Juvenile Selection in Tree Breeding: Some Mathematical Models

By H. Kang

North Central Forest Experiment Station,
Forestry Sciences Laboratory, P.O. Box 898, Rhinelander,
Wisconsin 54501, USA

(Received 14th September 1984)

Abstract

Four different juvenile selection models are developed and evaluated. When the heritability is assumed to be constant for all ages and the coefficient of juvenile mature correlation function is set at 308 (LAMBERT, 1980), the optimum selection age varied between 1.3 and 21.5 for the rotation age of 25 years.

The large differences in the optimum selection age were due to the differences in the models and the evaluation criteria used. Important findings of the models are as follows. 1) Basic model: When linear functions are used to represent the heritability ratio (g = h_2/h_3) and the genetic juvenile mature correlation (r), the intercept of g must be equal or greater than that of r to guarantee that the early selections are always beneficial. When a linear function and a log function were used to represent g and r (= 1 + Blnx), respectively, g did not influence the time of maximum annual genetic gain (x_{max}) significantly up to B values around 0.3. 2) Extended time models: These models delayed the x_{max} of the basic model. 3) Maximum future net worth model: The early selection x_{max} is independent of the rotation age (T), and an economic criterion k was the prime factor which influenced x_{max}. 4) Multiple cycle model: Under many simplifying assumptions this model reduced to the basic model.

It was concluded that extremely early selection time such as earlier than 1/3 of rotation age should be used with caution. Other shortcomings of the models were discussed.

Key words: Accelerated tree breeding, juvenile selection, heritability ratio, juvenile-mature correlation, tree breeding.

Zusammenfassung

Es wurden vier verschiedene Methoden zur Frühelektion entwickelt und ausgewertet. Unter gewissen Voraus-
In this paper, I will develop and evaluate various models of early selection. The objectives are 1) to define some elementary models which can be of further use, 2) to demonstrate that an array of selection ages can be obtained from different models but the answers do not converge to a narrow range of ages, and 3) to discuss the effects which influence optimum selection ages within and between the models.

The models to be discussed are not prediction models. They are rather simple, and assumptions used are unrealistic. The models are primarily useful for learning the dynamics of optimum selection age under various conditions. It is possible the models can be further improved and used for predictions in the future. Such application will however, require substantially more information on the changing nature of the parameters of the models. I will limit the discussion to a basic model and three modifications of the basic model.

**Basic Model**

One popular way to examine the question of optimum selection age has been to use genetic gain per year (annual genetic gain) as the measure for comparing selection at different ages (Dickerson and Haze, 1944). Let the annual genetic gain be \( G^*_t \), then

\[
\Delta G^*_t = \Delta G^*_t / t,
\]

where \( \Delta G^*_t \) is the standardized selection intensity, \( \Delta G^*_t \) and \( h^2_t \) are heritabilities at age \( t \) and maturity (T), respectively, \( \rho(t) \) is the genetic correlation between \( t \) and \( T \), and \( \sigma^2_T \) is the phenotypic standard deviation at \( T \).

By comparing the annual genetic gain \( (\Delta G^*_t) \) at different times, one can empirically determine the selection age at which \( \Delta G^*_t \) is maximum. If the information on the genetic juvenile maturity correlation \( \rho(t) \) and heritability \( (h^2_T) \) for all different time was available, then finding an accurate selection age with maximum \( \Delta G^*_t \) would be relatively simple matter. Currently, only limited information on such parameters is available.

Even if sufficient information is available learning the dynamic behavior of the optimum selection age under different models can be enhanced if the genetic correlation and heritability are expressed as functions of time. If we assume that such time dependent expressions exist and let

\[
g = g(t), \quad h^2(t), \quad \text{and} \quad \rho(t)
\]

represent \( g(t), h^2(t), \) and \( \rho(t) \) at time \( t \) \( \times \) \( f(t) \), then (1) can be rewritten as

\[
\Delta G^*_t = \Delta G^*_t h^2(t) / \sigma^2_T
\]

where \( Q = h^2_T \sigma^2_T / T \). Equation 2 will be referred to as the basic model in this paper. Note that \( \Delta G^*_t \) is replaced by \( \Delta G^*_t \). The assumptions of the model are as follows. Selections made at different ages are drawn from standard normal populations at different time, and the population size will remain constant over different selection periods. If the breeder uses a fixed proportion of individuals for all the selections the selection intensity \( i \) will remain constant. Only a single cycle of selection will be applied such that \( h^2_T \) will remain constant. These assumptions assure that \( Q = h^2_T \sigma^2_T / T \) is a positive constant, and \( \Delta G^*_t \) is maximum at \( x \) which maximizes the function \( y = g(x) \).

Clearly, \( y \) is a critical parameter indicating the relative merits of selecting at \( t < T \). \( y < 1 \) implies no advantage to early selection. Since \( y \) is a variable function of \( g, r \) and \( x \), its behavior reveals the effects of those parameters on the efficiency of early selection.

The first and second derivatives of \( y \) are:

\[
y' = \begin{cases} \log(\sqrt{g}) + \frac{x}{2} + \log(x) & \text{if } y > 0 \\ \log(\sqrt{g}) + \frac{x}{2} - \log(x) & \text{if } y < 0 \end{cases}
\]

\[
y'' = \begin{cases} \frac{1}{g} \left[ 2 \log(\sqrt{g}) + x \right] & \text{if } y > 0 \\ \frac{1}{g} \left[ 2 \log(\sqrt{g}) - x \right] & \text{if } y < 0 \end{cases}
\]

If we set \( y = 0 \), then

\[
y' = \frac{1}{g} \left[ 2 \log(\sqrt{g}) + x \right]
\]

One of the roots of (5) will be the maximum provided that \( y' \) at the \( x \) is less than zero. The root will be referred to as \( x_{\text{max}} \). The \( x_{\text{max}} \) will vary depending on the nature of the population which governs \( g \) and \( r \).

Although \( g \) and \( r \) depend on the populations, efforts were made to describe some general behavior of the heritability and genetic juvenile maturity correlation as time related functions. Namkong and Conkie (1976) and Franklin (1979) indicated that the heritability can change substantially among three different phases of tree growing processes (i.e., juvenile genotypic phase, mature phase, and codominance suppression phase). Although the classification did not generate any explicit time related heritability function, their idea could be viewed as a step toward learning such functions.

Lambeth (1980) suggested to use,

\[
r(x) = B_0 + B_1 \ln x
\]

where \( x = t / T \).

\[
r(x) = \text{genetic juvenile maturity correlation between } t \text{ and } T, \text{ and } B_0, B_1 \text{ are constants representing intercept with } x = 1 \text{ and regression coefficients of } r(x) \text{ on } \ln x, \text{ respectively. The log function by Lambeth seems to apply to many pine species and will be used extensively in this paper with the following modification.}
\]

\[
r(x) = 1 + B_2 \exp(-1/B_1) x + B_3 x
\]

\( B_2 = 1 \) is used because \( B_2 \) values other than 1 will result in discontinuity of \( r \) at \( x = 1 \). The domain of \( x, \exp(-1/B_1) \leq x \leq 1 \) will apply to all the pertinent equations in this paper unless specified differently.

In the remainder of this section I will use two sets of arbitrary functions to represent \( g \) and \( r \), and discuss some properties of \( x_{\text{max}} \).

**Case I.** Both \( g \) and \( r \) are linear functions:

If we assume that both heritability ratio \( (g) \) and genetic juvenile-mature correlation \( (r) \) are linear functions of time, we may write:

\[
g = A + (1-A)x, \quad \text{and} \quad r = D + (1-D)x, \quad \text{with the restrictions that } g = 1 \text{ at } x = 1, \quad \text{and} \quad 0 < x < 1, A \quad \text{and} \quad D \text{ are parameters that have to be estimated from actual data. Some of the data collected by Franklin (1979), such as from Sulllack and Ganse (1974), indicate that it is possible to approximate } g \text{ by a linear function depending on the definition of the rotation age. A linear function of } r \text{ seems to be a fairly good approximation of } (5) \text{ when } 0.5 \leq x \leq 1 \text{ (Figure 1). Using the linear functions.}
shows the value of $A$ necessary to cause minima at the specified values of $x$ and $D$. The most important special case of (8) is when $x = 1$. In this case any early selection will be guaranteed to generate $y$ greater or equal to 1. Then (8) reduces to

$$A = 1 - D.$$
Table 2. — Values of A which will guarantee the extreme values of y at given x and B.

<table>
<thead>
<tr>
<th>x</th>
<th>.05</th>
<th>.15</th>
<th>.25</th>
<th>.35</th>
<th>.45</th>
<th>.55</th>
<th>.65</th>
<th>.75</th>
<th>.85</th>
<th>.95</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.1</td>
<td>.008</td>
<td>.014</td>
<td>.021</td>
<td>.027</td>
<td>.032</td>
<td>.037</td>
<td>.042</td>
<td>.047</td>
<td>.052</td>
<td>.056</td>
<td>.061</td>
</tr>
<tr>
<td>.2</td>
<td>.048</td>
<td>.056</td>
<td>.064</td>
<td>.071</td>
<td>.077</td>
<td>.084</td>
<td>.091</td>
<td>.097</td>
<td>.106</td>
<td>.115</td>
<td>.123</td>
</tr>
<tr>
<td>.3</td>
<td>.765</td>
<td>.756</td>
<td>.717</td>
<td>.707</td>
<td>.697</td>
<td>.687</td>
<td>.677</td>
<td>.666</td>
<td>.655</td>
<td>.644</td>
<td>.633</td>
</tr>
<tr>
<td>.4</td>
<td>1.208</td>
<td>.690</td>
<td>.503</td>
<td>.437</td>
<td>.407</td>
<td>.391</td>
<td>.383</td>
<td>.379</td>
<td>.378</td>
<td>.378</td>
<td>.388</td>
</tr>
<tr>
<td>.5</td>
<td>1.124</td>
<td>1.163</td>
<td>.826</td>
<td>.691</td>
<td>.620</td>
<td>.576</td>
<td>.551</td>
<td>.533</td>
<td>.527</td>
<td>.521</td>
<td>.517</td>
</tr>
<tr>
<td>.6</td>
<td>2.66</td>
<td>2.414</td>
<td>1.956</td>
<td>.889</td>
<td>.744</td>
<td>.693</td>
<td>.643</td>
<td>.616</td>
<td>.587</td>
<td>.560</td>
<td>.531</td>
</tr>
<tr>
<td>.7</td>
<td>5.621</td>
<td>2.124</td>
<td>1.445</td>
<td>1.159</td>
<td>1.003</td>
<td>.907</td>
<td>.796</td>
<td>.736</td>
<td>.681</td>
<td>.635</td>
<td>.592</td>
</tr>
<tr>
<td>.8</td>
<td>8.795</td>
<td>2.721</td>
<td>1.769</td>
<td>1.385</td>
<td>1.188</td>
<td>.968</td>
<td>.861</td>
<td>.762</td>
<td>.675</td>
<td>.596</td>
<td>.529</td>
</tr>
<tr>
<td>.9</td>
<td>8.692</td>
<td>2.996</td>
<td>1.968</td>
<td>1.540</td>
<td>1.183</td>
<td>.983</td>
<td>.933</td>
<td>.878</td>
<td>.823</td>
<td>.772</td>
<td>.723</td>
</tr>
</tbody>
</table>

1) Boxed in values represent A which will guarantee maximum y. Values above the box represent A which will guarantee minimum y.

The necessary values of A = B for various values of x are shown in the following table.

<table>
<thead>
<tr>
<th>x</th>
<th>.1</th>
<th>.2</th>
<th>.3</th>
<th>.4</th>
<th>.5</th>
<th>.6</th>
<th>.7</th>
<th>.8</th>
<th>.9</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = B</td>
<td>≤ .323</td>
<td>.371</td>
<td>.402</td>
<td>.425</td>
<td>.443</td>
<td>.458</td>
<td>.478</td>
<td>.491</td>
<td>.5</td>
<td></td>
</tr>
</tbody>
</table>

When x = 1, Equation 14 is undefined. However, it can be shown that \( \lim_{x \to 1} B = \frac{1}{2} \) by repeated application of l'Hôpital's rule.

When B is greater than \( .5 \) or the value specified in the above table, A = B + m must be greater than B, or m > 0. By solving (13) for A = B + m we obtain,

\[
A = B + m = \frac{x}{1 - x} \frac{\text{Binx}}{1 + \text{Binx}}, \exp(-1/B) \leq x < 1.
\]

Figure 2. — Linear function \( g(x) = \frac{h(t)}{h_p} \) necessary to guarantee the presence of maximum y at specified \( x_{\text{max}} \) values. Note the drastic changes in A Values to accomplish the change of \( x_{\text{max}} \) by 0.05.

2a) B = .3: When \( x_{\text{max}} = 2 \), A = .765 and \( y = 2.628 \),
When \( x_{\text{max}} = .5 \), A = .528 and \( y = .744 \).
2b) B = .4: When \( x_{\text{max}} = 2 \), A = 2.008 and \( y = 2.503 \),
When \( x_{\text{max}} = .5 \), A = .967 and \( y = 1.364 \).

Table 3 shows the values of A necessary to guarantee \( y = 1 \) for various values of x and B. By comparing Tables 2 and 3 we can see that \( y = 1 \) exists in the earlier regions of x although \( x_{\text{max}} \) does not exist in the same area. In general A needed for \( y = 1 \) is smaller than that for \( x_{\text{max}} \) except for the boxed-in values (Figure 3). At these points y values are less than 1 at \( x_{\text{max}} \).

A special situation of Case 2 arises when A = 1. Table 1 showed that A has relatively little influence on \( x_{\text{max}} \) for B values up to around 3. In this case, assuming \( A = 1 \) simplifies the analysis substantially. Therefore,

\[
y = r/x = (1 + \text{Binx})/x, \quad y = (B - 1 - \text{Binx})/x.
\]
Table 3. — Values of A which will guarantee that \( y = 1 \) at given \( x \) and \( B \).

| \( x \) | .05 | .1 | .15 | .2 | .25 | .3 | .35 | .4 | .45 | .5 | .55 | .6 | .65 | .7 | .8 | .9 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| .1 | .022 | .033 | .041 | .048 | .056 | .059 | .063 | .067 | .071 | .074 | .078 | .081 | .084 | .086 | .089 | .096 |
| .2 | .079 | .095 | .108 | .119 | .128 | .136 | .143 | .150 | .155 | .161 | .166 | .171 | .175 | .179 | .182 | .194 |
| .3 | .467 | .248 | .234 | .223 | .237 | .242 | .248 | .253 | .258 | .263 | .267 | .271 | .276 | .280 | .287 | .294 |
| .5 | 2.524 | 1.200 | .753 | .548 | .564 | .544 | .530 | .521 | .515 | .510 | .506 | .502 | .505 | .507 | .508 | .509 |
| .6 | 7.021 | 1.640 | 1.115 | .916 | .814 | .753 | .712 | .684 | .663 | .647 | .635 | .628 | .621 | .615 | .608 | .607 |
| .7 | 10.949 | 2.297 | 1.493 | 1.192 | 1.037 | .943 | .805 | .735 | .676 | .630 | .588 | .561 | .545 | .530 | .516 | .509 |
| .8 | Negative values |
| .9 | 9.223 | 3.136 | 2.059 | 1.658 | 1.424 | 1.276 | 1.176 | 1.103 | 1.095 | .943 | .9 |
| 1 | 7.297 | 3.242 | 2.259 | 1.817 | 1.566 | 1.405 | 1.294 | 1.149 | 1.060 |

1) The boxes represent the combinations of \( x \) and \( B \) at which \( y \) values at \( x_{max} \) is less than 1.

The \( x_{max} \) occurs when

\[ x = \exp(1 - 1/B). \]

The values of \( x \) at \( B = .3 \) are shown in Table 4 for different rotation ages, and that at various \( B \) are shown in Table 1 on the bottom row with \( A = 1 \). An interesting feature of this special case is that the genetic juvenile-maturity correlation \( r \) is identical to \( B \) at \( x_{max} \). This can be shown by replacing \( x_{max} \) in (6) such that

\[ r(x_{max}) = 1 + B \ln(x_{max}) \]

\[ = 1 + B \ln[\exp(1 - 1/B)] \]

\[ = B. \]

Extended Time Model

In the basic model, \( g \) and \( r \) were either linear or logarithmic functions. In this section we will let \( g = h^2/T \), \( r = 1 + B \ln x \), and the denominator of the basic model (Equation 2) to vary. The extended time model is expressed as

\[ y = \frac{g(t)}{g(t) + k} \]

\[ = \frac{1 + B \ln x}{1 + B \ln x + k} \]

where \( k = x/B \), \( g = \ln x/T \), and \( y \) represents the actual amount of time added.

By differentiating \( y \) with respect to \( x \) and equating the derivative to zero we get,

\[ \ln x = -k/1 - 1/B, \]

or

\[ \ln x = \ln x + 1 = 1/B. \]

By changing the \( k \) values a wide variety of situations can be examined. Two cases of interest are as follow.

Case 1: LAMBETH (1980) extended \( T \) and \( t \) to \( T + 3 \) and \( t + 5 \), respectively. Such extension was necessary to accommodate time needed to make the regeneration, such as flower induction. His scheme can be expressed as,

\[ y = (1 + B \ln x)/(x + 5/T). \]

Note that \( T + 3 \) is a constant and can be incorporated into \( Q \) such that \( Q^* = h^2\ln(T + 3)/T \). The replacement of \( k = 5 \) in Equation 15 will generate,

\[ \ln x = -5/xT = 1 - 1/B. \]

The above equation is numerically evaluated for \( B = .3 \) (Table 4).

Case 2. Two-cycle model:

Another special case of interest is when \( k = T \). This represents the situation where the realized genetic gain in the subsequent production population is divided by the entire period of breeding and production. In this model the regeneration time, such as in Case 1, is ignored. Then the model becomes

\[ y = (1 + B \ln x)/(x + 1), \]

and the roots of

\[ \ln x = -1 + 1/B \]

will generate the desired \( x_{max} \). By using the first three terms of Taylor's series expansion about \( x = .5 \), (16) may be approximated as

\[ 20x^2 - 32x + 18.386 - 2/B = 0, B \leq .358 \]

with the solution \( x = .8 - .025 \times 160/B = 446.88 \). However, the approximation has a narrow range of application, say \( 2 \leq B \leq .31 \). For \( B \) values outside the above range the root of (16) is numerically evaluated (Table 4).

When \( B = .3 \), (16) becomes

\[ xln(x) + x - 1 = 0, \]

with the solution \( x = 1 \). Such evaluations show that, when compared to the basic model, the two-cycle model lengthens the regeneration time.

![Figure 3. — Curves of \( r = 1 + B \ln x \) with various values of \( B \). The values of \( r \) at \( x_{max} \) (starred points) are: .0, .1, .2, .3, .4, and .5 for points 1, 2, 3, 4, and 5, respectively.](image)

Table 4. — Comparison of the basic model \( (g(x) = 1) \) and extended models when \( B = .3 \).

<table>
<thead>
<tr>
<th>( T )</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>( x_{max} )</td>
<td>.077</td>
<td>.097</td>
<td>.097</td>
<td>0.097</td>
<td>.091</td>
<td>.097</td>
<td>.097</td>
</tr>
<tr>
<td>Model</td>
<td>( x_{max} )</td>
<td>1.94</td>
<td>2.92</td>
<td>3.91</td>
<td>3.39</td>
<td>3.45</td>
<td>3.55</td>
<td>5.7</td>
</tr>
<tr>
<td>LAMBETH</td>
<td>( x_{max} )</td>
<td>.256</td>
<td>.271</td>
<td>.212</td>
<td>.209</td>
<td>.188</td>
<td>.173</td>
<td>.129</td>
</tr>
<tr>
<td>model</td>
<td>( x_{max} )</td>
<td>5.12</td>
<td>7.78</td>
<td>6.36</td>
<td>6.67</td>
<td>7.52</td>
<td>8.1</td>
<td>8.65</td>
</tr>
<tr>
<td>Two</td>
<td>( x_{max} )</td>
<td>.558</td>
<td>.558</td>
<td>.556</td>
<td>.558</td>
<td>.556</td>
<td>.558</td>
<td>.556</td>
</tr>
<tr>
<td>cycle</td>
<td>model</td>
<td>11.36</td>
<td>14.2</td>
<td>17.04</td>
<td>19.98</td>
<td>22.72</td>
<td>25.56</td>
<td>28.4</td>
</tr>
</tbody>
</table>

1) \( x_{max} \) values of the two cycle model at different \( B \) values are .142, .33, .795, and 1 for \( B \) values of .1, .2, .3, and .5, respectively.
ens the \( x_{\text{max}} \) substantially (Table 4). When \( B = .3 \), the \( x_{\text{max}} \) is longer than 1/2 of the rotation. Furthermore, when \( B \geq .5 \), juvenile selection loses its advantage completely. The coefficient \( B \) for the loblolly pine example of Franklin (1979, Table 3) is .432. At this \( B \) value, \( x_{\text{max}} = .867 \).

Unless the juvenile-mature correlations are so high that \( B \) is less than .2, the early selection age suggested in the basic model will not be realized in the two-cycle model. Conceptually, it is possible to shorten the age when the annual genetic gain is maximum by reducing the rotation length (T). Such reduction in rotation length would tend to increase the genetic correlation between an earlier juvenile age and the rotation time. The reduction in rotation time, however, may not change \( B \) (Lambrin, 1980). Therefore the impact of reducing the rotation length is to reduce \( t_{\text{max}} \) but not \( x_{\text{max}} \).

**Maximum Future Net Worth Model (MFN)**

The basic model and extended time models use different practical parameters as the denominator for determining the annual genetic gain (\( \Delta G^* \) or \( \Delta G_{st}^* \)). The annual genetic gain, however, does not always generate practically meaningful answers. Consider, for example, the basic model with \( g(x) = 1 \). As was discussed before \( r(x) \) at \( x_{\text{max}} \) equals \( B \) (Figure 3), which means that the expected genetic gain at \( x_{\text{max}} \) (\( \Delta G_{\text{max}} \)) is only \( B \) fraction of that obtainable at the rotation age (i.e., \( x = \Delta G_{\text{max}} = B \Delta G_{\text{r}} \)). For small \( B \), \( \Delta G_{\text{r}} \) is maximum at \( x_{\text{max}} \) mainly because of the extremely small \( x \) used as the denominator. An extreme situation can be seen in Figure 3, when \( B = .1 \). The \( x_{\text{max}} = .0001 \) and \( r(x_{\text{max}}) = .1 \), meaning \( \Delta G_{\text{max}} = \Delta G_{\text{r}} / 10 \). However, the \( x \) at which \( \Delta G_{\text{r}} = .0001 \) is .0007 and that at which \( \Delta G_{\text{r}} = .75 \Delta G_{\text{r}} \) is .083. All the \( x \) values are very small, and selecting at \( x = .0001 \) to maximize \( \Delta G_{\text{r}} \) is a nonsense. The situation improves as \( B \) increases. When \( B = .3 \), the \( x \) values are 0.007 (\( x_{\text{max}} \)), 0.189, and 0.435 for (3/10)\( \Delta G_{\text{r}} \), 0.5\( \Delta G_{\text{r}} \), and 0.75\( \Delta G_{\text{r}} \), respectively. Nevertheless, the example demonstrates the need to use the concept of annual genetic gain with caution.

One way of avoiding such potential complication is to use the expected genetic gain at different times (\( \Delta G_{\text{r}} \) or \( \Delta G_{st}^* \)) directly, by incorporating it into an economic model. In this section the future net worth equation is utilized for that purpose.

Future net worth (FNW) is defined as follows (Lundgren, 1973):

\[
\text{FNW} = P \cdot V_t - c(1 + R)^t
\]

where \( P_t \) and \( V_t \) represent the price and quantity (or volume) of the products at time \( t \), \( R \) represents the interest rate, and \( c \) represents the initial cost. The FNW equation can be modified to incorporate the genetic gain at different period as follows.

\[
\text{FNW} = (1 + f(\Delta G^*_{\text{avg}})) V_t - c(1 + R)^t
\]

where \( c \) is replaced by \( c \cdot f(\Delta G^*_{\text{avg}}) \).

\( f(\Delta G^*_{\text{avg}}) \) is an unknown function of expected genetic gain \( \Delta G^*_{\text{avg}} \).

The assumption are: 1) The time between the selection and seed production is ignored. 2) \( f(\Delta G^*_{\text{avg}}) = \Delta G_{\text{avg}} \), which implies that the price of the improved seeds is proportional to the expected genetic gain at different time. 3) All the seeds produced are utilized, and the quantity of seed production is constant for the different selection ages. Then we can assume \( V_t = V_T \).

4) The subsequent operation cost (including the seed production) is constant for all different selection times.

By replacing \( f(\Delta G^*_{\text{avg}}) = |\ln f(\Delta G_{\text{avg}})| \cdot \ln(1/R) \) in (17) we get

\[
\text{FNW} = (1 + f(\Delta G^*_{\text{avg}})) V_t - c(1 + R)^t
\]

where \( c = c \cdot f(\Delta G^*_{\text{avg}}) \).

If we assume that \( g(x) = 1 \), and \( r(x) = 1 + \Delta G_{\text{r}} x \)

\[
\text{FNW} = (1 + f(\Delta G^*_{\text{avg}})) V_t - c(1 + R)^t
\]

Differenitizing \( x \) with respect to \( x \), and equating the resulting derivative to zero yields

\[
x(1 + R) = \frac{c}{(1 + R)^2} \]

Since \( x = 1/R \), (18) may be rewritten as

\[
x = \frac{c}{(1 + R)^2} = \frac{1}{R} \frac{c}{(1 + R)^2}
\]

From (19), we can see that \( x_{\text{max}} \) no longer depends on \( R \), the rotation age. Approximate solutions to (19) may be obtained by using the first two or three terms of Taylor's series expansion of \( x_{\text{max}} \) about \( \Delta G_{\text{r}} \). The use of the first three terms of the Taylor's series expansion yields

\[
x_{\text{max}} = \frac{1}{2} \left( 1 + \sqrt{1 + 4 \Delta G_{\text{r}} / 3} \right)
\]

where \( \Delta G_{\text{r}} = (1 + R)^2 / 2c 

Equation 20 tends to overestimate \( x_{\text{max}} \), but is fairly useful for \( k \geq 1 \). From Table 5, we can see that there are certain combinations of \( B \), \( R \), and \( k \) under which juvenile selection is not beneficial (for example, \( B = .3 \), \( R = .05 \), \( k = .1 \), and \( T = 20 \)), where \( k \) represents the ratio between the initial cost and the return at the regular rotation age \( T \).

The most important findings of MFN include, 1) the \( x_{\text{max}} \) is determined by both biological and non-biological factors, and 2) it might be easier to work with non-biological factors to adjust \( x_{\text{max}} \). The combined influence of \( k \) and \( B \) on the \( x_{\text{max}} \) is in the form:

\[
x = B / k
\]

where \( z \) is a combined factor influencing the \( x_{\text{max}} \). The explicit expression for \( z \) is,

\[
z = (1 + R)^2 / 2c
\]

and all the components in the numerator influence the amount of return than the cost. Therefore larger \( z \) values mean larger return/cost rate, which will retard the \( x_{\text{max}} \).

In equation (21), \( b_n, q_n \) and \( B \) are biological variables, and \( q_n, V_t \), and \( c \) are economic variables. In practice, it might be more practical to attempt to adjust the \( x_{\text{max}} \) by manipulating non-biological variables. For example, by increasing the selection intensity, the breeder can delay the \( x_{\text{max}} \). If advanced propagation methods, such as tissue culture, become practical, the \( V_t \) and \( h_{n \cdot r} \) will increase, and the return on the selection could peak at rotation age. Low initial investment as well as high cost of unimproved seeds will also lengthen the \( x_{\text{max}} \).

The \( x_{\text{max}} \) is also influenced by the interest rate (\( R \)). Equation 20 shows that higher interest rates will shorten the \( x_{\text{max}} \). To obtain \( x_{\text{max}} \) values under different interest rates, one may multiply \( \ln(1.05) / \ln(1 + R) \) to the \( x_{\text{max}} \) shown in Table 5.

The MFN model also suggests that early selection for the sake of an early return on investment only may not be
Table 5. — Values of $t_{\text{max}}$ for various $B$ values obtained from Equation 29.

<table>
<thead>
<tr>
<th>$B$</th>
<th>.1</th>
<th>.2</th>
<th>.3</th>
<th>.4</th>
<th>.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>.5</td>
<td>11.70</td>
<td>18.06</td>
<td>22.65</td>
<td>26.35</td>
<td>29.40</td>
</tr>
<tr>
<td>.25</td>
<td>10.17</td>
<td>13.14</td>
<td>16.23</td>
<td>19.53</td>
<td>22.06</td>
</tr>
<tr>
<td>.5</td>
<td>6.67</td>
<td>8.11</td>
<td>9.20</td>
<td>10.26</td>
<td>11.16</td>
</tr>
<tr>
<td>.75</td>
<td>3.43</td>
<td>4.41</td>
<td>5.66</td>
<td>6.80</td>
<td>7.66</td>
</tr>
<tr>
<td>.9</td>
<td>2.81</td>
<td>3.75</td>
<td>4.75</td>
<td>5.55</td>
<td>6.02</td>
</tr>
<tr>
<td>.1</td>
<td>1.87</td>
<td>3.46</td>
<td>4.86</td>
<td>5.11</td>
<td>7.24</td>
</tr>
</tbody>
</table>

1) $R = \text{interest rate,}$  
$k = c/(bV_{t+1})qV_T$,  
$c = \text{initial cost,}$  
$T = \text{rotation age,}$  
$B = \text{the slope of } t(x) \text{ in log scale.}$

always desirable. For example, assume that a breeder practiced selection at an early age to obtain an early return on investment. Also assume that he has to bear additional costs arising from early flower induction. In order to make up for the additional investment, MFN model suggests the selection be made sooner. The further shortened selection time will require even more investment for flower induction research and development. Obviously, this will not lead to a realistic maximum. In many tree breeding programs such as flower induction can be justified for reasons other than juvenile selection combined with mass propagation, and the investment may not be considered as an additional cost for making early selections.

Although MFN model avoids the problem of the basic model and extended time models, it has its own shortcomings. In addition to genetic assumptions such as $g(x) = 1$, and $r(x) = 1 - B \ln x$, the model requires many economic simplifying assumptions. It is unlikely that assumptions such as no time lag between selection and seed production, $V_t = V_T$, $h(\Delta G_t) = \Delta G_t$, and no subsequent operational cost difference between different selection times can be justified in the real world. The model is, however, valuable because it demonstrates the need for balancing economic factors and biological factors to determine the optimum selection time. It also indicates a direction of future development in economically realistic prediction models.

### Multiple Cycle Model

There are two approaches to juvenile selection in tree breeding (KANG, 1982). The first deals with improving the juvenile trait itself, and the performance of the individuals at a later stage is not important: Hence, juvenile-mature correlation is not a pertinent issue. This approach reflects the recent interest in short-rotation forestry, which tend to change the quality of the wood supply.

In the second approach, juvenile selection is considered as a technique that influences the tempo of overall breeding activities. For a given time, say 50 years, the tree breeder has a choice of different selection cycles. For example, jack pine flowers as early as 12 months (RUDOLPH, 1979), so the breeder may then ask how selection cycles, say 6, 4, 3, 2, and 1, will influence the total genetic gain at the end of 24 years (Figure 4).

Since most breeding techniques are applied to breeding stocks during selection and crossing time, the more generations occurring within a fixed period the more opportunities to apply breeding techniques will be. Therefore, many well established crop breeding techniques can be more readily applied to tree breeding if a short multi-cycle option is used. For example, in jack pine (or other species with the possibility of early flower induction) the development of inbred lines, with or without selection, can be achieved in less than 50 years, which is a typical single rotation time for many tree species.

In the multiple-cycle model $b_{rT}$ and $\sigma_T$ can no longer be considered fixed, because population parameters change over different selection cycles. Furthermore the distribution of $g(x)$ and $r(x)$ will vary over different generations, therefore, two types of time have to be considered: time within a single generation ($x$), and the number of generations ($n$). The general expression of the accumulative genetic gain for the multiple cycle model is:

$$
\alpha_{\text{max}}^n = 1 + \sum_{j=1}^{n} \frac{n-j}{j} g_j r_j (x) e_j
$$

where $j$ indexes the generation. Because $g_j(x)$ and $r_j(x)$ vary, $\alpha_{\text{max}}^n$ will also vary over different generation. As will be

![Figure 4](image-url)
assumed later, only when the distributions are considered fixed over generations the two different types of time, \( x \) and \( n \), will have interchangeable meanings.

Currently, little knowledge on the genetic causal components, such as the number of loci, linkage structure, and interaction among alleles etc., of quantitative traits of trees is available. Therefore, developing a prediction model of multiple-cycle selection is even more difficult than that in the previous models. It is, however, possible to develop theoretical models which can be used to understand the changes of the parameters over generations. These questions will be discussed further in another paper.

In this section the assumption that \( h^2_r, s^2_r, g(x), \) and \( r(x) \) do not vary over generations is made. The assumption will make it possible to connect \( x \) and \( n \) in an explicit form. Under the assumption Equation 22 can be rewritten as:

\[
\Delta G_{x} = nh^2_r g(x)(r(x))s^2_r r(x)
\]

If we further impose the restriction that different selection schemes will be evaluated at time \( T \), then the number of cycles (of effective generation) one can have during time \( T = T/t \) is given by:

\[
\Delta G_{x} = (1/x)nh^2_r g(x)(r(x))s^2_r r(x)
\]

where \( Q = h^2_r s^2_r r/T \) as was the case in Equation 2. Because both \( T \) and \( Q \) are constants, \( x_{\text{max}} \) can be obtained by finding \( x \) which maximizes \( y = g(x) \). Once \( x_{\text{max}} \) is found, the number of generations which maximizes \( \Delta G_{x} \) can be determined by the relation

\[
\Delta G_{x} = (1/x_{\text{max}})
\]

The changing nature of \( x_{\text{max}} \) has been discussed in the basic model and extended time models.

Another worthwhile question is: how many short cycles are necessary to obtain a level of genetic gain equivalent to that available with a single long cycle? This question is equivalent to asking for the \( x \) which guarantees \( y = 1 \) in the basic model. This question is useful if early selection is viewed as an activity that influences the tempo of overall tree breeding efforts. Because the question does not seek for \( x_{\text{max}} \), the answer is likely to point towards smaller \( x \) and larger \( n \). Larger \( n \) will generate more chances to apply different breeding techniques. In a qualitative sense, larger \( n \) might generate more overall gain from breeding than smaller \( n \) which maximizes \( \Delta G_{x} \). To obtain the answer we equate \( \Delta G_{x} \) with \( \Delta G_{t} \) and solve for:

\[
\Delta G_{x} = nh^2_r g(x)(r(x))s^2_r = h^2_r s^2_r r(x)
\]

\[
g(x)r(x) = 1
\]

\[
n = \frac{1}{g(x)r(x)}
\]

If we assume as before that \( g(x) = 1 \), and \( r(x) = 1 + Blnx \)

\[
n = \frac{1}{1 + (1 + Blnx)} \]

To determine \( n \) it is necessary to find the intermediate root of:

\[
T(1-x) + (T+r)Blnx = 0
\]

where \( s \) represents an arbitrary constant such as the time required for seed maturation. Note that \( x = 1 \) is a trivial solution. The replacement of the intermediate root in (23) will yield the desired \( n \).

Let \( n^* = (T+r)/(t+s) \), for species that require \( s \) years for seed maturation. If \( n^* > n \), repeated juvenile selection at age \( t \) will yield a greater genetic gain than selection at \( T \).

It is also possible to use \( B \) to compare different selection cycles. Let \( r^* = 1/n^* \), then \( n^* > n \) is equivalent to saying that \( r > r^* \), or

\[
1 + Blnx = (t+s)/(t+2), \quad \cdots \quad (24)
\]

for species with 2 years for seed maturation. By solving equation 24 for \( B \) we get

\[
B < B_{\text{min}} \quad \cdots \quad (25)
\]

where \( B_{\text{min}} = .662x/(x-1)/lnx \).

For species with 1 year for seed maturation, we get

\[
B_{\text{min}} = .939x/lnx \quad \cdots \quad (26)
\]

Table 6 shows \( B_{\text{min}} \) for various values of \( x \). By comparing the actual \( B \) values and the table values, we can determine at which age juvenile selection begins to be acceptable.

Table 7 shows \( r(x) \), \( n \), and \( n^* \) for lobolly pine at various ages (data from Frankling, 1979). Underlines represent situations where \( n^* > n \).

<table>
<thead>
<tr>
<th>t</th>
<th>n</th>
<th>n^*</th>
<th>n^*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year</td>
<td>367</td>
<td>365</td>
<td>345</td>
</tr>
<tr>
<td>2 years</td>
<td>345</td>
<td>345</td>
<td>345</td>
</tr>
</tbody>
</table>

Table 7. — Genetic correlation, \( r(t) \), \( n^* \), and \( n^* \) for lobolly pine at various ages (data from Frankling, 1979). Underlines represent situations where \( n^* > n \).

Table 6. — \( B_{\text{min}} \) at various values of \( x \).

<table>
<thead>
<tr>
<th>x</th>
<th>n</th>
<th>n^*</th>
<th>n^*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

LAMBERT (1980) used \( B = .308 \) for several Pinaceae species. At \( B = .308 \) juvenile selection at any time greater than or equal to 1/10 of the rotation time will be advantageous (Table 6). Therefore, in most practical situations, repeated juvenile selection will lead to genetic gains equivalent to single generation turnover at maturity. This finding, together with the increase in the number of opportunities to apply breeding techniques in multiple short cycles, make juvenile selection an attractive option.
Table 8. — Selection time (in years) suggested under various models when the rotation age is 25.

<table>
<thead>
<tr>
<th>Model</th>
<th>0.125</th>
<th>0.25</th>
<th>0.308</th>
</tr>
</thead>
<tbody>
<tr>
<td>Franklin (1978)</td>
<td>12.5</td>
<td>12.5</td>
<td>12.5</td>
</tr>
<tr>
<td>Basic model (x)</td>
<td>5.7</td>
<td>5.6</td>
<td>2.9</td>
</tr>
<tr>
<td>Basic model (x+1)</td>
<td>10.5</td>
<td>10.6</td>
<td>7.7</td>
</tr>
<tr>
<td>Lambeth (1980)</td>
<td>10.7</td>
<td>9.5</td>
<td>7.8</td>
</tr>
<tr>
<td>Two Cycle</td>
<td>21.5</td>
<td>19.8</td>
<td>17.0</td>
</tr>
<tr>
<td>MNF x=1/1</td>
<td>1.7</td>
<td>1.7</td>
<td>21.8</td>
</tr>
<tr>
<td>x+.5</td>
<td>10.5</td>
<td>10.1</td>
<td>5.5</td>
</tr>
<tr>
<td>x+1</td>
<td>3.0</td>
<td>3.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Multiple cycle (x+2)</td>
<td>3.2</td>
<td>3.2</td>
<td>1.3</td>
</tr>
</tbody>
</table>

1) The age suggested by the model is 27.4.
2) The age suggested by the model is 26.3.

Discussion

To determine a precise optimum selection age of a population one needs a rather detailed information on the performance of the population over a wide range of locations as well as time. When such information is present, it is not critical to develop models described in this paper to determine the optimum selection age. When such information is lacking, but general information for early selection is desired, the models could be useful, with the understanding that the models generate elementary information which could be of some value in decision making, but not prediction values. In this section I will discuss subjects which are useful for understanding the dynamics of t_max, and the components that influence optimal selection age. No efforts will be made to discuss any particular numbers generated from the models.

The models presented in this paper are by no means complete or realistic. To understand the validity of the models it is necessary to examine their shortcomings. We will use four different, but interdependent factors to examine the models. They are: 1) the available information on basic parameters, 2) the validity of approximating the parameters by using some time dependent functions, 3) the structure of the models, and 4) the nature of the evaluation criteria.

Information on Parameters: Two most important parameters are the heritability ratio, g, and the genetic juvenile-mature correlation, r. The information on both parameters is limited, but sufficient to develop simple models on single cycles of selection. Recent papers by Franklin (1979) and Lambeth (1980) summarized the available information. The information on g is not as well structured as that of r. Although the information on g and r come from similar sets of data base, the specific nature of r, which we will discuss later, makes it easier to incorporate it into models. The available information is not sufficient to make the early selection ages of different models to converge to a small range of numbers (Table 8). Note, however, that more information does not necessarily imply a convergence to early selection ages.

Further work on economics of early selection is necessary to generate information on economic parameters, such as k, as described in the MNF model. Such efforts will also generate economic models which are more realistic than MNF.

Little information on the nature of g, and r over different generations is available. It is unlikely that the information will be generated for a majority of conifer species in the near future. It is however, possible to obtain information from species such as jack pine and willows. Some information can be generated by theoretical means. This subject will be discussed in a separate paper.

Approximation of the Parameters: In this paper the heritability ratio, g, was approximated by a linear function of time. It is possible to use other functions to represent g and r. For example Franklin’s phase concept can be approximated by using a step function or a quadratic function of time. Such functions might influence the x_max to move toward the middle of x values rather than to the early part of x as was shown in the basic model. However, if x_max is the criterion of interest in the basic model, the nature of the function might not matter for r values up to somewhere around 0.3.

Furthermore Defining a function to approximate g also depends on the definition of the rotation age. If the mature stage in the sense of Franklin (1979) is considered as the rotation age, a linear approximation might be acceptable. If the codominance suppression phase is considered as the rotation age, a quadratic function would be more useful. If the changes in the heritability is not too drastic among different phases, the assumption that g = 1 might be considered as a good approximation.

The choice of the function to approximate the genetic juvenile mature correlation, r, does not depend on the definition of the rotation age. The closer the time is to the reference point the higher the correlation will tend to be, regardless of the time chosen as the reference point. This is why the information on r is structured better than that on g although both come from the similar date base.

A number of different functions may be taken to represent r. Both a linear function and a log function were used here. The linear function is not as descriptive as the log function, but the simplicity makes it useful. The linear function of Figure 1 was obtained by connecting two points, (75, 1 + Binx) and (1, 1) of the log function, y = 1 + Binx. In practice, however, data points between x = .5 and x = 1 may be used to fit a linear regression line which passes through (1, 1). It is possible to use a quadratic or cubic functions in place of the linear function. In that case the points between a smaller number, say x = .3, and x = 1 may be fitted, provided that the r tends to be high in general. The use of cubic function does not simplify the approximation. It carries too many parameters (i.e. coefficients). The log function as defined here carries only one parameter, B.

Of the functions that may be used to approximate r, the log function seems to offer the best results, but its usefulness for small x values is questionable. As x approaches zero the function drops quickly, and enters into the region of negative correlation (Figure 3). This problem is shared by other functions also, except that they migh overestimate the correlation. Regardless of the function used one should be careful about the interpretation of the values at small x. At this time it is not advisable to choose a small x as optimum selection age even if the model indicates to that direction. This is the area where the added information on r could make a difference. It can help to choose a function that is suitable for the approximation, and to determine how close to zero the function can be extended.

Structure of the Models and Evaluation Criteria: The two factors are treated together because they are very in-
Provenance Study of Douglas-Fir in the Pacific Northwest Region
IV. Field Performance at Age 25 Years*

By T. L. White and K. K. Ching**

(Received 17th September 1984)

Summary

Twenty-five-year results are reported for 14 Pacific Northwest provenances of Douglas-fir growing in 5 thinned, fast-growing plantations. In all plantations, the southern Oregon provenance was the poorest performer for all response variables (height, diameter at breast height, height: diameter ratio, volume per hectare, and site index). It produced 30 percent less volume than the average of the other provenances. Differences among the other 13 provenances were smaller and often not statistically significant. Provenance × plantation interactions were generally not significant when the southern Oregon source was omitted from the analyses. These results seemingly contradict experiments with seedlings and young plantations which indicate an extremely tight adaptation of Douglas-fir populations to the local environment in which they evolved. Of the several possible explanations for the apparent discrepancies, the most likely is a lack of experimental precision in the present study.

Key words: Douglas-fir, provenance, yield, genotype × environment interaction.

* Paper 1967 of the Forest Research Laboratory, Oregon State University, Corvallis, Oregon 97331.
** Supervisor, Western Research, International Paper Company, Portland, Oregon 97206, USA, and Professor Emeritus, Department of Forest Science, Oregon State University, Corvallis, Oregon 97331, USA.