MODELLING OF SELECTION AND MATING DECISIONS IN TREE BREEDING PROGRAMS

Hardwood trees from the temperate forests of southern Australia are an important source of timber for high quality paper. Two species in particular, *Eucalyptus globulus* and *Eucalyptus nitens* are well suited to this purpose and are now widely grown in commercial plantations. These plantations have been established by professional tree breeders using seedlings derived originally from broadly based collection of seed in natural forests. To increase productivity it is desirable to select trees that grow quickly and give high yields of top quality timber. Nevertheless it is important to maintain genetic diversity in the breeding population and thereby retain a robust capacity to adapt to changing environmental factors. In this article we formulate a number of related mathematical models for the selection and mating processes and discuss the consequences of these models. We recommend a relatively simple scheme which can be implemented on an IBM compatible PC using standard algorithms.

1. Introduction

Tree breeding aims to maximise the rate of genetic progress with each generation. To achieve this goal, tree breeders have to take a number of key decisions:

- which of the available trees should be selected for use in the breeding program;
- to what extent should each selected tree contribute to the next generation; and
- which mate pairs should be used.

After a brief review of relevant literature (Brisbane and Gibson, 1995; Jansen and Wilton, 1985; Meuwissen and Woolliams, 1994; and Sedgley and Griffin, 1989) and discussion with representatives from the Cooperative Research Centre in Temperate Hardwood Forestry, the Southern Tree Breeding Association and Northern Forest Products, the MISG team suggested three alternative models. *The basic tree selection model* was developed in detail and a corresponding selection scheme has already been implemented using real data. This scheme will

- select the number of times each tree should be mated in order to maximise the total breeding value of the selected trees subject to an appropriate penalty for their collective pairwise relatedness.

Two other models were developed in principle. *The entropy model* makes novel use of a standard entropy function to measure the genetic diversity of the breeding population and a corresponding scheme would

- select the number of times each tree should be mated in order to maximise the entropy of the selected trees while maintaining an acceptable total breeding value.

The third model is more complicated but may be useful in the longer term to predict the improvement in breeding values over a number of generations. In recognizing that the breeding value for the progeny of a particular mating is a random variable with expected value equal to the average breeding value of the parents the MISG team realised that if a large number of progeny are produced from each particular cross and only the best progeny are selected then the expected breeding value of the selected progeny will be somewhat higher than the average breeding value of the parents. *The enhanced tree selection model* incorporates these ideas and a corresponding scheme would

- calculate the number of progeny to be produced and the minimum acceptable breeding value for each cross in order to maximise the total expected breeding value for the selected progeny subject to the incorporation of sufficient genetic diversity in the progeny.

2. Measuring the genetic content of a population

We suppose the current population $\mathcal{P}_t$ at time t is descended from an original population $\mathcal{P}_0$ of N individuals denoted by $\alpha_1, \alpha_2, \ldots, \alpha_N$. Thus we write

$$\mathcal{P}_0 = \{\alpha_1, \alpha_2, \ldots, \alpha_N\}. \tag{1}$$

Each individual α is represented by a *pedigree* composed of an ordered pair of pedigrees of the male and female parents respectively. Thus when $\alpha \in \mathcal{P}_t$ and $\beta \in \mathcal{P}_t$ we denote the offspring by $\alpha \times \beta \in \mathcal{P}_t$ and write

$$p(\alpha \times \beta) = (p(\alpha), p(\beta)) \tag{2}$$

for the corresponding pedigree. For the original population $\mathcal{P}_0$ we write

$$p(\alpha_j) = j. \tag{3}$$

The *genetic content* $\boldsymbol{\nu}(\gamma) = (\nu_j(\gamma)) \in \Re^N$ of each individual $\gamma \in \mathcal{P}_t$ is a measure of the relative contribution of each member of the original population to the pedigree of that individual and is calculated from the genetic content of the parents by the formula

$$\boldsymbol{\nu}(\alpha \times \beta) = \frac{1}{2}\boldsymbol{\nu}(\alpha) + \frac{1}{2}\boldsymbol{\nu}(\beta). \tag{4}$$

For the original population $\mathcal{P}_0$ we assume that

$$\boldsymbol{\nu}(\alpha_1) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \boldsymbol{\nu}(\alpha_2) = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \boldsymbol{\nu}(\alpha_N) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \tag{5}$$

which means we effectively assume that the original population is composed of unrelated individuals. The *genetic measure* $\boldsymbol{\nu}(\mathcal{P}_t) = (\nu_j(\mathcal{P}_t)) \in \Re^N$ of each original ancestor in the current population $\mathcal{P}_t$ is defined by

$$\boldsymbol{\nu}(\mathcal{P}_t) = \frac{1}{M_t} \sum_{\alpha \in \mathcal{P}_t} \boldsymbol{\nu}(\alpha) \tag{6}$$

where $\boldsymbol{\nu}(\alpha) \in \Re^N$ is the genetic content of the individual α and where $M_t = M(\mathcal{P}_t)$ is the number of individuals in the current population. It is easy to see that

$$\sum_{j=1}^{N} \nu_j(\alpha) = 1 \tag{7}$$

for all $\alpha \in \mathcal{P}_t$ and hence

$$\sum_{j=1}^{N} \nu_j(\mathcal{P}_t) = 1. \tag{8}$$

The *degree of relatedness* of two individuals is measured by their common ancestry and is defined by

$$a(\alpha, \beta) = \sum_{j=1}^{N} \min(\nu_j(\alpha), \nu_j(\beta)). \tag{9}$$

If the genetic contents of α and β are equal then clearly $a(\alpha, \beta) = 1$. Otherwise $0 \leq a(\alpha, \beta) < 1$. For our discussions we are assuming that the individuals in the original population are completely unrelated with $a_{ij} = a(\alpha_i, \alpha_j) = 0$ for $j \neq i$.

The measurement scheme described above is illustrated by the simple example in Section 8.

3. The basic tree selection model

We wish to select m trees from a population with M individuals in order to maximise the genetic merit of the selection subject to an appropriate limit on the relatedness of the m selected trees. We suppose that the population $\mathcal{P} = \mathcal{P}_t$ is denoted by $\beta_1, \ldots, \beta_M$ and that we use m_i clones of the individual β_i for each $i = 1, 2, \ldots, M$. Thus we require

$$\sum_{i=1}^{M} m_i = m. \tag{10}$$

We assume that the estimated breeding value of β_i is given by v_i and define an objective function

$$\begin{aligned} V &= \sum_{i=1}^{M} v_i m_i - k \sum_{i=1}^{M} (\sum_{j=1}^{M} a_{ij} m_j)^2 \\ &= \mathbf{v}^T \mathbf{m} - k \|A\mathbf{m}\|^2 \end{aligned} \tag{11}$$

where $\mathbf{v} = (v_i) \in \Re^M$, $\mathbf{m} = (m_j) \in \Re^M$, $k \in \Re$ with $k > 0$ is a known constant and the symmetric matrix $A = (a_{ij}) \in \Re^{M \times M}$ is the *relatedness matrix* defined by

$$a_{ij} = a(\beta_i, \beta_j). \tag{12}$$

The penalty term expresses the total pairwise relatedness of the selected individuals. It is convenient to define the *associated relatedness matrix* $B = (b_{ij}) \in \Re^{M \times M}$ by $B = A^T A$ in which case we can write

$$\begin{aligned} V &= \sum_{i=1}^{M} v_i m_i - k \sum_{i=1}^{M} \sum_{j=1}^{M} m_i b_{ij} m_j \\ &= \mathbf{v}^T \mathbf{m} - k \mathbf{m}^T B \mathbf{m}. \end{aligned} \tag{13}$$

If we define

$$x_i = \frac{m_i}{m} \tag{14}$$

then we can rewrite the objective function as

$$\begin{aligned} V &= V(\mathbf{x}) \\ &= m \sum_{i=1}^{M} v_i x_i - k m^2 \sum_{i=1}^{M} \sum_{j=1}^{M} x_i b_{ij} x_j \\ &= m \mathbf{v}^T \mathbf{x} - k m^2 \mathbf{x}^T B \mathbf{x} \end{aligned} \tag{15}$$

where $\mathbf{x} = (x_i) \in \Re^M$ and we seek to maximise $V(\mathbf{x})$ on a region $\mathcal{F} \subseteq \Re^M$ defined by the constraints

$$\sum_{i=1}^{M} x_i = 1 \tag{16}$$

and $x_i \geq 0$ for each $i = 1, 2, \ldots, M$. If we regard the variables x_i as continuous variables then the problem can be solved using Lagrange multipliers. Because $V(\mathbf{x})$ is concave and because $\mathbf{x}$ is restricted to a convex subset $\mathcal{F} \subseteq \Re^M$ any local maximum is also a global maximum. Because $V(\mathbf{x})$ is continuous and $\mathcal{F}$ is compact there must be at least one point where the global maximum is achieved. We define a Lagrangean function

$$\begin{aligned} \mathcal{V} &= \mathcal{V}(\mathbf{x}) \\ &= m\left[\sum_{i=1}^{M} v_i x_i - km\sum_{i=1}^{M}\sum_{j=1}^{M} x_i b_{ij} x_j + \lambda(1 - \sum_{i=1}^{M} x_i) + \sum_{i=1}^{M} \pi_i x_i\right] \\ &= m\left[\mathbf{v}^T\mathbf{x} - km\mathbf{x}^T B\mathbf{x} + \lambda(1 - \mathbf{1}^T\mathbf{x}) + \boldsymbol{\pi}^T\mathbf{x}\right] \end{aligned} \tag{17}$$

where $\mathbf{1} = (1) \in \Re^M$ and where $\lambda \in \Re$, $\boldsymbol{\pi} = (\pi_i) \in \Re^M$ are Lagrange multipliers with $\lambda \geq 0$, $\boldsymbol{\pi} \geq \mathbf{0}$. By applying the Kuhn-Tucker equations

$$\frac{\partial \mathcal{V}}{\partial x_i} = 0 \tag{18}$$

for each $i = 1, 2, \ldots, M$ and the complementary slackness conditions we obtain necessary conditions for a local maximum. We have

$$v_i - 2km\sum_{j=1}^{M} b_{ij} x_j - \lambda + \pi_i = 0 \tag{19}$$

for each $i = 1, 2 \ldots, M$ and

$$\lambda(1 - \sum_{i=1}^{M} x_i) + \sum_{i=1}^{M} \pi_i x_i = 0. \tag{20}$$

In vector form we can write

$$\mathbf{v} - 2kmB\mathbf{x} - \lambda\mathbf{1} + \boldsymbol{\pi} = \mathbf{0} \tag{21}$$

and

$$\lambda(1 - \mathbf{1}^T\mathbf{x}) + \boldsymbol{\pi}^T\mathbf{x} = 0. \tag{22}$$

Let $\mathcal{S} \subseteq \mathcal{F}$ be the solution set. If we assume that $\mathbf{x}_a$, $\mathbf{x}_b$ are distinct solutions with $V(\mathbf{x}_a) = V(\mathbf{x}_b) = V_{max}$ then we have

$$\mathbf{v} - 2kmB\mathbf{x}_a - \lambda_a\mathbf{1} + \boldsymbol{\pi}_a = \mathbf{0} \tag{23}$$

and

$$\mathbf{v} - 2kmB\mathbf{x}_b - \lambda_b \mathbf{1} + \boldsymbol{\pi}_b = \mathbf{0} \tag{24}$$

and some elementary manipulations allow us to deduce that

$$2km(\mathbf{x}_a - \mathbf{x}_b)^T B(\mathbf{x}_a - \mathbf{x}_b) + \boldsymbol{\pi}_a{}^T \mathbf{x}_b + \boldsymbol{\pi}_b{}^T \mathbf{x}_a = 0. \tag{25}$$

Since each term in this equation is non-negative it follows that each term is zero. Now we can see that

$$\begin{aligned} V(\theta\mathbf{x}_a + [1-\theta]\mathbf{x}_b) &= \mathbf{v}^T(\theta\mathbf{x}_a + [1-\theta]\mathbf{x}_b) \\ &\qquad -km(\theta\mathbf{x}_a + [1-\theta]\mathbf{x}_b)^T B(\theta\mathbf{x}_a + [1-\theta]\mathbf{x}_b) \\ &= \theta V(\mathbf{x}_a) + (1-\theta)V(\mathbf{x}_b) + km(\mathbf{x}_a - \mathbf{x}_b)^T B(\mathbf{x}_a - \mathbf{x}_b) \\ &= V_{max} \end{aligned} \tag{26}$$

for each $\theta \in [0,1]$ and hence each point between $\mathbf{x}_a$ and $\mathbf{x}_b$ is also a solution. Therefore the solution set $\mathcal{S}$ is convex. To find $\mathcal{S}$ it is necessary to solve equations (21) and (22) but we can only do this if we first nominate which variables x_i will be zero at the solution point. These variables are then deleted from the problem and we reformulate a corresponding *reduced order* problem in the same form as the original. For convenience we will use the same notation for the reduced order problem. We can now assume that all variables x_i are non-zero at the solution point. We decompose each vector into orthogonal components in the null space $\mathcal{N} = \mathcal{N}(A)$ and the range space $\mathcal{R} = \mathcal{R}(A^T)$ where $A \in \Re^{L \times L}$ is now the reduced order relatedness matrix. For each $\mathbf{u} \in \Re^L$ we write $\mathbf{u} = \mathbf{u}_n + \mathbf{u}_r$. Because $\mathbf{x} > \mathbf{0}$ it follows that $\boldsymbol{\pi} = \mathbf{0}$ and the equation

$$\mathbf{v} - 2kmB\mathbf{x} - \lambda \mathbf{1} = \mathbf{0} \tag{27}$$

can be rewritten as

$$(\mathbf{v}_n + \mathbf{v}_r) - 2kmA^T A(\mathbf{x}_n + \mathbf{x}_r) - \lambda(\mathbf{1}_n + \mathbf{1}_r) = \mathbf{0} \tag{28}$$

from which we deduce that

$$\mathbf{v}_n - \lambda \mathbf{1}_n = \mathbf{0} \tag{29}$$

and

$$\mathbf{v}_r - 2kmA^T A\mathbf{x}_r - \lambda \mathbf{1}_r = \mathbf{0}. \tag{30}$$

If equation (29) is satisfied we can calculate a unique solution for $\mathbf{x}_r$ in equation (30). For every feasible $\mathbf{x} = \mathbf{x}_n + \mathbf{x}_r \in \mathcal{F}$ we can use equation (30) to deduce that

$$\lambda = \mathbf{v}_r{}^T \mathbf{x}_r - 2km\|A\mathbf{x}_r\|^2 \tag{31}$$

and hence

$$\begin{aligned} V(\mathbf{x}) &= \mathbf{v}^T\mathbf{x} - km\|A\mathbf{x}\|^2 \\ &= \mathbf{v}_n{}^T\mathbf{x}_n + \mathbf{v}_r{}^T\mathbf{x}_r - km\|A\mathbf{x}_r\|^2 \\ &= km\|A\mathbf{x}_r\|^2 + \lambda(\mathbf{1}_n{}^T\mathbf{x}_n + \mathbf{1}_r{}^T\mathbf{x}_r) \\ &= km\|A\mathbf{x}_r\|^2 + \lambda \end{aligned} \tag{32}$$

and the solution set is given by

$$\mathcal{S} = \{\mathbf{x}|\mathbf{x} = \mathbf{x}_n + \mathbf{x}_r \text{ where } \mathbf{x}_n \in \mathcal{N}\} \cap \mathcal{F}. \tag{33}$$

If this set is empty then it means we have nominated the wrong variables to take zero values and we must begin the solution process again with a different set of zero variables. If $\mathcal{S}$ is non-empty it is still possible that we may have set too many variables equal to zero and may not have found the complete solution set. In general we need to find feasible solutions which minimise the number of zero variables.

The solution scheme is illustrated by the simple example in Section 9.

The choice of the constant k in the penalty term will influence the solution. We can consider the problem from a different point of view. Suppose that the matrix B is positive definite. Then equation (27) can be solved to give

$$\begin{aligned} \mathbf{x} &= \mathbf{x}_k \\ &= \frac{1}{2km}B^{-1}(\mathbf{v} - \lambda_k\mathbf{1}). \end{aligned} \tag{34}$$

Since condition (16) can be rewritten in the form

$$\mathbf{1}^T\mathbf{x}_k = 1 \tag{35}$$

we now have

$$\frac{1}{2km}\mathbf{1}^TB^{-1}(\mathbf{v} - \lambda_k\mathbf{1}) = 1. \tag{36}$$

From equation (27) we deduce that

$$\mathbf{v}^T\mathbf{x}_k - 2km\mathbf{x}_k{}^TB\mathbf{x}_k - \lambda_k = 0 \tag{37}$$

and if we define

$$\bar{v}_k = \mathbf{v}^T\mathbf{x}_k \tag{38}$$

and choose k so that

$$\mathbf{x}_k{}^TB\mathbf{x}_k = r \tag{39}$$

where r is the acceptable level of relatedness and is essentially the *risk factor* then equation (37) shows that

$$\lambda_k = \bar{v}_k - 2kmr. \tag{40}$$

If equation (34) is used to rewrite equation (39) in the form

$$\frac{1}{4k^2m^2}(\mathbf{v}-\lambda_k\mathbf{1})^T B^{-1}(\mathbf{v}-\lambda_k\mathbf{1}) = r \tag{41}$$

then equations (36) and (41) can be solved by a suitable iterative scheme to determine k and λ_k.

Note that in the case where the population consists of a number of unrelated families the matrix B takes a block diagonal form and hence B^{-1} also has this same structure. It follows from the above formulae that the determination of $\mathbf{x}$ can be made separately for each family.

4. A practical solution algorithm for the basic tree selection model

If we define the *relatedness vector* $\mathbf{g} = (g_i) \in \Re^M$ by setting

$$\begin{aligned} \mathbf{g} &= B\mathbf{m} \\ &= mB\mathbf{x} \end{aligned} \tag{42}$$

then

$$g_i = \sum_{j=1}^{M} b_{ij}m_j \tag{43}$$

is the *relatedness coefficient* for the tree β_i. We can rewrite the objective function in the form

$$\begin{aligned} V &= \sum_{i=1}^{M}(v_i - k\sum_{j=1}^{M} b_{ij}m_j)m_i \\ &= \sum_{i=1}^{M}(v_i - kg_i)m_i \\ &= (\mathbf{v}-k\mathbf{g})^T\mathbf{m} \end{aligned} \tag{44}$$

and suggest a possible elementary procedure for calculation of the maximum value of V subject to the constraint (10) and such that each component of $\mathbf{m} = (m_i) \in \Re^M$ is a non-negative integer. Essentially we start from the most highly valued trees and consider the vector

$$\mathbf{w} = \mathbf{v} - k\mathbf{g} \tag{45}$$

of *modified breeding values* where the modification is a penalty for the degree of relatedness in the selected trees. For convenience we will assume that $v_1 \geq v_2 \geq \ldots \geq v_M$. We calculate

$$V = \mathbf{w}^T\mathbf{m} \tag{46}$$

and apply some scheme of iterative improvement. The most rudimentary scheme would simply replace one currently selected tree with a tree that is not currently selected. To this end we take $p < q$ and suppose that m_p is reduced by 1 and m_q is increased by 1. We have

$$\mathbf{m}^T = (m_1, m_2, \ldots, m_p, \ldots, m_q, \ldots, m_M) \tag{47}$$

and

$$(\mathbf{m} + \Delta\mathbf{m})^T = (m_1, m_2, \ldots, m_p - 1, \ldots, m_q + 1, \ldots, m_M) \tag{48}$$

and hence

$$\begin{aligned} \Delta V &= [\mathbf{v} - kB(\mathbf{m} + \Delta\mathbf{m})]^T(\mathbf{m} + \Delta\mathbf{m}) - [\mathbf{v} - kB\mathbf{m}]^T\mathbf{m} \\ &= (v_q - v_p) - 2k(g_q - g_p) - k(b_{qq} + b_{pp} - 2b_{pq}) \\ &= (v_q - v_p) - 2k[(g_q - g_p) + (1 - b_{pq})] \end{aligned} \tag{49}$$

where g_q and g_p are the relatedness coefficients for trees β_q and β_p before the change. Since we assume that $v_q < v_p$ then we see that $\Delta V > 0$ only if tree β_q is less closely related to the other trees than is tree β_p.

Two simulated annealing algorithms which we call *SingleTree* and *MultipleTree* have been written to perform the required optimisation and have been tested on real data.

In *SingleTree* we allow only $m_i = 0$ or $m_i = 1$ but the heart of both programs is the same: initially we choose $\mathbf{m}$ so that condition (10) is satisfied. In other respects the non-negative integer components of $\mathbf{m}$ are arbitrarily chosen. For each pair of suitable components p and q we investigate decreasing m_p by 1 and increasing m_q by 1. This is restricted to p such that $m_p > 0$ and, in the case of *SingleTree*, to q such that $m_q = 0$. The change ΔV in V is computed by equation (49), and the change is made if $\Delta V > 0$. On the other hand, if $\Delta V \leq 0$, the change is made with probability given by

$$p(\Delta V, s) = \frac{e^{-(\Delta V)^2}}{\sqrt{sC + 1}} \tag{50}$$

where s is the number of steps already taken for which $\Delta V \leq 0$ and C is initially set to 0.02, for fairly slow *cooling*. There is nothing magical about this cooling schedule; the main thing is that $p(\Delta V, s)$ has the desirable features of going to 0 rather slowly as s increases and rather quickly as $|\Delta V|$ increases.

A sample set of test data was supplied containing 730 trees with some information on their pedigrees that permitted calculation of a nominal modified relatedness matrix $B = (b_{ij})$ in equation (12). *SingleTree* and *MultipleTree* were tested in two sample situations, both of which appear to find the true optimum easily.

SingleTree was used with $m = 20$ and $M = 730$, and with $k = 0.1$. Starting with two permutations of the input data, firstly unsorted and secondly sorted by $v_i > v_{i+1}$ *SingleTree* gave the same solution after only about 200,000 iterations, taking less than a minute and with no further improvement after an overnight run.

In *MultipleTree* we allow m_i to take any non-negative integer values. This program was tested with $m = 60$ and $M = 730$ and a range of k values from 0.01 to 0.2 and seemed to give consistent results. Later tests with different cooling functions showed that choosing $C \approx 10$ gives the best results for k up to about 0.1. More extensive testing should be done to check the most appropriate values for C. Different cooling schedules all seemed to give the same apparently optimal solutions after a few hundred thousand iterations and these solutions could not be improved upon even if millions of iterations were tried.

In summary, the fact that *MultipleTree* gives the same solution when run with different cooling schedules suggests that for these problems the simulated annealing method is actually finding the true maximum of the objective function, and in a very short time. The time was less than a minute for the values of k, M and m in the test examples mentioned above. This is to be expected when the state space is asymmetric. In this regard the region of the space containing the trees with the highest breeding values, where the optimum would be expected to lie, acts in an attractive manner. For the same reason, one would expect genetic algorithms also to perform well.

Further evidence that the true maximum is being found by simulated annealing was obtained when the *SingleTree* problem was reformulated as an integer linear programming problem. Since the linear programming algorithm uses approximately $M^2/2$ variables it was necessary to make the problem size more managable by restricting our attention to the top 100 trees, ranked by v_i. This was regarded as a fairly safe restriction since with $m = 20$, all trees used in the solution found by *SingleTree* were ranked in the top 80 trees. After several hours, a standard package confirmed the solution found by *SingleTree*. The *MultipleTree* problem is much harder to write as an integer linear programming problem because each quadratic term $b_{ij}m_im_j$ in equation (13) can take on more than two possible values.

5. Existence of a solution in non-negative integers for the basic tree selection model in the case of a population of unrelated clones.

It is common practice in seed orchards to plant equal numbers of essentially unrelated clones. We ask whether the expected breeding value could be increased

by using more of the superior clones. In this case we have

$$a_{ij} = 0 \tag{51}$$

for $j \neq i$ and the objective function defined in equation (11) reduces to

$$V = \sum_{i=1}^{M} [v_i m_i - k m_i^2] \tag{52}$$

and is a special case of the more general form

$$V = \sum_{i=1}^{M} [v_i m_i - f(m_i)] \tag{53}$$

which must be maximised subject to condition (10) and subject to the restriction that each component of $\mathbf{m} = (m_i) \in \Re^M$ is a non-negative integer. Once again we assume that $v_1 \geq v_2 \geq \ldots \geq v_M$. If $\mathbf{m} \in \Re^M$ is a feasible selection and if $m_p < m_q$ for some $p < q$ then we can see that the new selection with m_p and m_q interchanged will change V by an amount

$$\Delta V = (v_p - v_q)(m_q - m_p). \tag{54}$$

If $v_p > v_q$ then $\Delta V > 0$ and the new selection is better. Since there are only a finite number of alternatives it follows from this reasoning that an optimal solution exists and that the optimal solution will satisfy

$$m \geq m_1 \geq m_2 \geq \ldots \geq m_M. \tag{55}$$

This problem can be solved very effectively using dynamic programming with the total number of iterations required of the order of M^3. Although it is not necessary to argue that condition (55) holds it is nevertheless true that the use of this condition makes the solution scheme more efficient. We formulate the problem in the following way. Let $P(j,t)$ denote the problem of selecting the numbers

$$m_{j+1} \geq m_{j+2} \geq \ldots \geq m_M \geq 0 \tag{56}$$

with

$$\sum_{i=j+1}^{M} m_i = t \tag{57}$$

so that

$$V_j = \sum_{i=j+1}^{M} [v_i m_i - f(m_i)] \tag{58}$$

is maximised. If we write

$$V(j,t) = \max\{V_j \mid m_{j+1} \geq \ldots \geq m_M \geq 0 \text{ and } m_{j+1} + \ldots + m_M = t\} \tag{59}$$

then it is clear that

$$V(j,t) = \max_{m_{j+1}} [v_{j+1}m_{j+1} - f(m_{j+1})] + V(j+1, t-m_{j+1}). \quad (60)$$

The dynamic programming solution uses equation (60) to solve the problems $P(M,t), P(M-1,t), \ldots$ in sequence for all $t = 0, 1, \ldots, m$.

6. The entropy model

The *genetic diversity* of the current population $\mathcal{P}_t$ can be measured by the *entropy* which is defined by

$$H(\mathcal{P}_t) = (-1)\sum_{j=1}^{N} \nu_j(\mathcal{P}_t) \log \nu_j(\mathcal{P}_t) \quad (61)$$

where $\boldsymbol{\nu}(\mathcal{P}_t)$ is the genetic measure defined in equation (6). For the original population we have

$$\nu_j(\mathcal{P}_0) = \frac{1}{N} \quad (62)$$

for each $j = 1, 2, \ldots, N$ and hence

$$H(\mathcal{P}_0) = \log N. \quad (63)$$

It is clear that

$$H(\mathcal{P}_t) \leq H(\mathcal{P}_0) \quad (64)$$

and although it is not true to say that the entropy of the population decreases with time we observe that if

$$\nu_j(\mathcal{P}_t) = 0 \quad (65)$$

for $j = j_1, j_2, \ldots, j_n$ then

$$H(\mathcal{P}_s) \leq \log(N-n) \quad (66)$$

for all $s \geq t$. These ideas are also illustrated by the simple example in Section 8.

As before we suppose the current population $\mathcal{P} = \mathcal{P}_t$ is denoted by $\beta_1, \ldots, \beta_M$ and that we use m_i clones of the individual β_i for each $i = 1, 2, \ldots, M$. We suppose that the genetic content of β_i is denoted by

$$\boldsymbol{\nu}(\beta_i) = \begin{pmatrix} \nu_{i1} \\ \nu_{i2} \\ \vdots \\ \nu_{iN} \end{pmatrix} \quad (67)$$

and we note that

$$\sum_{j=1}^{N} \nu_{ij} = 1. \tag{68}$$

The entropy of the effective population is given by

$$H = (-1)\sum_{j=1}^{N} \nu_j \log \nu_j \tag{69}$$

where

$$\nu_j = \sum_{i=1}^{M} \nu_{ij} x_i. \tag{70}$$

It is convenient to begin by solving an unconstrained problem. We show that the entropy of the effective population is maximised by choosing $x_1, x_2, \ldots, x_M$ in such a way that

$$\nu_1 = \nu_2 = \ldots = \nu_N = \frac{1}{N}. \tag{71}$$

Because of the conditions (10) and (68) we have

$$\begin{aligned} H &= H(x_1, x_2, \ldots, x_{M-1}) \\ &= (-1)\sum_{j=1}^{N} \nu_j \log \nu_j \end{aligned} \tag{72}$$

where

$$\nu_j = \sum_{i=1}^{M-1} (\nu_{ij} - \nu_{Mj}) x_i + \nu_{Mj} \tag{73}$$

for each $j = 1, 2, \ldots, N-1$ and where

$$\nu_N = 1 - \sum_{j=1}^{N-1} \nu_j. \tag{74}$$

Now we calculate

$$\frac{\partial H}{\partial x_i} = (-1)\sum_{j=1}^{N-1} (\nu_{ij} - \nu_{Mj}) \log\left(\frac{\nu_j}{\nu_N}\right) \tag{75}$$

and solve the equations

$$\frac{\partial H}{\partial x_i} = 0 \tag{76}$$

for each $i = 1, 2, \ldots, M-1$. It is clear that this gives $M-1$ linear equations in the $N-1$ unknowns

$$\log\left(\frac{\nu_1}{\nu_N}\right), \ldots, \log\left(\frac{\nu_{N-1}}{\nu_N}\right). \tag{77}$$

If $M \geq N$ and the rank of this set is equal to $N-1$ then the unique solution is given by

$$\log\left(\frac{\nu_j}{\nu_N}\right) = 0 \tag{78}$$

for each $j = 1, 2, \ldots, N-1$. Thus, in this generic case, we can see that H is maximised when $\nu_1 = \nu_2 = \ldots = \nu_N$.

Strictly speaking we should solve the above problem subject to the additional constraint $\mathbf{x} \geq \mathbf{0}$ but because the solution is not conditional on these restrictions we can formally omit them. This is an important point in relation to our subsequent arguments with the real constrained selection problem.

We now wish to maximise the entropy H subject to the constraints $V \geq V_0$, where V_0 is the minimum allowed total breeding value, $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{1}^T\mathbf{x} = 1$. We define a Lagrangean function

$$\begin{aligned} \mathcal{H} &= \mathcal{H}(x_1, x_2, \ldots, x_{M-1}) \\ &= H + \kappa\left[\sum_{i=1}^{M-1}(v_i - v_M)x_i + v_M - V_0\right] + \sum_{i=1}^{M-1}(\lambda_i - \lambda_M)x_i + \lambda_M \end{aligned} \tag{79}$$

where $\kappa \in \Re$ and $\boldsymbol{\lambda} \in \Re^M$ are Lagrange multipliers with $\kappa \geq 0$, $\boldsymbol{\lambda} \geq \mathbf{0}$ and apply the Kuhn-Tucker equations

$$\frac{\partial \mathcal{H}}{\partial x_i} = 0$$

for each $i = 1, 2, \ldots, M-1$ and the complementary slackness conditions to obtain the necessary conditions

$$(-1)\sum_{j=1}^{N-1}(\nu_{ij} - \nu_{Mj})\log\left(\frac{\nu_j}{\nu_N}\right) + \kappa(v_i - v_M) + \lambda_i - \lambda_M = 0 \tag{80}$$

for each $i = 1, 2, \ldots, M-1$ and

$$\kappa\left[\sum_{i=1}^{M-1}(v_i - v_M)x_i + v_M - V_0\right] + \sum_{i=1}^{M-1}(\lambda_i - \lambda_M)x_i + \lambda_M = 0. \tag{81}$$

Therefore we again have $M-1$ linear equations in the $N-1$ unknowns

$$\log\left(\frac{\nu_1}{\nu_N}\right), \ldots, \log\left(\frac{\nu_{N-1}}{\nu_N}\right) \tag{82}$$

but since the equations are now non-homogeneous we can no longer expect an unconstrained solution. In fact, if we assume that $\lambda_i = 0$ for N different values of i we can obtain a solution for the corresponding x_i in which each of these x_i is strictly positive. All other x_i are set to zero and the corresponding λ_i and κ are determined from the remaining equations.

The important insight is that we expect only N of the variables $x_1, \ldots, x_M$ will be non-zero.

More information about this technique can be obtained from a technical report by Howlett *et al.* (1996).

7. The enhanced tree selection model

In the previous models an implicit assumption about breeding values is that the breeding value of the offspring of a particular mating is the average of the breeding values of the two parents and hence we have assumed that

$$v(\beta_i \times \beta_j) = \frac{1}{2}v(\beta_i) + \frac{1}{2}v(\beta_j). \tag{83}$$

This is tantamount to saying that the best parents will produce the best offspring and on average this is observed to be true. Although this is a useful assumption in models with an unbiassed selection of the next generation of breeding trees the assumption is not appropriate when we choose only the best progeny from each cross and when we wish to make projections about the consequent improvement in breeding values over several generations.

It is clear that the genetic structure of the offspring is not determined uniquely for each mating pair. Each individual in the population carries a fixed number of chromosomes and the offspring of a particular cross receives these chromosomes from one or other of the parents. The allocation of the chromosomes is essentially a random process. If there are k different chromosomes then there are 2^k different genetic combinations that could be obtained. It is therefore more reasonable to regard the breeding value V_{ij} for the progeny $\beta_i \times \beta_j$ as a random variable. Since the breeding value is determined by a large number of independent characteristics we could assume that V_{ij} is a normal random variable with probability density function given by

$$\begin{aligned} f_{ij}(v) &= N[\mu_{ij}, {\sigma_{ij}}^2](v) \\ &= \frac{1}{\sigma_{ij}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{v-\mu_{ij}}{\sigma_{ij}}\right)^2\right]. \end{aligned} \tag{84}$$

Appropriate values for the parameters should really be determined by experiment. On the other hand we could assume that the breeding value V_i of each individual β_i is a random variable and that the random variable V_{ij} is determined from the random variables V_i and V_j by a simple formula. Indeed, if we assume that

$$V_{ij} = \frac{1}{2}V_i + \frac{1}{2}V_j \tag{85}$$

then we could argue that

$$\begin{aligned} \mu_{ij} &= E[V_{ij}] \\ &= E[\frac{1}{2}V_i + \frac{1}{2}V_j] \\ &= \frac{1}{2}\mu_i + \frac{1}{2}\mu_j \end{aligned} \tag{86}$$

and

$$\begin{aligned} {\sigma_{ij}}^2 &= E[(V_{ij} - \mu_{ij})^2] \\ &= E[\frac{1}{4}(V_i - \mu_i)^2 + \frac{1}{4}(V_j - \mu_j)^2 + \frac{1}{2}(V_i - \mu_i)(V_j - \mu_j)] \\ &\sim \frac{1}{4}{\sigma_i}^2 + \frac{1}{4}{\sigma_j}^2 + \frac{1}{2}a_{ij}\sigma_i\sigma_j \end{aligned} \tag{87}$$

where $A = (a_{ij}) \in \Re^{M \times M}$ is the relatedness matrix and where we have assumed for the sake of argument that

$$E[(V_i - \mu_i)(V_j - \mu_j)] \sim a_{ij}\sigma_i\sigma_j. \tag{88}$$

If n_{ij} is the number of progeny produced from $\beta_i \times \beta_j$ and if w_{ij} is the minimum acceptable breeding value from this cross then the expected number of progeny selected for the next breeding population will be given by

$$\begin{aligned} m_{ij} &= n_{ij}F_{ij}(w_{ij}) \\ &= n_{ij}\int_{w_{ij}}^{\infty} f_{ij}(v)dv \end{aligned} \tag{89}$$

and the associated expected mean breeding value will be

$$\begin{aligned} v_{ij} &= \int_{w_{ij}}^{\infty} v f_{ij}(v)dv \\ &= \frac{\sigma_{ij}}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{w_{ij} - \mu_{ij}}{\sigma_{ij}}\right)^2\right] + \mu_{ij}F_{ij}(w_{ij}). \end{aligned} \tag{90}$$

We need to select n_{ij} and w_{ij} for $i, j = 1, 2, \ldots, M$ and $i \leq j$ such that we maximise the overall expected breeding value

$$E = \sum_{1 \leq i \leq j \leq M} m_{ij}v_{ij} \tag{91}$$

subject to the equality constraint

$$\sum_{1 \leq i \leq j \leq M} m_{ij} = m \tag{92}$$

and the inequality constraint

$$H(\mathcal{P}_s) \geq \log r \tag{93}$$

where m is the total number of progeny we wish to select, $\mathcal{P}_s$ is the population of selected progeny and r is some fixed number with $0 \leq r < N$. This problem can be formulated as a standard finite dimensional constrained optimisation problem.

The main importance of this model is that it is not only a model for making a good selection for the next generation of breeding stock but is also a model that allows us to estimate the expected improvement in the breeding population from one generation to the next. This could be useful for forward projections concerned with the economic viability of the breeding operation.

8. A simple example

Let $\mathcal{P}_0 = \{\alpha_1, \alpha_2, \alpha_3\}$ and suppose that the following individuals are added to the population

$$\begin{aligned} \alpha_4 &= \alpha_1 \times \alpha_2, \\ \alpha_5 &= \alpha_4 \times \alpha_3 = (\alpha_1 \times \alpha_2) \times \alpha_3, \\ \alpha_6 &= \alpha_5 \times \alpha_2 = ((\alpha_1 \times \alpha_2) \times \alpha_3) \times \alpha_2. \end{aligned} \tag{94}$$

The pedigrees of each individual are

$$p(\alpha_1) = 1, \quad p(\alpha_2) = 2, \quad p(\alpha_3) = 3, \tag{95}$$

and

$$p(\alpha_4) = (1,2), \quad p(\alpha_5) = ((1,2),3), \quad p(\alpha_6) = (((1,2),3),2); \tag{96}$$

and the genetic contents of each individual are

$$\boldsymbol{\nu}(\alpha_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\nu}(\alpha_2) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\nu}(\alpha_3) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{97}$$

and

$$\begin{aligned} \boldsymbol{\nu}(\alpha_4) &= \frac{1}{2}\boldsymbol{\nu}(\alpha_1) + \frac{1}{2}\boldsymbol{\nu}(\alpha_2) = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}, \\ \boldsymbol{\nu}(\alpha_5) &= \frac{1}{2}\boldsymbol{\nu}(\alpha_4) + \frac{1}{2}\boldsymbol{\nu}(\alpha_3) = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{2} \end{pmatrix}, \\ \boldsymbol{\nu}(\alpha_6) &= \frac{1}{2}\boldsymbol{\nu}(\alpha_5) + \frac{1}{2}\boldsymbol{\nu}(\alpha_2) = \begin{pmatrix} \frac{1}{8} \\ \frac{5}{8} \\ \frac{1}{4} \end{pmatrix}. \end{aligned} \tag{98}$$

The genetic measure $\nu(\mathcal{P}) = (\nu_j(\mathcal{P}) \in \Re^3$ is given by

$$\begin{aligned} \nu_1(\mathcal{P}) &= \frac{1}{6}(1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}) = \frac{5}{16} \\ \nu_2(\mathcal{P}) &= \frac{1}{6}(1+\frac{1}{2}+\frac{1}{4}+\frac{5}{8}) = \frac{19}{48} \\ \nu_3(\mathcal{P}) &= \frac{1}{6}(1+\frac{1}{2}+\frac{1}{4}) = \frac{7}{24}. \end{aligned} \tag{99}$$

The relatedness of individuals in the current population is defined by the relatedness matrix $A \in \Re^{6\times 6}$ given by

$$A = \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{4} & \frac{5}{8} \\ 0 & 0 & 1 & 0 & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 & \frac{1}{2} & \frac{5}{8} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & \frac{1}{2} & 1 & \frac{5}{8} \\ \frac{1}{8} & \frac{5}{8} & \frac{1}{4} & \frac{5}{8} & \frac{5}{8} & 1 \end{pmatrix} \tag{100}$$

and the entropy is

$$\begin{aligned} H &= \frac{5}{16}\log\left(\frac{16}{5}\right) + \frac{19}{48}\log\left(\frac{48}{19}\right) + \frac{7}{24}\log\left(\frac{24}{7}\right) \\ &\approx 1.0897. \end{aligned} \tag{101}$$

Note that

$$H_0 = \log 3 \approx 1.0986. \tag{102}$$

9. Another simple example

Consider a population

$$\mathcal{P} = \{\beta_1, \beta_2, \beta_3, \beta_4\} \tag{103}$$

with $v_1 = v_2 = 6$, $v_3 = 4$ and $v_4 = 1$ and with relatedness matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{104}$$

If we choose k so that $km = 2$ then

$$V = 6(x_1 + x_2) + 4x_3 + x_4 - 2\left[2(x_1 + x_2)^2 + {x_3}^2 + {x_4}^2\right] \tag{105}$$

and

$$\mathcal{V} = V + \lambda(1 - \mathbf{1}^T\mathbf{x}) + \boldsymbol{\pi}^T\mathbf{x} \tag{106}$$

and if $\mathbf{x} > \mathbf{0}$ then $\boldsymbol{\pi} = \mathbf{0}$ and the Kuhn-Tucker equations give

$$\mathbf{v} - 4B\mathbf{x} - \lambda\mathbf{1} = \mathbf{0}. \tag{107}$$

If we wish to resolve the various vectors into orthogonal components in $\mathcal{N}$ and $\mathcal{R}$ then we have

$$\mathbf{v}_n = \mathbf{0}, \qquad \mathbf{v}_r = \begin{pmatrix} 6 \\ 6 \\ 4 \\ 1 \end{pmatrix}$$

$$\mathbf{x}_n = \begin{pmatrix} \frac{x_1 - x_2}{2} \\ \frac{-x_1 + x_2}{2} \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{x}_r = \begin{pmatrix} \frac{x_1 + x_2}{2} \\ \frac{x_1 + x_2}{2} \\ x_3 \\ x_4 \end{pmatrix}$$

$$\mathbf{1}_n = \mathbf{0}, \qquad \mathbf{1}_r = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \tag{108}$$

In this case equation (29) is automatically satisfied and equation (30) becomes

$$\begin{pmatrix} 6 \\ 6 \\ 4 \\ 1 \end{pmatrix} - 4 \begin{pmatrix} 2 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{x_1 + x_2}{2} \\ \frac{x_1 + x_2}{2} \\ x_3 \\ x_4 \end{pmatrix} - \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{0} \tag{109}$$

which gives

$$\mathbf{x}_r = \begin{pmatrix} \frac{6-\lambda}{16} \\ \frac{6-\lambda}{16} \\ \frac{4-\lambda}{4} \\ \frac{1-\lambda}{4} \end{pmatrix}. \tag{110}$$

We can now use condition (10) to calculate

$$\lambda = \frac{8}{5} \quad \text{and} \quad \mathbf{x}_r = \begin{pmatrix} \frac{11}{40} \\ \frac{11}{40} \\ \frac{3}{5} \\ -\frac{3}{20} \end{pmatrix} \tag{111}$$

which is not feasible. If we write

$$\mathbf{e}_n = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix} \tag{112}$$

then we find that the solution set

$$\mathcal{S} = \{\mathbf{x}|\mathbf{x} = \mathbf{x}_r + \theta \mathbf{e}_n \text{ for } \theta \in \Re\} \tag{113}$$

does not intersect the feasible set $\mathcal{F}$. On the other hand if we assume that $x_4 = 0$ and delete x_4 from the problem then equation (30) becomes

$$\begin{pmatrix} 6 \\ 6 \\ 4 \end{pmatrix} - 4 \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{x_1+x_2}{2} \\ \frac{x_1+x_2}{2} \\ x_3 \end{pmatrix} - \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{0} \tag{114}$$

which gives

$$\mathbf{x}_r = \begin{pmatrix} \frac{6-\lambda}{16} \\ \frac{6-\lambda}{16} \\ \frac{4-\lambda}{4} \end{pmatrix} \tag{115}$$

and we can use condition (10) to calculate

$$\lambda = 2 \text{ and } \mathbf{x}_r = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{2} \end{pmatrix} \tag{116}$$

which is feasible. If we write

$$\mathbf{e}_n = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \tag{117}$$

the entire solution set is

$$\mathcal{S} = \{\mathbf{x}|\mathbf{x} = \mathbf{x}_r + \theta \mathbf{e}_n \text{ for } \theta \in \Re\} \tag{118}$$

and the set of feasible solutions is given by

$$\mathcal{S} \cap \mathcal{F} = \{\mathbf{x}|\mathbf{x} = \mathbf{x}_r + \theta \mathbf{e}_n \text{ for all } \theta \in [-\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4}]\} \tag{119}$$

with $V(\mathbf{x}) = V_{max} = \frac{7}{2}$ for all $\mathbf{x} \in \mathcal{S} \cap \mathcal{F}$.

10. Discussion

In summary we simply observe that *The basic tree selection model* has been fully analysed and tested and appears to offer an effective selection procedure for the next generation of breeding stock. *The entropy model* although not fully developed offers an alternative procedure that we believe would simply change the emphasis of the selection rather than the nature. *The enhanced tree selection model* is a proposal that could be used as a basis for a more comprehensive scheme that is concerned with longer term analysis over several generations.

Acknowledgements

The moderators, Colin Thompson and Phil Howlett, would like to thank the industry representatives, Nuno Borrahlo, Mike Powell, Richard Kerr, and Peter Gore, and the other team members, Nick Wormald, Bill Whiten, Jerzy Filar, Genevieve Mortiss, Ernie Chow, Jemery Day, Natashia Boland, Pierre Prado, David Peel, Matthew York and Peter Watson for their special inspirations and excellent co-operation. We were particularly grateful for the willingness of team members to undertake the various specific tasks that were handed out from time to time. We would also like to thank Tim Thwaites for compiling the executive summary.

References

J.R. Brisbane and J.P. Gibson, "Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions" *Theory and Application of Genetics* **91** (1995), 421–431.

P.G. Howlett, J.A. Filar, C.J. Thompson and G. Mortiss, "An entropy model for selection and mating decisions in tree breeding programs", *School of Mathematics Report, University of South Australia* (in preparation).

G.B. Jansen and J.W. Wilton, "Selecting mating pairs with linear programming techniques", *Journal of Dairy Science* **68** (1985), 1302–1305.

T.H.E. Meuwissen and J.A. Woolliams, "Response versus risk in various breeding schemes", *Proceedings 5th. World Conference of Genetics applied to Livestock Production, University of Guelph, Canada* (1994), 236–243.

M. Sedgley and A.R. Griffin, *Sexual Reproduction of Tree Crops* (Academic Press, 1989).