greater benefit than if each process was treated separately.

Calculation of the contribution vector

Two algorithms were considered. The first is based on the approach of WRAY and GODDARD (1994) who aimed to maximise cumulative net response over a fixed time period. Their objective function which enabled this to occur will be the same used in this study:

$$R_t/t \approx \Delta G_L - Q\Delta F \qquad (1)$$

where R_t is the cumulative net response at generation t, ΔF is the rate of inbreeding per generation, and ΔG_L is the asymptotic rate of gain per generation expected in an infinite population. An asymptotic rate is used because the reduction in between family additive genetic variance through selection disequilibrium (BULMER, 1971) can be accounted for. Finally,

$$Q = \Delta G_L(t-1)/2 + D \qquad (2)$$

where D is the depression in performance per unit of inbreeding and t is the number of generations in the future at which net response is maximized. The scalar Q can be regarded as a cost or penalty factor for the increase in inbreeding resulting from selection. The objective function (1) is now replaced by a selection criterion (S_t) at any generation t within the defined time period. The selection criterion is written in terms of the coancestry of selection candidates and their contributions to the next generation. The following selection criterion is applicable to tree breeding. Firstly ΔG_L is substituted by the level of additive genetic merit in the next generation, ΔG_{t+1}, expected from selection decisions made at generation t. This is given by

$$\Delta G_{t+1} = c_t' b_t \qquad (3)$$

where b_t contains the estimated breeding values, deviated from the mean, of all available candidates at generation t. An approximation for ΔF arises by considering the following equation (FALCONER and MACKAY, 1996):

$$\Delta F = \frac{F_{t+1} - F_t}{1 - F_t}. \qquad (4)$$

The mean inbreeding in progeny of any generation will depend on the amount of common ancestry in the parents. If F_t is small, and hence $1-F_t$ can be ignored, ΔF could be approximated by $\bar{C}_{t+1} - \bar{C}_t$, where $\bar{C}$ represents average coancestry in the parents. This approximation is dependent on the assumption that the mate allocations only produce small fluctuations in the value for ΔF. Because $\bar{C}_t$ is fixed by the time selection of parents for generation $t+1$ is made, it can also be ignored. Thus $\Delta F \approx \bar{C}_{t+1}$. Because a numerator relationship coefficient between two individuals is twice the coancestry between them, and assuming discrete generations,

$$\bar{C}_{t+1} = c_t' A_t c_t/2, \qquad (5)$$

where A_t is a $n_c \times n_c$ matrix containing the numerator relationships amongst all candidates at generation t and is a subset of the numerator relationship matrix for the entire population. $\bar{C}_{t+1}$ defined like this is the same as average coancestry defined by COCKERHAM (1967) and used by LINDGREN et al. (1996), except the average of all pair-wize coancestries is a weighted average. Following LINDGREN and KANG (1998) group coancestry is prefered over average coancestry, since the latter can often imply the exclusion of self-coancestry. A_t does include the relationship of an individual with itself. Thus the selection criterion, S_t, is given by

$$S_t = c'_t b_t - Q(c'_t A_t c_t)/2 \qquad (6)$$

Maximising S_t is achieved by imposing the constraint

$$c'_t 1 = 1$$

and the LAGRANGEan function

$$S_t = c'_t b_t - Q(c'_t A_t c_t)/2 - (c'_t 1 - 1)\lambda$$

where 1 is a vector of ones of same magnitude as c_t. By differentiating S_t with respect to c_t and setting to zero yields:

$$b_t - QAc_t - 1\lambda = 0.$$

Rearranging terms to obtain a solution for c yields:

$$c_t = \frac{A^{-1}}{Q}(b_t - 1\lambda). \qquad (7)$$

Multiplying both sides of the equation by 1', we get

$$1 = \frac{1}{Q} 1' A^{-1} b_t - \frac{1}{Q} 1' A^{-1} 1 \lambda$$

and solving for λ yields:

$$\lambda = \frac{1' A^{-1} b_t - Q}{1' A^{-1} 1}. \qquad (8)$$

The second algorithm is based on the algorithm of MEUWISSEN (1997) who maximised the level of additive genetic merit in the following generation, ΔG_{t+1}, subject to a constraint on the increase in group coancestry which approximates the future level of inbreeding and is given by equation (5). Thus the selection criterion S is simply

$$S_t = c'_t b_t, \qquad (9)$$

subject to

$$c'_t A_t c_t / 2 = \bar{C}_t + \Delta F \approx \bar{C}_{t+1} \qquad (10)$$

and

$$c'_t 1 = 1$$

LINDGREN et al. (1997) define status effective number (N_s) as half the inverse of group coancestry and is equivalent to the actual census number of a set of non-inbred, unrelated genotypes. Thus N_s will tend to decline with generations as inbreeding and coancestry rise. LINDGREN and MULLIN (1997) suggest status effective number is a more easily understood concept than group coancestry and the concept also circumvents the ambiguity concerning what average or group coancestry refers to. Genetic merit could be maximized subject to a constraint on the status effective number desired for each future generation. This could be acheived by setting the value

of $c'_t A c_t / 2$ equal to $0.5/N_{st}$ where N_{st} is the desired status effective number in generation t.

To find the optimal c_t to maximise S_t a different LAGRANGEan function is used:

$$S_t = c'_t b_t - (c'_t 1 - 1)\lambda_1 - (c'_t A_t c_t - 2\bar{C}_{t+1})\lambda_2$$

Equating the first derivative of S_t with respect to c_t to zero and solving for c_t yields:

$$c_t = \frac{A^{-1}(b_t - 1\lambda_1)}{2\lambda_2} \qquad (11)$$

Multiplying both sides of the equation by 1' yields an equation similar to equation (8) for λ_1:

$$\lambda_1 = \frac{1' A^{-1} b_t - 2\lambda_2}{1' A^{-1} 1}. \qquad (12)$$

Using the constraint $c'_t A c_t = 2\bar{C}_{t+1}$ it follows that

$$(13)$$

$$\lambda_2^2 = \frac{b'_t \left[A^{-1} - A^{-1} 1 (1' A^{-1} 1)^{-1} 1' A^{-1} \right] b_t}{8\bar{C}_{t+1} - 4(1' A^{-1} 1)^{-1}}.$$

A value for λ_2 is obtained prior to calculating λ_1.

The following algorithm was used for both methods to obtain the vector c_t which maximised the selection criterion S_t.

1. Select the best $n = n_c$, that is, select the complete set of available candidates.

2. Compute necessary lambdas and use in either 7 or 11 to compute c_t.

3. Rank the candidates on their c_i values. If all elements of c_t are positive then stop the search and the current c_t is optimal for maximising S_t, otherwise remove the individual with the most negative c_i value and repeat from step 2 with $n = n - 1$.

MEUWISSEN (1997) has given proof that this algorithm yields the optimal c_t within the solution space $c_t \geq 0$.

Allocation

The following procedure is then used to translate the genetic contribution vector, c_t, into an allocation matrix, X_t, which defines the mating program at generation t. The first step is to divide each c_i by c_{min}, the minimum contribution, and round up to the nearest integer which yields the number of contribution units (ζ_i). As an example *table 2* lists 6 selections, with their corresponding genetic contributions and contribution unit numbers. *Table 3* lists additive genetic relationships (a_{ij}) between all trees.

Table 1. – Phenotypic variances, heritabilities (on diagonal), genetic correlation (above diagonal) and phenotypic correlation (below diagonal) used for diameter at breast height (DBH) and pilodyn penetration score (PIL). These are the traits simulated in the study.

Trait	Phenotypic variance	Correlations DBH	PIL
DBH	6.1	.25	.27
PIL	2.8	.11	.60

Table 2. – Example of 6 selected individuals with their genetic contributions and contribution unit numbers.

Individual	c_i	ζ_i
1	.25	5
2	.25	5
3	.20	4
4	.15	3
5	.10	2
6	.05	1

Table 3. – Matrix containing additive genetic relationships between individuals listed in table 2.

1.	0	.25	.25	0	.25
	1.	0	0	0	.25
		1.	.25	.25	0
Symmetric			1.	.5	0
				1.	0
					1

There are many possible strategies in making the allocations. In this study it was decided to minimise the group coancestry between mates (i.e. the inbreeding of the progeny) and to minimise the variance of family size, subject to achieving the prescribed genetic contributions from each parent. The problem reduces to finding values for $n_s \times n_s$ variables x_{ij} which indicate the number of contribution units from each mating. Where $x_{ij} = 0$ implied the individuals were not mated. A suitable objective function to be minimised is as follows:

$$\sum_i \sum_j a_{ij} x_{ij} + \frac{\sum_i \sum_j x_{ij}^2 - \left(\sum_i \sum_j x_{ij}\right)^2/n_s}{n_s - 1}, \quad (14)$$

subject to

$$\sum_j x_{ij} + \sum_j x_{ji} = \zeta_i$$

for all i and where $i \neq j$. To obtain a solution a heuristic algorithm, based on interchange and iterative improvement, was used. This approach is sometimes referred to as "simulated annealing", due to the method first appearing in statistical mechanics (KIRKPATRICK et al., 1983). The problem is represented by

$$\min\{v(\xi): \xi \in \Xi\} \quad (15)$$

where ξ is an arrangement or configuration with sample space Ξ and value $v(\xi)$, determined from (14). Denoting the current arrangement as ξ^i, the algorithm proceeds by generating a rearranged configuration ξ'. In this study random numbers, uniformly distributed in the interval (0,1), were used to rearrange a fraction of the individuals into new matings. If the rearranged configuration improves the value of the objective function it is kept, that is

if $v(\xi') < v(\xi^i)$, then $\xi^{i+1} = \xi'$.

On the other hand if the value is greater, $v(\xi') > v(\xi^i)$ the rearranged configuration becomes the new configuration with propability ρ, where

$$\rho = \exp\left\{\frac{v(\xi') - v(\xi^i)}{\tau_j}\right\}.$$

where τ is a control parameter for the jth round of iteration. This step is a feature of the METROPOLIS algorithm (METROPOLIS et al., 1953), which incorporates controlled uphill steps in the search for better solutions. The motivation in accepting a configuration with smaller objective value is that it avoids getting stuck in a local but not a global optimum. The process proceeds like this for a set number of iterations. When a prerequisite number of successes, $v(\xi') < v(\xi^i)$, or the iteration limit has been reached, a new round of iteration begins ($j \rightarrow j+1$). The parameter τ_j is decreased at the rate $\tau_{j+1} = \tau_j(1-\alpha)$, with α usually in the vicinity (0.01, 0.1). In statistical mechanics τ_j is referred to as the "annealing temperature". The system is first optimised at a high effective temperature and slowly the temperature is lowered until the system "freezes" and no further changes occur. A starting value for τ is chosen such that the value is larger than the largest expected value for $|v(\xi') - v(\xi^i)|$ (PRESS et al., 1989). In this study α was 0.01. It may be noted with this method that an exact solution is never assured. However as the size of the optimisation problem increases, an exact solution becomes increasingly irrelevant, and the performance of the algorithm dominates the analysis of practical applications.

An initial configuration was obtained by the following algorithm.

1. Initialise X by setting to zero.
2. Rank selections according to genetic contributions.
3. For $i = 1$ to $(n_s - 1)$ do

$$m = \sum_j^{i-1} x_{ji}$$

    $k_i = \zeta_i - m$
    $j = i + 1$
    while k_i greater than zero
        $k_j = \zeta_j$
        if k_i greater than or equal to k_j
            $x_{ij} = k_j$
            $k_i = k_i - k_j$
        else
            $x_{ij} = k_i$
            $k_i = 0$
        end if
        $j = j + 1$
    end while
end do.

In order for the last parent to have the correct ζ_i value, some approximation may be required in allocating an extra unit to the penultimate parent. Applying this algorithm to the individuals listed in *table 2* resulted in an allocation matrix shown in *table 4*. Rows of the allocation matrix are assigned to female parents and columns to male parents. In the annealing algorithm a new configuration is obtained by completing a series of shuffles. One shuffle is achieved by randomly choosing a row, say i. Within this row a filled (i, j) and unfilled cell (i, k) are randomly chosen. Then within the kth column a filled cell is randomly chosen (l, k). The shuffle is completed when a contribution unit is subtracted from both (i, j) and (l, k) and added to (i, k) and (l, j). The optimum configuration as deter-

168

mined by the annealing algorithm is shown in *table 5*. The total weighted coancestry was 0.5 initially and the initial variance of x_{ij} values was 3.67. The optimum allocation had a total weighted coancestry of 0.375 and a variance of x_{ij} values of 0.285. If the objective function was to minimize total pairwize weighted coancestry only ($\Sigma_i \Sigma_j a_{ij} x_{ij}$) this would result in fewer matings each producing many non or less inbred individuals. In order to reduce the variance of family size it is more optimal to produce more matings each producing fewer, but similar numbers of non or less inbred individuals. *Table 6* shows the optimum configuration under the sole objective of reducing $\Sigma_i \Sigma_j a_{ij} x_{ij}$. This configuration also had a total weighted coancestry of 0.375 but a variance of x_{ij} values of 2.0.

Table 4. – Allocation matrix, **X**, prior to optimisation.

	1	2	3	4	5	6
1	0	5	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	3	1	0
4	0	0	0	0	0	0
5	0	0	0	0	0	1
6	0	0	0	0	0	0

Table 5. – Allocation matrix, *X*, following optimisation.

	1	2	3	4	5	6
1	0	2	0	2	1	0
2	0	0	0	0	0	0
3	0	2	0	1	0	1
4	0	0	0	0	0	0
5	0	1	0	0	0	1
6	0	0	0	0	0	0

Table 6. – Allocation matrix, *X*, following optimisation with the sole objective of reducing total weighted pairwize coancestry.

	1	2	3	4	5	6
1	0	1	0	3	1	0
2	0	0	0	0	0	0
3	0	4	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	1
6	0	0	0	0	0	0

The final step is to obtain actual numbers of progeny to be collected from each mating. It is assumed that the limiting resource for the breeding program is the number of individuals that can be field tested, denoted *N*. If the number of progeny per contribution unit is given by $N/\Sigma_i \Sigma_j x_{ij}$ then the actual numbers of progeny from each mating are obtained by multiplying each x_{ij} by this amount.

Simulation

Populations were simulated with discrete generations and a multi-trait breeding objective was considered. A founder population of 200 individuals was considered sufficient for a typical forest tree breeding population (BURDON et al., 1977). Additive

genetic values or breeding values ($g(i)_j$, stated as the breeding value of the *i*th individual for the *j*th trait) for 2 traits were sampled from a multivariate normal distribution $MN(\mu, G)$, where μ is a 1 x 2 vector of zeros and G is a 2 x 2 positive definite additive genetic variance covariance matrix, with elements $G_{jk} = COV(g_j, g_k)_{t=0}$. For descendents $G_{jk} = 1/2(1-\bar{F}_{ds})$ $COV(g_j, g_k)_{t=0}$ where $\bar{F}_{ds}$ is the average inbreeding coefficient of the female parent, *d* and the male parent, *s*, and $\mu = [\mu_1, \mu_2]$ where $\mu_1 = 1/2g(d)_1 + 1/2g(s)_1$, which is the mid-parental breeding value for the 1st trait and likewise for the 2nd trait. A record for each trait for an individual was simulated as the sum of the breeding value, an environmental effect common to all individuals located in the same site and an individual environmental effect. Both environmental effects were sampled from multivariate normal distributions, with mean vectors containing zeros and variance covariance matrices with elements $E_{jk} = COV(e_j, e_k)$.

Variances and covariances used were current estimates from the Southern Tree Breeding Association's (STBA) *E. globulus* data base and are shown in *table 1*. Selection criteria are currently diameter at breast height (DBH) and pilodyn penetration score (PIL), while traits in the breeding objective are volume, density and pulp yield. Weightings for selection criteria were derived using the economic weightings currently used in the breeding objective.

In the founder generation the 200 parents were each mated 4 times and 15 progeny were generated from each mating. Estimates of breeding value (EBVs) were computed using a multivariate BLUP. The 6000 progeny were ranked on EBV and the number of possible candidates was restricted to the top 800 individuals. One of 5 options was then used to select the parents. The options were:

Option 1

Selection on a variable number of parents with a genetic contribution vector calculated using the WRAY and GODDARD (1994) algorithm. A breeding horizon of 12 generations was used. The depression in performance per unit of inbreeding, *D*, was assumed to be zero. Simulated annealing was used to assign matings.

Option 2

Selection on a variable number of parents with a genetic contribution vector calculated using the MEUWISSEN (1997) algorithm. Simulated annealing was used to assign matings. The group coancestry of the parents of generation *t* was limited to $(t-1)\Delta F$ for $t = 2, \dots, 6$ where $\Delta F = 0.017$ giving an effective population size of approximately 30.

Option 3

Same as for Option 2 except that the group coancestry of the parents of generation *t* was limited to $0.5/N_{st-1}$ for $t = 2, \dots, 6$. The following values for N_{st} were used: $N_{s1} = 150$; $N_{s2} = 100$; $N_{s3} = 50$; $N_{s4} = N_{s5} = 25$. The aim of this option is to show that the decline in status effective number can be fixed. To reduce computing requirements it was decided to restrict the number of individuals selected per family to 2. Thus fewer individuals are needed to achieve the desired status effective number and operations needed to determine the contribution vector (see equations 11, 12 and 13) are far less time consuming.

Unrestricted truncation selection

The top 200 ranked individuals were selected on EBV. Each parent was then mated 4 times with no restrictions on matings between close relatives. Fifteen progeny were obtained from each mating.

Restricted truncation selection

The number of individuals selected per family was restricted to 2. The number of parents remained at 200, with each parent mated 4 times and 15 progeny collected from each mating.

For options 1 to 3 simulated annealing, as described above, was used to assign matings and determine the number of progeny from each mating. The total number of progeny each generation was exactly 6000 for truncation selection and on average 6000 for other options.

For option 1 a value for Q is needed prior to calculating the contribution vector. Equation (2) was redefined as

$$Q_j = \Delta G_j (t - j - 1)/2, \qquad (16)$$

where Q_j is calculated at each generation j over the t generations defined as the breeding horizon, and where ΔG_j is the expected, asymptotic rate of gain per generation, based on information at generation j. This formulation for Q_j implies that there is no interest in selection response after generation t. A value for ΔG_j was predicted using a selection index developed by KERR (1998) that approximates a multivariate BLUP individual tree model. This index depends on predictions for selection intensity, accuracy of the selection index and the genetic variance, pertaining to individuals in generation j. The value arrived at for Q can only be approximate for several

reasons. When the number of selections can vary the selection intensity can only be deduced once the selection criterion (6) has been maximized and the vector c_j has been defined. However maximizing (6) requires a value for Q, which in turn needs a prediction for selection intensity. In this study a prediction for selection intensity was obtained on the basis of the previous generation's selection decisions. WRAY and GODDARD (1994) noted the stability of selection response in their simulations over quite differing values used for Q.

Only 6 generations of response was simulated. This represents approximately 50 years in Eucalypt breeding programmes and is considered sufficient for a long term breeding program. Specifying a breeding horizon of 12 generations in option 1 ensures there is sufficient genetic variance for continuing selection beyond the actual simulated horizon.

Results

Table 7 shows the mean true breeding value (TBV) for the objective of reducing cost per tonne of pulp, at generations 2 to 6 for each option. *Table 7* also shows the mean inbreeding in the progeny and the group coancestry between parents of the progeny at each generation. Each value is the mean over 50 replicates and standard errors are shown in parentheses. *Figure 1* shows status numbers by generation for each scheme. *Table 9* shows the average number of parents selected each generation, the mean number of matings made, the mean number of matings per individual and the mean number of progeny per mating for options 1 to 3. Under truncation selection the number of parents selected is held constant at 200, and each individual is mated 4 times which amounts to 400 matings per generation. The number of progeny per mating is also fixed at 15. All means were computed over 50 replicates.

Table 7. – True breeding value (TBV) in progeny for the objective of reducing cost per tonne of pulp, mean inbreeding of the progeny (F_t) and group coancestry of parents ($\bar{C}_t$) at each generation for: option 1 – maximizing selection response given a variable target group coancestry; option 2 – maximizing response given a fixed target group coancestry; option 3 – maximizing response given a fixed target group coancestry which is related to status number; unrestricted truncation selection and restricted truncation selection. Values are averaged over 100 replicates (standard errors in parentheses).

Option	Generation (t)	TBV	$\bar{C}_t$	F_t
Option 1 (Variable $\bar{C}_t$)	2	20.6 (.4)	.038 (.001)	.001 (.000)
	3	35.1 (.5)	.067 (.002)	.014 (.002)
	4	50.1 (.6)	.100 (.003)	.052 (.003)
	5	64.3 (.7)	.137 (.003)	.092 (.003)
	6	78.6 (.8)	.175 (.004)	.132 (.004)
Option 2 (Fixed $\bar{C}_t$)	2	19.4 (.3)	.017 (.000)	.000 (.000)
	3	33.9 (.3)	.033 (.000)	.008 (.000)
	4	48.0 (.3)	.050 (.000)	.013 (.001)
	5	62.2 (.3)	.067 (.000)	.029 (.001)
	6	76.1 (.3)	.083 (.000)	.047 (.001)
Option 3 (Fixed $\bar{C}_t$)	2	9.0 (.1)	.003 (.000)	.002 (.000)
	3	19.6 (.1)	.005 (.000)	.004 (.000)
	4	32.9 (.2)	.010 (.000)	.007 (.000)
	5	47.1 (.2)	.020 (.000)	.013 (.000)
	6	56.2 (.2)	.020 (.000)	.017 (.000)
Unrestricted truncation selection	2	17.3 (.2)	.020 (.001)	.021 (.001)
	3	31.0 (.3)	.051 (.002)	.050 (.002)
	4	43.9 (.3)	.076 (.004)	.075 (.004)
	5	56.7 (.3)	.102 (.006)	.102 (.006)
	6	69.6 (.4)	.131 (.007)	.131 (.007)
Restricted truncation selection	2	14.1 (.2)	.007 (.000)	.006 (.000)
	3	26.6 (.2)	.014 (.000)	.014 (.000)
	4	38.5 (.3)	.022 (.001)	.022 (.001)
	5	50.1 (.4)	.030 (.001)	.029 (.001)
	6	61.6 (.4)	.038 (.001)	.038 (.001)

Table 8. – Changes in mean inbreeding between generations under option 2 with and without mate allocation.

$F_t - F_{t-1}$	**With mate allocation**	**Without mate allocation**
$F_3 - F_2$	.008	.015
$F_4 - F_3$	.005	.015
$F_5 - F_4$	.016	.016
$F_6 - F_5$	.018	.018

The first comparison to be noted is between unrestricted truncation selection and option 1 in which selection decisions are made according to their effect on the optimal rate of inbreeding. Option 1 resulted in a 13% to 19% relative advantage in aggregate genetic merit over the generations simulated. Several additional advantages are noted. Firstly, option 1 requires fewer numbers of parents selected and fewer controlled pollinations. On average the requirements for option 1 are: 35 parents and between 100 and 160 controlled pollinations per generation (see *Table 9*) compared with 200 parents and 400 controlled pollinations under truncation selection. This implies a significant financial cost saving. However, option 1 also results in an increase in the group coancestry of the breeding population and the greatest reduction in status effective number of all options (see *Figure 1*). It should also be noted that the increased coancestry does not translate to increased inbreeding. In fact, mean inbreeding in the progeny is less under option 1, relative to unrestricted truncation selection, in all but the last generation. It must be remembered that under this option it is assumed that selection stops in generation 12 and a decreasing Q in the latter generations results in a less stringent constraint on the rate of inbreeding. Finally, standard errors for mean TBV are higher under option 1 (see

Generation (*t*)	No. parents selected	Total no. matings	No. matings per parent	No. progeny per mating
Option 1				
1	36	160	8.6	42
2	37	159	8.4	42
3	35	135	7.7	47
4	33	125	7.6	51
5	27	95	7.0	73
Option 2				
1	92	427	9.3	14
2	75	364	9.8	17
3	69	306	9.0	20
4	66	310	9.4	20
5	66	293	8.9	21
Option 3				
1	534	491	1.8	12
2	399	493	2.5	12
3	202	487	4.8	12
4	119	457	7.7	13
5	220	485	4.4	12

[1]) Mean of 100 replicated simulations of the option.

Table 7). Thus option 1 carries with it more risk as measured by variance of response.

Option 2 provides a more risk averse alternative. Under this option selection decisions are made according to their effect on maintaining a predefined limit on group coancestry. From *table 7* group coancestry in the parental population increases in increments of 0.017 which was the defined rate. The advantage in genetic gain (9% to 12%), relative to unrestricted truncation selection was less than that under option 1 but was achieved while simultaneously reducing the group coancestry compared to unrestricted truncation selection. Option 2 without the mate allocation step was also tested. Instead of using simulated annealing, the male and female parents were individually selected in mating pairs with probabilities defined by their genetic contributions. The gains in aggregate genetic merit were the same with and without optimal mate allocations but different levels of inbreeding were observed. Values for $F_t - F_{t-1}$

Figure 1. – Decline in status effective number from 5 generations of selection for each of the options in making selection decisions.

are shown in *table 8* with and without optimal mate allocations. These results tend to confirm the finding that minimizing pairwize coancestry only serves to delay the onset of inbreeding by at least one generation (Toro et al., 1988). It also confirms that $\bar{C}_t - \bar{C}_{t-1}$ provides a reasonable approximation to ΔF. Perhaps the main role of the mate allocation step, in addition to delaying inbreeding, is in providing a convenient mechanism for translating genetic contributions into a mating program.

All options presented so far result in a dramatic decrease in status effective number in the first generation (see *Figure 1*). Option 3 was designed such that status effective number fell from 200 to 150 in the first generation of selection, then fell in increments of 50 and 25 until a status effective number of 25 was reached. For generations 4 and 5 the status effective number remained at 25. From *table 7* it can be seen this predefined decrease in status effective number is achieved with a much reduced increase in aggregate genetic merit relative to other options. The option also requires a much larger census number. From *table 9* it can be seen that on average over 500 parents are required in generation 1 (an increase of 300 from the base population of 200), 400 in generation 2 and 200 in generations 3 and 5. In addition, close to 500 controlled pollinations are required each generation which is an increase of 25% relative to truncation selection. To ease computational burden option 3 was implemented restricting the number of full sibs selected per family to 2. Truncation selection in which a maximum of 2 full sibs were able to be selected per family was also tested. Restricted truncation selection although resulting in more genetic gain relative to option 3, could not guarantee a similar decline in status effective number (see *Figure 1*). In the first generation the status effective number fell to 71, which was more than options 1 an 2 and unrestricted truncation selection, but was significantly less than 150 under option 3. Of all options, option 3 offers the least risk as measured by the magnitude of the standard errors for response in aggregate merit and inbreeding. It should be noted that the same status effective number can only be maintained for perhaps one or
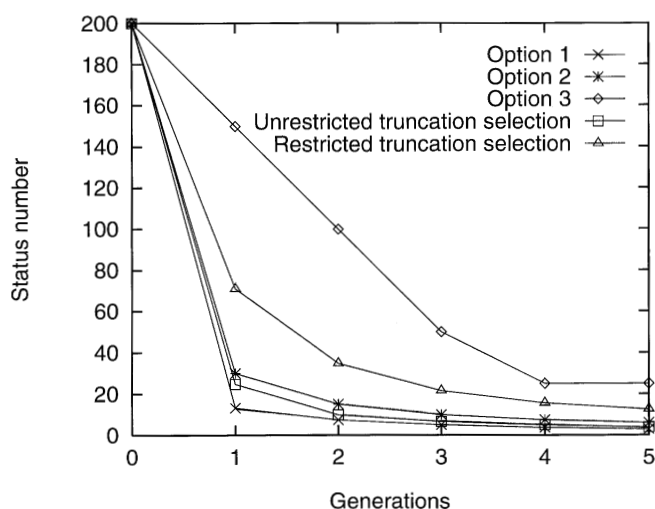
two generations of a breeding program. Coancestry eventually has to increase.

Discussion

A major decision in forest tree breeding is how large to make the breeding population. Given that a founder population of at least 200 unrelated genotypes is desirable (BURDON et al., 1977) how many selected parents should there be in advanced generations? Increased intensity of selection will yield greater short term genetic gain but reduce effective population size more rapidly, thus eroding the potential for long term genetic gain through reduced genetic variability and increased rates of inbreeding.

Beyond knowing the size of the selected group there are questions regarding the number of families represented and the number of individuals selected from any family. When an optimal index selection method such as BLUP is used and if selecting for a lowly heritable trait it can be demonstrated that a few families will dominate the rankings (JARVIS et al., 1995). This is called selective advantage (ROBERTSON, 1961) and will increase with the degree of use of family information in the index. Though a few families will dominate the rankings, the problem is how to prevent the few families from dominating the selections. A restriction on the numbers of individuals selected per family will result in more families and lower ranked individuals being selected. However, the level of this restriction needs to be ascertained.

Finally once the selected group has been decided, how are the selections to be mated and should each mating leave equal numbers of progeny? A case against unequal progeny could be made in that having some families with many progeny will increase the probability of these sibships dominating the rankings in the next selection cycle. Again it is a case of preventing them from dominating the selections.

This study has presented a methodology which effectively answers all the above questions in a non empiric manner. The methodology assumes the only resource limitation is the capacity of the breeding organisation to field test progeny. Within the methodology various options were presented which differ in the manner by which group coancestry in the breeding population is weighted against additive genetic gain. In the first option presented (option 1) group coancestry is weighted by a factor Q. Q is determined by the effect selection decisions have on future genetic variance and, if applicable, on inbreeding depression. Option 1 results in what seems at first startling low numbers of select parents: from 200 unrelated founder parents to approximately 36 parents in the 1st generation and 27 in the 6th generation. However, the study revealed that the judicious choice of these parents, their genetic contributions to the next generation, and the individuals to which they are mated, resulted in substantially more genetic gain than is achieved when leaving the number of selected parents at 200 and practising truncation selection. However, there is a cost attached to this option. Group coancestry in the breeding population is increased substantially. Minimizing pairwize coancestry in the mate allocation step can delay inbreeding over the initial generations but the increased coancestry will ultimately mean higher rates of inbreeding when the time horizon is approached (12 generations in this study). The tendency for the rate of inbreeding to increase as the time horizon is approached could be eliminated by always defining the time horizon to 12 generations into the future.

In forest tree breeding, breeding populations are generally not part of the production populations. Thus it may be permissable to allow high rates of inbreeding in the breeding population. It is also assumed under option 1 that there is no inherent interest in the breeding population after the breeding horizon is completed. It is conceivable that more than one breeding population could be established: a smaller population bred solely for the purposes of achieving the maximum genetic gain over a specific horizon; and a larger population which is bred for management of genetic resources. Alternatively, two separate breeding populations or sublines could be established, both which are bred for maximum productivity. Individuals from different sublines can then be crossed to produce deployment families which are non-inbred.

If genetic gain and genetic diversity are to be considered jointly in the one breeding population, then maximizing gain while maintaining a defined rate of increase in group coancestry would be preferable. In option 2 the defined rate of increase in group coancestry was related to change in mean inbreeding between generations, while in option 3 it was related to change in status effective number. LINDGREN et al. (1997) point out that the rate of change of inbreeding between generations, and hence effective population size, is a limited concept for conveying genetic diversity information at the current generation. The concept of status effective number is better for this purpose. Thus options 2 and 3 are essentially the same. The only difference is the degree to which group coancestry changes between generations. Does it change in accordance with a fixed rate of change in inbreeding; or does it change in accordance with a fixed rate of change in status effective number?

The essential difference between options 2 and 1 is whether the rate of inbreeding in the population is optimised or fixed. Apart from this they are essentially the same algorithm. At the same group coancestry they will yield the same genetic contributions. Once the vector of contributions is obtained both options use the same algorithm to determine the mating combinations. The algorithms used to calculate the contribution vector are similar to that described by WRAY and GODDARD (1994) and MEUWISSEN (1997) but redefined to take into account that forest trees are monoecious. These authors have defended their respective algorithms against earlier work, and a similar defense is perhaps not needed in this study. WRAY and GODDARD generally noted the arbitrary or empiric manner in which earlier algorithms calculated cost factors in penalising expected gain against expected coancestry. MEUWISSEN argued that no previous work presented a selection rule which restricted group coancestry to a predefined value.

The annealing algorithm used to determine the optimum mate combinations offers great flexibility. It was assumed in the simulation that an individual could either be male or female parent. However there may be restrictions in that certain individuals cannot provide pollen or female flowers. This could be taken into account by preventing certain cells in the allocation matrix from having non zero values. In the simulation individuals were involved in mutiple-pair matings under no fixed design. Certain designs could be implemented by restricting the annealing to a half-diallel design, for example.

It should be noted that the objective function for determining mate allocations was somewhat arbitrary. Initially we chose minimum coancestry matings as the sole objective because of the advantages in delaying inbreeding in the first stages of a breeding program; namely not having phenotypic values confounded with inbreeding depression and having all test plants with the same level of inbreeding (viz no inbreeding). However because of the tendency for the algorithm to produce fewer families with many progeny each, an additional term for minimizing variance of family size was included. Also the decision

to translate a genetic contribution to a contribution unit integer number seems reasonable since it would be very difficult to match each individual's genetic contribution, expressed as a fraction, exactly with the required number of progeny. It must be stressed that the avoidance of matings between close relatives only delays the first increment of inbreeding and any reduction in the subsequent rate of inbreeding is negligible. WANG (1998) has recently shown that excluding sib matings, under equal family size selection, can in fact result in higher inbreeding rates in the long term, compared with random mating. The main goal in the present study has been to demonstrate that simulated annealing provides a convenient mechanism for translating genetic contributions into a mating configuration. More research is needed in defining the most appropriate objective function for the optimal mating configuration.

The options presented, particularly options 2 and 3, can also be applied to selecting clones for a deployment seed orchard. LINDGREN and MATHESON (1986) described an algorithm which selected n out of m candidate clones and calculated the contribution of each clone selected to the total seed collected. If, for example, production managers could recommend a minimum value for status effective number then option 3 could be used to calculate the number of clones to be deployed and the contribution of each clone. As MEUWISSEN (1997) has pointed out prediction error variances (PEV) could be used in lieu of additive genetic relationships to constrain the variance of response to a predefined value.

An understated aspect of this study has been the use of BLUP in the prediction of breeding values. WEI (1995) suggests that often phenotypic selection is preferable to BLUP selection in forest tree breeding. The arguments are as follows. Phenotypic selection does not use information from relatives, hence the probability of co-selection of relatives is less. Phenotypic selection is also less complex than BLUP and does not require knowledge of male and female parent. Thus there is less risk of incurring mistakes and less cost and time involved in its execution. However we would argue that BLUP is necessary to correctly account for fixed effects. Use of BLUP can provide estimates of genetic trend which are at least close to unbiased and which have smaller sampling variances than estimates based on least squares estimates of differences from control populations. An estimate of genetic trend is important for breeders to question the effectiveness of their breeding program. With the exponential growth in personal computing capability and the availability of public domain BLUP software it is doubtful whether BLUP selection does incur more time and cost.

In summary, the study has presented a methodology which offers a mechanism to unify all facets of the breeding program. Values for parameters such as: breeding population size; genetic contributions of each individual; mating combinations; and the number of progeny tested per mating are all determined within a two step methodology. It remains for the breeder to decide if genetic gain is to be maximized subject to: a rate of inbreeding which is optimal for a specific time horizon; or a rate of inbreeding which is fixed; or alternatively a rate of decline in status effective number.

References

BAKER, R. J.: Assortative mating and artificial selection. Heredity **31**, 231–238 (1973). — BULMER, M. G.: The effect of selection on genetic variability. American Naturalist **105**, 201–211 (1971). — BURDON, R. D., SHELBOURNE, C. J. A. and WILCOX, M. D.: Advanced selection strategies. In: Proc. 3rd World Consultation on Forest Tree Breeding. Canberra, Australia. pp. 1137–1147 (1977). — COCHRAN, W. G.: Improvement by means of selection. In: Proc. 2nd Berkeley Sym. Math. Stat. and Prob.. pp. 449–470 (1951). — COCKERHAM, C. C.: Group inbreeding and coancestry. Genetics **56**, 89–104 (1967). — FALCONER, D. S. and MACKAY, T. F. C.: Introduction to Quantitative Genetics. (4th edition). Longman, Essex (1996). — GRUNDY, B. A., CABALLERO, A., SANTIAGO, E. and HILL, W. G.: A note on using biased parameter values and non-random mating to reduce rates of inbreeding in selection programmes. Anim. Prod. **59**, 465–468 (1994). — HENDERSON, C. R.: A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics **32**, 69–83 (1976). — JARVIS, S. F., BORRALHO, N. M. G. and POTTS, B. M.: Implementation of a multivariate BLUP model for genetic evaluation of *Eucalyptus globulus* in Australia. In: Eucalypt plantations: improving fibre yield and quality. POTTS, B. M., BORRALHO, N. M. G., REID, J. B., CROMER, R. N., TIBBITS, W. N. and RAYMOND, C. A. (*eds.*): Proc. CRC-IUFRO Conf., Hobart, 19 to 24 Feb. CRC for Temperate Hardwood Forestry, Hobart (Australia). pp. 141–145 (1995). — KERR, R. J.: Asymptotic rates of response from forest tree breeding strategies using best linear unbiased prediction. Theor. Appl. Genet. **96**, 484–493 (1998). — KIRKPATRICK, S., GELATT, C. D. and VECCHI, M. P.: Optimization by simulated annealing. Science **220**, 671–680 (1983). — LINDGREN, D.: Optimal utilization of genetic resources. For. Tree Improv. **23**, 51–67 (1991). — LINDGREN, D., GEA, L. D. and JEFFERSON, P. A.: Status number for measuring genetic diversity. Forest Genet. **4**, 69–76 (1997). — LINDGREN, D. and KANG, K. S.: Status number – a useful tool for tree breeding. Research Report of the Forest Genetics Research Institute, Korea (1998). — LINDGREN, D. and MATHESON, C. A.: An algorithm for increasing the genetic quality of seed from seed orchards by using the better clones in higher proportions. Silvae Genet. **35**, 173–177 (1986). — LINDGREN, D. and MULLIN, T. J.: Balancing gain and relatedness in selection. Silvae Genet. **46**, 124–129 (1997). — LINDGREN, D., WEI, R.-P. and BONDESSON, L.: Optimal selection from families. Heredity **70**, 619–621 (1993). — METROPOLIS, N., ROSENBLUTH, A., ROSENBLUTH, M., TELLER, A. and TELLER, E.: Equations of state calculations by fast computing methods. J. Chem. Phys. **21**, 1087–1092 (1953). — MEUWISSEN, T. H. E.: Maximizing the response of selection with a predefined rate of inbreeding. J. Anim. Sci. **75**, 934–940 (1997). — PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. and VETTERLING, W. T.: In Numerical Recipes, the Art of Scientific Computing. Fortran Edition. Cambridge University Press (1989). — ROBERTSON, A.: Inbreeding in artificial selection programmes. Genet. Res. **2**, 189–194 (1961). — TORO, M., NIETO, B. and SALGADO, C.: A note on minimization of inbreeding in small scale selection programmes. Livest. Prod. Sci. **20**, 317–323 (1988). — WEI, R.-P.: Predicting Genetic Diversity and Optimizing Selection in Breeding Programmes. PhD Dissertation, Swedish University of Agricultural Sciences (1995). — WRAY, N. R. and GODDARD, M. G.: Increasing long term response to selection. Genet. Sel. Evol. **26**, 431–451 (1994). — ZHENG, Y. Q., LINDGREN, D., ROSVALL, O. and WESTIN, J.: Combining genetic gain and diversity by considering average coancestry in clonal selection of norway spruce. Theor. Appl. Genet. **95**, 1312–1319 (1997).