Optimal Utilization of Clones and Genetic Thinning of Seed Orchards

By L. Bondesson[1]) and D. Lindgren[2])

Swedish University of Agricultural Sciences,
S-901 83 Umeå, Sweden

(Received 15th November 1991)

Abstract

A mathematical method is developed for assigning clones utilization factors (= relative numbers of copies) given an upper bound on the utilization factor of each clone. The mean genetic value is maximized under a diversity constraint. The optimal utilization for a clone is linearly related to the genetic value of the clone when this value is not too low and the upper bound is inactive. It is argued that this method is a useful tool for genetic thinning of seed orchards. The method is extended to the situation when the clones can be organized into groups of similar clones. For each group an optimal linear utilization can be found; in the simplest case the slopes are the same for all the groups. The method is applied to a Norway spruce seed orchard with 36 clones for which predicted breeding values are available.

Key words: Diversity, genetic value, truncation selection, optimization, Kuhn-Tucker theorem, Norway spruce.

1. Introduction

Genetic improvement may be defined as a process where genetic value is improved while diversity is considered. Diversity can be described in terms of effective population size. For seed orchards genetic diversity is of big importance since forests have a very long life and are difficult to manage after planting. Risk of pests and diseases should be reduced already at stand establishment.

According to Crow and Kimura (1979), truncation selection is known to be the most efficient form of directional selection. However, this view has later been revised. Taking into account both genetic value and diversity, there are more efficient ways to assign clones (genotypes) utilization factors (= relative number of copies). This has been demonstrated by examples (Kang and Namkoong, 1988; Kang, 1989), by numerical optimization techniques (Toro and Nieto, 1984), and by mathematical derivation of optimal utilization factors (e. g. Lindgren and Matheson, 1986; Lindgren et al., 1989; Bondesson, 1989).

In our earlier treatments, it was assumed that there is no upper limit on the utilization of a clone. In the present study this restriction is removed in order to make the method more practical. Moreover, the method is extended to cover also the case when the clones have a group structure.

The results of this study are of particular relevance to genetic thinning of clonal seed orchards. Seed orchards are improved by removing ramets belonging to clones with low breeding values. In that situation the utilization of a clone is strictly constrained by the existing number of ramets in the seed orchard. The number of ramets can be decreased but not increased. When a seed orchard is thinned the selection is often weak. In this case, in particular, truncation selection is inefficient.

This paper is organized as follows: In Section 2 the concepts of mean genetic value and diversity are defined. It is argued that each clone should be utilized so that the mean genetic value is maximized for a fixed value of the diversity. In Section 3 the optimal solution is derived with the help of the Kuhn-Tucker theorem. Moreover, an example is given. The case when the clones have a group structure is considered in Section 4. In this case the diversity measure is modified. Finally, in Section 5, there is some discussion.

2. Mean Genetic Value and Diversity

Let C_1, C_2, ..., C_n be n available clones with genetic values denoted g_1, g_2, ..., g_n. At the moment these values are assumed to be well-determined by progeny testing. The clones can be used to lay out an orchard producing seeds (or cuttings). Alternatively, they can already be in a seed orchard that should be thinned. The clones can be used in different frequencies p_1, p_2, ..., p_n. These numbers are non-negative and sum to 1. We may interpret p_i as the relative number of ramets of C_i, for instance.

When a seed orchard is laid out there is a priori no other restrictions on the frequencies than $0 \leq p_i \leq 1$, $i = 1, 2, ..., n$. If it is impossible to produce many ramets of some clones, the upper bound 1 has to be replaced by lower upper bounds b_i, $i = 1, 2, ..., n$. When a seed orchard is thinned we must have $0 \leq p_i \leq b_i = p^0_i/(1-T)$, where p^0_i is the frequency of C_i before thinning and T is the desired degree of thinning. For example, if we want to thin 50%, then $p_i \leq 2 \cdot p^0_i$.

The mean genetic value in the new or thinned seed orchard will be $G = \Sigma p_i g_i$. Assuming random crossings and additivity of genetic values and noting that G also equals $\Sigma\Sigma\ p_i p_j (g_i + g_j)/2$, we see that G is the relevant measure of what the seed orchard can produce. Negative effects of selfing are here neglected, cf. Remark 2.2 below. It is desirable that the frequencies are chosen so that G is made as large as possible.

It is not only important to have a large value of G in a seed orchard. It is also important to have high diversity there. High diversity will give a general protection against all possible unknown future calamities for the progenies.

There are several ways of defining the diversity mathematically. The simplest and most natural such measure is $1/\Sigma\ p^2_i$, which equals n when all the clones have equal frequencies. The measure is often called "effective population size", cf. Robertson (1961). In our case it is actually the effective number of clones. Another name for it, or its logarithm, is "Simpson diversity" (Pielou, 1982). As a measure of the relative diversity, compared to the case when the n clones have equal frequencies, we here use

$$\alpha = \frac{1}{n \cdot \Sigma p_i^2}$$

---

[1]) Department of Biometry and Forest Management

[2]) Department of Forest Genetics and Plant Physiology

It is used instead of $1/\Sigma p^2_i$ mainly to simplify the presentation. We have $1/n \leq \alpha \leq 1$ with $\alpha = 1$ if and only if $p_i \equiv 1/n$.

Remark 2.1.

The choice of diversity measure can be motivated as follows. In the future the progenies (one generation) of the n clones will be affected by accumulated random damages $X_1, X_2, \ldots, X_n$. Neglecting a possible common component of these random variables, we may in the simplest case assume them to be stochastically independent with a common variance. Then the variance of the total damage, $\Sigma X_i p_i$, is proportional to Σp^2_i. The variance should be made small, i. e. $1/\Sigma p^2_i$ or, equivalently, α should be made large. Thoughts about inbreeding in the future forests give also support for the measure α. The number of crossings between progenies of ramets belonging to the same clone can be assumed to be proportional to Σp^2_i and should be kept low; cf. e.g. LINDGREN (1991).

High α and high G are two goals that are in conflict with each other. The approach we make in this paper is to fix α in advance at some target value and then try to choose the frequencies of the clones in such a way that G is maximized. We are thinking of a situation where a law or rule prescribes that there must be some minimum diversity. The rule may prescribe that the seed orchard should contain a minimum number of different clones.

Alternatively, one may instead fix G at some value and try to maximize the diversity. However, it turns out that the mathematical form of the solution does not differ from the above one, cf. Remark 3.2. A third approach would be to try to maximize for example G — constant · Σp^2_i (cf. Remark 2.2) under no other constraints than the natural ones. However, it is difficult to choose an appropriate value of the constant.

Remark 2.2.

Selfing (within the orchard) may lead to a loss that should be taken into account. The loss can be assumed to be proportional to or linearly related to Σp^2_i. Fortunately, to maximize G — constant · $\Sigma_p{}^2_i$ given the diversity is equivalent to maximizing G given the diversity.

Remark 2.3.

In general g_i is an estimated value of a true genetic value G_i. It can be assumed that $g_i = G_i + \delta_i$, where δ_i is a random error with mean zero and constant variance σ^2 and with at least approximately a Normal distribution. We want to maximize $G' = \Sigma\, p_i G_i$. We make a Bayesian approach and assume that the G_i's have a prior distribution. Then we can maximize the conditional expected value of G' given $g_1, g_2 \ldots, g_n$. Assuming that the G_i's stem from a Normal population with mean μ and variance τ^2, we have, with $p^2 = \tau^2/(\tau^2+\sigma^2)$,

$$E[G'|g_1, \ldots g_n] = \Sigma\, p_i E[G_i|g_i] = \Sigma\, p_i(\mu + \rho^2 \cdot (g_i - \mu)) = (1-\rho^2)\mu + \rho^2 \cdot \Sigma\, p_i g_i.$$

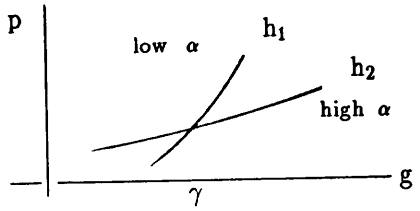
It follows that in the case of measurement errors $E[G'|g_1, \ldots, g_n]$ is maximized when $G = \Sigma\, p_i g_i$ is maximized. It may be added that the relation $E[G_i|g_i] \approx u + p^2 (g_i - \mu)$ holds whenever the distribution of G_i is not extremely non—Normal. It should also be added that $u + p^2\, (g_i - \mu)$ is the best linear predictor of G_i whatever the distribution of G_i is.

We end this section with a simple but important lemma to be invoked in Section 3. It has also an independent interest.

Lemma 2.1.

Let h_1 and h_2 be 2 nonnegative increasing functions such that $h_1(g) \leq h_2(g)$ for $g < \gamma$, $h_1(g) \geq h_2(g)$ for $g > \gamma$ and $h_1(\gamma) = h_2(\gamma)$. Then a population with frequencies $p_i = h_1\, (g_i)$, $i = 1, \ldots, n$, has a strictly lower diversity than a population with frequencies $p_i = h_2(g_i)$, $i = 1, \ldots, n$. It is assumed that $\Sigma\, h_1(g_i) = \Sigma\, h_2(g_i) = 1$ and that $h_1(g_i) \neq h_2(g_i)$ for at least some i.

The following *figure 1* illustrates the lemma.



Proof

We must show that $\Sigma\, h^2{}_1(g_i) > \Sigma\, h^2{}_2(g_i)$. Let Σ^- (Σ+) indicate summation over the clones for which $g_i < \gamma$ ($g_i > \gamma$). The assumptions and some simple estimates give

$$\Sigma\, h_1^2(g_i) - \Sigma\, h_2^2(g_i) = \Sigma\, (h_1(g_i) - h_2(g_i)) \cdot (h_1(g_i) + h_2(g_i))$$

$$= \Sigma^-\, (h_1(g_i) - h_2(g_i)) \cdot (h_1(g_i) + h_2(g_i)) + \Sigma^+\, (h_1(g_i) - h_2(g_i)) \cdot (h_1(g_i) + h_2(g_i))$$

$$> \Sigma^-\, (h_1(g_i) - h_2(g_i)) \cdot (h_1(\gamma) + h_2(\gamma)) + \Sigma^+\, (h_1(g_i) - h_2(g_i)) \cdot (h_1(\gamma) + h_2(\gamma)) = 0.$$

The last equality follows from the fact that $\Sigma\, h_1(g_i) = \Sigma\, h_2(g_i) = 1$.

Remark 2.4.

It is not hard to show that the assumption that h_1 and h_2 be increasing can be replaced by the weaker assumption that $h_1(g) + h_2(g)$ exceeds (does not exceed) $h_1(\gamma) + h_2(\gamma)$ for $g > \gamma$ ($g < \gamma$) whenever $h_1(g) \neq h_2(g)$.

3. Optimal Utilization

The problem we have encountered reads mathematically as follows:
Maximize $G = \Sigma\, p_i g_i$ under the constraints

$$\Sigma\, p_i = 1, \quad \Sigma\, p_i^2 = 1/(n\alpha), \quad \textit{and} \quad 0 \leq p_i \leq b_i. \tag{3.1}$$

Here the desired diversity α is specified in advance as a number. The bounds b_i, $i = 1, 2, \ldots, n$, are also given numbers. We denote the maximum of the b_i's by b. All values of α in the interval (0,1] may not be possible. If $b_i \equiv b = n^{-1}/(1-T)$, then $\alpha \geq \alpha_{\min}$, where $\alpha_{\min} \approx 1-T$. The minimum diversity $\alpha_{\min}$ is obtained for truncation selection, i. e. when only the best clones are used and in equal proportions (excepted the clone with genetic value closest to the truncation point).

Now comes the first main result of this paper.

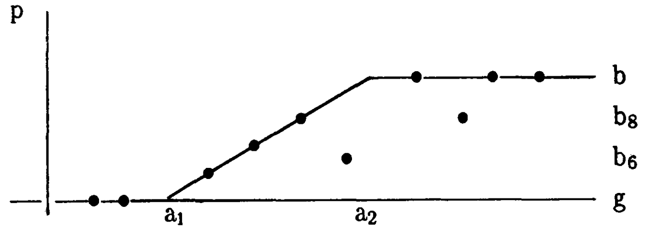
Theorem

The optimal solution to the maximization problem has the form $p_i = \min\, (\tilde{p}_i, b_i)$, where

$$\tilde{p}_i = \begin{cases} 0 & \text{for } g_i < a_1 \\ \beta\cdot(g_i - a_1) & \text{for } a_1 \leq g_i \leq a_2 \\ b & \text{for } g_i > a_2 \end{cases} \qquad (3.2)$$

Moreover, except in degenerate cases, a_1, a_2, and β ($= b/(a_2 - a_1)$) are uniquely determined by the constraints (3.1).

The following *figure 2* (for $n = 10$) illustrates the theorem.



Proof

Assuming that there is least one p-vector satisfying all the constraints (3.1), we first of all notice that there is an optimal solution since the set of admissible p-vectors is a bounded and closed region. Then the theorem is mainly a consequence of the KUHN-TUCKER theorem; see, e.g., LUENBERGER (1973, p. 233). Trying equivalently to minimize $-\Sigma p_i g_i$ and writing the 2n constraints $0 \leq p_i \leq b_i$ as $-p_i \leq 0$ and $p_i - b_i \leq 0$, we define the Lagrange function as follows:

$$L = -\Sigma p_i g_i + \lambda_1(\Sigma p_i - 1) + \lambda_2(\Sigma p_i^2 - \frac{1}{n\alpha}) + \Sigma u_i(-p_i) + \Sigma v_i(p_i - b_i).$$

Here λ_1, λ_2, u_1, ..., u_n and v_1, ..., v_n are the so-called Lagrange multipliers. Let now (p_1, ..., p_n) be the optimal solution. The KUHN-TUCKER theorem asserts that there are real constants $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ and non-negative $\tilde{u}_i$'s and $\tilde{v}_i$'s such that when the corresponding parameters have these values

$$\frac{\partial L}{\partial p_i} = 0, \quad \tilde{u}_i\cdot(-p_i) = 0, \quad \text{and} \quad \tilde{v}_i\cdot(p_i - b_i) = 0; \qquad i = 1, 2, ..., n.$$

In particular, we have $\tilde{u}_i = 0$ if $p_i > 0$ and $\tilde{v}_i = 0$ if $p_i < b_i$. Since $\partial L/\partial p_i = -g_i + \tilde{\lambda}_1 + 2\tilde{\lambda}_2 p_i - \tilde{u}_i + \tilde{v}_i$, we get

$$p_i = (2\tilde{\lambda}_2)^{-1}(g_i - \tilde{\lambda}_1 + \tilde{u}_i - \tilde{v}_i).$$

It is obvious that $\tilde{\lambda}_2 > 0$. We see that the optimal solution is linear in g_i for $0 < p_i < b_i$ as then $\tilde{u}_i = \tilde{v}_i = 0$. Clearly the line crosses the g-axis to the right of the rightmost clone with $p_i = 0$ since for all clones with p_i equal to zero $\tilde{u}_i \geqq 0$ and $\tilde{v}_i = 0$. Analogously, all clones for which $p_i = b_i$ are on or below the line; cf. the illustrating figure above. This gives the first part of the theorem.

To verify the uniqueness, we invoke Lemma 2.1. If there had been two different linear solutions yielding the same values of (Σp_i, Σp^2_i), then these 2 lines must cross each other at some point (γ, p) such that $0 < p < b$ since otherwise the sum of the p_i-values cannot be equal for the two solutions. If $b_i \equiv b$, Lemma 2.1 gives immediately a contradiction. Moreover, by the remark following Lemma 2.1, it can be verified that a contradiction is obtained also when $b_i \not\equiv b$. Of course, the line is not uniquely determined if there is no or only one clone C_i such that $0 < p_i < b_i$.

Remark 3.1.

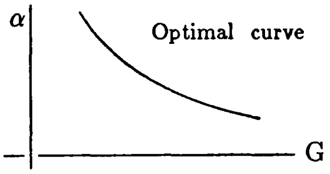
In the proof it has implicitly been assumed that the optimum point (p_1, ..., p_n) is a regular point; cf. LUENBERGER (1973, p. 233). In fact, this can be verified to be true except in the degenerate case when the line is not unique.

Remark 3.2.

The particular case when $b_i \equiv 1$ has earlier been treated in LINDGREN et al. (1989) and BONDESSON (1989). As to the 2 completely different proofs there, some questions have been raised about them. It is a hope that the short new proof, covering a more general case, will remove all doubts.

Remark 3.3.

The Kuhn-Tucker theorem also shows that if instead Σp^2_i is minimized (i.e., α is maximized) under the constraints $\Sigma p_i = 1$, $\Sigma p_i g_i = G$ (fixed), and $0 \leq p_i \leq b_i$, then the optimal solution still has the form (3.1). The following *figure 3* illustrates this fact.



For each value of α there is an optimal G-value. The optimal α-value corresponding to that G-value equals the original α.

There are several ways of calculating the optimal solution:

i) By a trial and error procedure. Different pairs (a_1, a_2) are tested and for each such pair Σp_i and Σp^2_i are calculated. We are satisfied when these numbers are sufficiently close to 1 and $1/(n\alpha)$, respectively. A systematic iterative search can be made so that the optimal solution is rapidly found.

ii) By first calculating the line $p_i = \beta_0 + \beta \cdot g_i$ for which $\Sigma p_i = 1$ and $\Sigma p^2_i = 1/(n\alpha)$. If all the remaining constraints are satisfied, the optimal solution is obtained. If not, the clone C_i for which $\max(-p_i, p_i - b_i)$ has its largest value is removed from the further calculations. Its p-value is set to 0 if $p_i < 0$ and otherwise to b_i. The procedure is then repeated with obvious modifications of the target values of Σp_i and Σp^2_i. The procedure is repeated until a solution with admissible p-values is obtained. In general this simple procedure leads to the desired solution though there may be exceptions. A correct solution is characterized by the property that all points (g_i, p_i) for which $p_i = b_i$ are on or below the full line whereas all points for which $p_i = 0$ are on or above the full line. The procedure is included in the package SELE that is distributed for nothing by D. LINDGREN on request.

iii) By using some general program for quadratic optimization. For example, the widely used package MATLAB has an addition, Optimization Toolbox, that contains such a program. It only produces the optimal p-values, not the parameters for the broken line.

Example 3.1 (36 clones)

For 36 clones of Norway spruce in seed orchard 52 at Maglehem in Skåne (Southern Sweden) the following genetic values (*Tab. 1*) (in increasing order) are available:

79.8	80.7	84.1	84.2	85.9	88.9	89.9	91.3	93.3
94.0	94.2	95.6	96.2	96.9	97.4	100.1	100.4	100.7
102.0	102.2	103.1	103.4	103.7	103.7	104.6	105.9	107.8
109.0	109.9	110.3	110.3	111.9	112.7	114.6	115.2	116.0

These values are predicted breeding values for height. They are expressed as percentages of a grand mean. The data are taken from KARLSSON and DANELL (1989, *Table A5*). We assume that all the clones have the same number ($\approx$ 60) of ramets in the seed orchard and that the genetic contribution to the seed orchard crop is equal to the fraction of ramets (to make the example simple). Suppose we want to thin the orchard by 50% by removing half of the ramets. The upper bound $b = (36)^{-1} / (1—0.5) = 1/18$ must then be imposed on the frequencies p_i. The smallest possible diversity there can be after the thinning is $\alpha = 0.5$ corresponding to thinning by truncation. By a program based on the technique in (ii) above, the following maximum G-values were obtained for some different values of α (Tab. 2). Comparisons are made with thinning by truncation for the same diversity and with the case when the upper bound on the p_i's is removed ($b = 1$). If the clones can be reproduced, we can have $b = 1$.

α	a_1	a_2	G_{max}	G_{trunc}	G_{max} (b=1)
0.5	101.35	101.35	108.13	108.13	109.10
0.6	93.78	107.51	107.59	106.76	107.76
0.7	87.21	113.71	106.40	105.26	106.40
0.8	80.30	119.72	104.94	103.75	104.94
0.9	70.35	129.65	103.29	102.00	103.29
0.95	56.92	143.08	102.27	101.04	102.27
0.98	30.81	169.18	101.41	100.41	101.41
1.00	$-\infty$	$+\infty$	100.00	100.00	100.00

The optimal frequencies are obtained as $p_i = (18)^{-1} \cdot (g_i—a_1)/(a_2—a_1)$ for $a_1 < g_i < a_2$. For $g_i < a_1$, $p_i = 0$ and for $g_i > a_2$, $p_i = 1/18$. We see that by using optimal thinning instead of thinning by truncation the mean genetic value can be improved by about 1% for α in the interval (0.6, 0.98). This is a rather substantial improvement. In particular, one should notice that the improvement is big also for α-values rather close to 1 (i.e. when the selection is weak). As $g_{min} = 79.8$, we also see that for $\alpha \geqq 0.81$, no clones are completely removed when optimal thinning is made. For $\alpha = 0.81$, thinning by truncation leads to the complete removal of $(1—0.81) \cdot 36 \approx 7$ clones. Optimal thinning keeps more clones in the orchard.

Example 3.2 (36 clones; ctd)

Assume that in the seed orchard the clones C_{20} ($g_{20} = 102.2$) and C_{30} ($g_{30} = 110.3$) are available only in the proportions 0.01 and 0.015, respectively, whereas the other clones have equal frequencies (i.e., 0.975/34). The original diversity is $\alpha = 0.982$. After 50% thinning the maximum frequency for all the clones except C_{20} and C_{30} is $b = 2 \cdot 0.975/34 = 0.05735$. For C_{20} and C_{30} there are the bounds $b_{20} = 0.02$ and $b_{30} = 0.03$. The program referred to above gives the following results (Tab. 3) which only slightly differ from the earlier results.

α	a_1	a_2	G_{max}	p_{20}	p_{30}	p_{18}
0.505	100.40	100.93	108.05	0.02	0.03	0.0324
0.6	93.31	107.26	107.48	0.02	0.03	
0.7	86.74	113.28	106.29	0.02	0.03	
0.8	80.01	119.03	104.84	0.02	0.03	
0.9	69.98	128.93	103.21	0.02	0.03	
0.998	$-\infty$	$+\infty$	100.14	0.02	0.028	(= all other p_i)

It should be noted that the maximum possible diversity, 0.998, equals 1 divided by $36 \cdot ((0.98/35)^2 \cdot 35 + (0.02)^3)$. The minimum possible diversity, 0.505, is obtained by setting p_i equal to the upper bound b for the 16 clones C_{19}, $C_{21}—C_{29}$, $C_{31}—C_{36}$, $p_{18} = 1 — 16 \cdot b —0.02 — 0.03 = 0.0324$, and p_{20} and p_{30} as above.

4. Groups of clones

It may be so that we know that the clones genetically resemble each other in some way or another. The solutions presented in Section 3 assume that in some sense all the clones are equally unsimilar. To cover also the more general situation, a new measure of the diversity is needed. In general the diversity can be defined to be inversely proportional to a quadratic form $Q = \Sigma\Sigma\, c_{ij}p_ip_j$, where c_{ij} ($0 \leqq c_{ij} < 1$) measures the similarity between C_i and C_j.

We may interpret c_{ij} as the correlation between the accumulated future damages X_i and X_j for the progenies of the clones C_i and C_j; cf. Remark 2.1. If all the X-variables have the same variance, the variance of the total damage ΣX_ip_i is proportional to Q above.

Alternatively and differently, we may interpret c_{ij} as a negative effect of crossings within the orchard (to make it simple) between the two clones C_i and C_j.

We consider from now on the important special case when there are m groups of clones such that all clones within a group are judged to be equally similar. We denote the clones C_{ij}, $j = 1, ..., n_i$; $i = 1, ..., m$. The corresponding g-values and frequencies are denoted g_{ij} and p_{ij}, respectively. We assume that for group i the similarity between any two clones can be described by a single number c_i. These numbers must be chosen more or less subjectively;

cf. the discussion in Section 5. The relative diversity is then given by

$$\alpha = \frac{1}{n \cdot Q} \quad \text{with} \quad Q = \sum_i c_i \Big(\sum_j p_{ij}\Big)^2 + \sum_i (1-c_i) \sum_j p_{ij}^2 \tag{4.1}$$

or, equivalently, $Q = \sum_i \sum_j c_i p_{ij} (\sum_k p_{ik} - p_{ij}) + \Sigma\Sigma p^2_{ij}$.

The factor 1/n is chosen so that comparison is made with the case when $c_i \equiv 0$ and all the clones have equal frequencies. If $c_i \equiv 0$, then $\alpha = 1/(n \cdot \Sigma\Sigma\, p^2_{ij})$ as in Section 2. If the clones within each group are extremely similar, we should have $c_i \equiv 1$ and $\alpha = 1/(n \cdot \Sigma\, P^2_i)$ with $P_i \equiv \sum_j p_{ij}$ (the frequency of group i).

Remark 4.1.

The maximum value of α can be shown to be $\alpha_{\max} = \frac{1}{n} \cdot \Sigma\, (c_i + \frac{1-c}{n_i})^{-1}$. That value is realized by letting the group frequency P_i be proportional to $(c_i + \frac{1-c_i}{n_i})^{-1}$ and letting, within each group, all the clones have equal frequencies. We have $\alpha_{\max} \leqq 1$ with equality if and only if $c_i \equiv 0$. Moreover, $\alpha_{\max}$ is larger than or equal to the diversity obtained for $p_{ij} \equiv 1/n$.

It can be shown by the KUHN-TUCKER theorem that $G = \Sigma\Sigma\, p_{ij} g_{ij}$ takes its maximum value under the constraints $0 \leqq p_{ij} \leqq b_{ij}$, $\Sigma\Sigma\, p_{ij} = 1$, and a fixed value of the diversity, only if p_{ij} has the form

$$p_{ij} = \text{MEDIAN}\left(0,\ \frac{\beta_0}{1-c_i} + \frac{\beta}{1-c_i} g_{ij} - \frac{c_i}{1-c_i} P_i,\ b_{ij}\right). \tag{4.2}$$

The right-hand side contains unknown p-values since P_i does. However, we see that the optimal solution is provided by essentially m broken lines, one line for each group. The slopes are $\beta/(1-c_i)$, i = 1, ..., m. Thus the lines are parallel if and only if the c_i's are equal. If $c_i \approx 1$ only the best clones in group i are preserved. The line for group i crosses the g-axis at $a_{i1} = (c_i P_i - \beta_0)/\beta$.

The relation (4.2) also holds if instead the diversity is maximized given the value of G.

In general the solution can be found numerically by the simple iterative removal technique described in (ii) in Section 3. For each iteration a system of equations has to be solved. The technique is simplest to implement when G instead of α is fixed in advance; cf. Remark 4.2 below. Varying G, any desired value of α can be reached.

Example 4.1 (36 clones; ctd)

Actually the clones in Example 3.1 are grouped into selections of plus trees from 12 different locations (stands), numbered 1, 2, ..., 12. From each of the locations 5, 6, 8, and 11 only one clone was selected and these 4 clones are therefore collected into an artificial group, R. The genetic values are now as follows (*Tab. 4*):

1	100.7	114.6					
2	88.9	91.3	95.6	102.2	107.8	109.0	
3	85.9	110.3					
4	79.8	80.7	94.0	103.7	104.6	111.9	115.2
7	84.1	94.2	103.7				
9	89.9	96.9	103.1	103.4	105.9		
10	97.4	102.0	109.9				
12	96.2	100.1	110.3	116.0			
R	84.2	93.3	100.4	112.7			

(There are no evident differences between the groups. In fact the "between groups" variance component is estimated to zero by the ANOVA-method.)

Assuming again that originally all the clones have equal frequencies, we get the upper bound b = 1/18 on the frequencies for a thinning by 50%. Using the iterative removal technique with fixed G-values, we get for some different c-values results as shown below (*Tab. 5*). The G-values are chosen so that they correspond to the diversities 0.5, 0.6, ..., 1.0 when $c_i \equiv 0$. The c_i-values are chosen to be equal for all the groups except the R-group; for the R-group $c_R = 0$.

G	α				
	$c_i \equiv 0$	$c_i \equiv 0.1$	$c_i \equiv 0.2$	$c_i \equiv 0.3$	$(c_R = 0)$
108.13	0.50	0.433	0.382	0.341	
107.59	0.60	0.513	0.454	0.410	
106.40	0.70	0.584	0.510	0.458	
104.94	0.80	0.651	0.561	0.499	
103.29	0.90	0.712	0.604	0.532	
100	1.00	0.768	0.641	0.558	

The minimum diversities are in the first row. The last main row gives the maximum diversities; the corresponding G-values may deviate very slightly from 100. We see that the diversities decrease considerably when the common c-values increase; this is quite natural.

The optimal α-values corresponding to a specific G-value are realized in different ways for different c-values. To illustrate, we choose G = 106.4. Then, with a_1 and a_2 defined as in Section 3, we get the following *table 6*.

	$c_i \equiv 0$		$c_i \equiv 0.1$		$c_i \equiv 0.2$		$c_i \equiv 0.3$	
	a_1	a_2	a_1	a_2	a_1	a_2	a_1	a_2
1	87.21	113.71	86.68	111.45	87.68	109.13	89.43	107.26
2	"	"	89.09	113.86	91.35	112.81	93.81	111.65
3	"	"	85.20	109.97	84.72	106.17	84.99	102.82
4	"	"	90.77	115.54	93.96	115.41	97.05	114.88
7	"	"	85.39	110.16	85.69	107.14	87.10	104.94
9	"	"	88.61	113.38	90.69	112.14	92.98	110.81
10	"	"	87.55	112.32	89.36	110.81	91.57	109.40
12	"	"	89.39	114.16	92.04	113.49	94.50	112.33
R	"	"	82.37	109.89	79.06	105.87	76.95	102.43
slope	0.002097		0.002243		0.002590		0.003115	

The slopes presented do not refer to the R-group; for the R-group the slope of the line is obtained by multiplying the given slope by $1—c_i$. We see that when the common c-values increase, the frequencies of the clones in the R-group increase whereas the clones with low g-values in the big groups (2, 4, and 9) get lower frequencies.. This is quite natural and desired. For example, for group 4, we have as easily can be verified (*Tab.* 7):

c_i	p_{41}	p_{42}	p_{43}	p_{44}	p_{45}	p_{46}
0	0	0	0.0142	0.0346	0.0365	0.0518
0.1	0	0	0.0072	0.0290	0.0310	0.0474
0.2	0	0	0.0001	0.0252	0.0276	0.0465
0.3	0	0	0	0.0208	0.0236	0.0465

Remark 4.2.

In this section it has implicitly been assumed that the g_{ij}-values are perfect estimates of the true genetic values G_{ij}. If not, some modification is needed. Assuming that $G_{ij} \sim N(\mu_i, \tau^2)$ and $g_{ij} \sim N(G_{ij}, \sigma^2)$, cf. Remark 2.3, we have, with $\rho^2 = \tau^2/(\tau^2+\sigma^2)$,

$$E[G_{ij}|g_{ij}] = (1-\rho^2)\mu_i + \rho^2 g_{ij} \approx (1-\rho^2)\bar{g}_i + \rho^2 g_{ij} ,$$

where $\bar{g}_i$ is a mean for group i. We infer that the proper thing to do is to replace the g_{ij}-values by the values $(1—\rho^2)\bar{g}_i + \rho^2 g_{ij}$ before the optimization is performed. An estimate of ρ^2 is needed but it may be rather rough.

Remark 4.3.

Finally it is appropriate to give more details on how the iterative removal technique is performed. In each step we must for the unremoved clones solve the system of equations

$$(1-c_i)p_{ij} + c_i P_i = \beta_0 + \beta \cdot g_{ij}$$

with all the p_{ij}'s and β_0 and β unknown, cf. (4.2). Here $P_i = \Sigma\, p_{ik}$, where the summation is performed over all clones in group i. Let $—\gamma_i$ be the sum of the p-values (0 or b_{ij}) for the removed clones in group i. Hence

$$p_{ij} + c_i \cdot \Sigma' p_{ik} = \gamma_i + \beta_0 + \beta \cdot g_{ij},$$

where ' indicates that the summation (in group i) excludes all removed clones and also the clone C_{ij}. For any values of β_0 and β, we have for each group a system of linear equations. Its solution can be shown to be:

$$p_{ij} = A_i \cdot (\gamma_i+\beta_0+\beta \cdot g_{ij}) + B_i \cdot \Sigma'(\gamma_i+\beta_0+\beta \cdot g_{ik}), \tag{4.3}$$

where $A_i = \frac{1}{1—c_i} \cdot \frac{(n_i—2)c_i+1}{(n_i—1)c_i+1}$ and $B_i = —\frac{c_i}{1—c_i} \cdot \frac{1}{(n_i—1)c_i+1}$ and here n_i is the number of unremoved clones in group i. Let now the G-value be fixed. We then have $\Sigma\Sigma\, g_{ij}p_{ij} = G$ and $\Sigma\Sigma\, p_{ij} = 1$ (summation over all clones). Since the p-values are known for the removed clones, we also know the values of these two latter sums when the summation is taken over the unremoved clones only. Using then (4.3), we get two (linear) equations that determine β_0 and β. Solving these equations and inserting the values of β_0 and β into (4.3), we get the desired p-values. The iterative procedure stops if all the inequalities $0 \leqq p_{ij} \leqq b_{ij}$ are satisifed. Otherwise the clone C_{ij} for which $\max(—p_{ij}, p_{ij}—b_{ij})$ has its largest value is removed and the procedure is repeated.

5. Discussion

No groups

Some extra work is needed to apply our method instead of truncation selection. The increase in mean genetic value in a general situation has been studied by LINDGREN (1991). The advantage of our method is considerable when the selection is weak as it usually is when genetic thinning is applied to seed orchards.

For such a thinning, spacing is a concern besides the breeding values. Actually, optimal thinning may also be more practical than truncation selection. For spacing reasons, the manager may want to keep some ramets of non-superior clones. The method permits that.

Related works

CROW and KIMURA (1979) proved that truncation selection will result in a higher mean genetic value than any broken line with the same average fitness. We think that the factor they kept constant is not the most relevant one when comparing the alternatives.

Toro and Nieto (1984) addressed the problem of minimizing genetic drift in a breeding program. They searched for optimal unequal contributions of selected individuals to the next generation under the constraint of a constant selection intensity. They made no analytical derivations but applied quadratic programming to get numeric solutions. Their approach should give the same numerical results as ours in Section 3, as in principle the assumptions are equivalent, though differently expressed. Using simulations, Toro and Nieto found that if different expected contributions are used in a breeding programme, then the long term inbreeding effect will also be smaller than for truncation selection. Evidently, this is an advantage.

Kang and Namkoong (1988) showed that there are situations where, for a given selection intensity, truncation selection generates a smaller inbreeding effective population size than other selective breeding schemes. They did not try to optimize. Somewhat mysteriously, Kang (1989) found by numerical experiments that for breeding values following the normal distribution increasing linear weighting functions did not lead to improved selection. The mystery disappears if his weighting functions are compared to the optimal functions of the form $\beta \cdot (g—a)\ \varphi(g)$, where φ is the standard normal density. These latter functions are decreasing in the right tail.

Groups

We have studied diversity when there is a resemblance within groups of clones. There is no well-established genetic theory for expressing "effective number" or "diversity" when there are clones with different degree of similarity as far as origin is concerned. In spite of that this is a common situation. The theory is better established for the case that there is a single population with a family structure (Falconer, 1981, Chapter 9).

We interpret the similarities as correlations, cf. the beginning of Section 4. Letting the groups be families, we should then have $c_i = 1$ if all the clones within a group are identical, $c_i \approx 0.5$ if they are full sibs, and $c_i \approx 0.25$ if they are half-sibs. Thus the optimization method can easily be implemented in practice when the clones in a seed orchard can be arranged in family groups.

It ought to be possible to assign c-values also to other types of groups than families, like the different origins in the Maglehem case. The c-values above for families indicate that these values should be kept rather low (unless there is coancestry). The c-values could possibly be inferred from measured values of a lot of characteristics (maybe on a molecular level!) that might be relevant from the damage point of view or the inbreeding point of view. At least the first author of this paper, who prefers to think in terms of future damages, means that the within group correlations for the g-values should not be used.

We suggest that optimal thinning (with separate lines for the groups) is practically applied on seed orchards where for some reason the clones are grouped.

Acknowledgement

We thank Per Lindström for information about packages for optimization and Bertil Matérn and a referee for comments on the text.

References

Bondesson, L.: Mathematical principles of optimal utilization of clones under diversity constraints. Swedish University of Agricultural Sciences, Section of Forest Biometry. Report 27 (1989). —— Crow, J. F. and Kimura, K.: Efficiency of truncation selection. Proc Natl. Acad. Sci., USA 76: 396—399 (1979). —— Falconer, D. S.: Introduction to Quantitative Genetics. 3rd ed. Longman, Essex (1989) —— Kang, H.: Inbreeding effective population size under some artifical selection schemes. II. Normal distribution of breeding values. Forest Science 35: 303—318 (1989). —— Kang, H. and Namkoong, G.: Inbreeding effective population size under some artificial selection schemes. I. Linear distribution of breeding values. Theor. Appl. Genet. 75: 333—339 (1988). —— Karlsson, B. and Danell, Ö.: Genetic parameters, predicted breeding values and potential gains for clones in a Norway spruce seed orchard. In: Stener, L.-G. and Werner, M. (eds): Norway spruce; provenance, breeding and genetic conservation. The Institute for Forest Improvement, Uppsala. Report 11: 90—112 (1989). —— Lindgren, D.: Optimal utilization of genetic resources. Forest Tree Improvement 23: 49—67 (1991). —— Lindgren, D., Libby, W. J. and Bondesson, F. L.: Deployment to plantations of numbers and proportions of clones with special emphasis on maximizing gain at a constant diversity. Theor. Appl. Genet. 77: 825—831 (1989). —— Lindgren, D. and Matheson, A. C.: Increasing the genetic quality of seed from seed orchards by using the better clones in higher proportions. Silvae Genetica 35: 173—177 (1986). —— Lindgren, D. and Werner, M.: Gain generating efficiency of different Norway spruce seed orchard designs. (Including an appendix by Ö Danell). In Stener, L.-G, and Werner, M. (eds): Norway spruce; provenances, breeding and genetic conservation. The Institute for Forest Improvement, Uppsala. Report 11: 189—206 (1989). —— Luenberger, D. G.: Introduction to Linear and Non-linear Programming. Addison-Wesley, Reading, Massachusets (1973). —— Pielou, E. C.: Diversity indices. In: Encyclopedia of Statistical Sciences, Vol. 2. Wiley, New York (1982). —— Robertson, A.: Inbreeding in artificial selection programmes. Genet Res., Camb. 2: 189—194 (1961). —— Toro, M. A. and Nieto, B. M.: A simple method for increasing the response to artificial selection. Genet. Res. Camb. 44: 347—349 (1984).

Improvement of Larch Micropropagation by Induced Short Shoot Elongation in Vitro

By U. Kretzschmar

Federal Research Centre for Forestry and Forest Products,
Institute of Forest Tree Breeding, D-15377 Waldsievers-
dorf, Germany

(Received 29th June 1992)

Abstract

In vitro shoot propagation in larch using the ability of axillary buds to elongate and form long shoots is mainly limited by occurrence of short shoot formation. To increase the rate of shoot propagation a stem elongating treatment was used on short shoots from hybrid larch clones. For in-