Theor Appl Genet (1986) 73:53–60

# Theoretical studies on the necessary number of components in mixtures

## 4. Number of components and juvenile-mature correlations

**M. Hühn**

Institut für Pflanzenbau und Pflanzenzüchtung, Universität Kiel, Olshausenstrasse 40–60,
D-2300 Kiel, Federal Republic of Germany

Received April 19, 1986; Accepted June 30, 1986
Communicated by P. M. A. Tigerstedt

**Summary.** Theoretical studies on the necessary numbers of components in mixtures (for example multiclonal varieties or mixtures of lines) have been performed according to the relations between the juvenile-mature correlations of mixtures and their number of components. For the juvenile-mature correlation $r_E$ based upon the values of the single components (= component means at juvenile and mature ages) and the juvenile-mature correlation $r_M$ based upon the means of mixtures of different components we usually will have $r_M > r_E$. Furthermore, $r_M$ will increase with an increasing number of components in the mixtures. The effectiveness of an early selection will be mainly determined by the magnitude of the juvenile-mature correlation. If we have $r_M > r_E$ an improvement of early testing can be realized by using mixtures instead of single components. But, what are the necessary numbers of components so that $r_M$ will be sufficiently high to enable an effective early selection of mixtures? Some relations between $r_E$ and $r_M$ can be obtained and conclusions have been derived.

The statistical approach 'significant difference between $r_E$ and $r_M$ for a given numerical value of $r_M$' leads to estimates for the necessary number n of components dependent on $r_M$, $\alpha$, $r_E$ and N where: N = total number of components, which are available for the composition of mixtures and $\alpha$ = error probability. For different tree species $r_E$ can be estimated by an appropriate formula which depends on T with T = time (in years) from planting date until the mature age.

Lambeth's formula, for example, has been developed for height growth in pines. For this situation numerical calculations are performed using $r_M = 0.90$ and $\alpha = 0.05$. The necessary numbers n for $T = 5$, $T = 10$, $T = 20$ and $T = 50$ are: 6, 9, 10 and 12 (for N = 50); 13, 17, 20 and 23 (for N = 100); 26, 34, 40 and 46 (for N = 200); 38, 51, 60 and 69 (for N = 300); 64, 85, 100 and 114 (for N = 500) and 128, 171, 199 and 228 (for N = 1,000). The dependence of these necessary numbers n of components on different type I errors $\alpha$ and different levels of $r_M$ have been investigated numerically.

**Key words:** Mixtures – Number of components – Juvenile-mature correlations – Early selection

### Introduction and problem

An increasing interest in clonal forestry may be partly explained by essential improvements of methods of vegetative propagation in the last years – for example, by using cuttings and tissue culture techniques. Most forest tree breeders propose the development of multiclonal varieties (= mixtures of different clones artificially created with definite proportions) to maintain some genetic diversity in the stands. The following concepts and results have been worked out with regard to this field of applications, but they are also of an extended validity for agricultural crop science and plant breeding: 1) including multilines (= mixtures of isolines that differ by single, major genes for reaction to a pathogen), 2) including mixtures of an arbitrary number of pure lines which are more different among each other than isolines and 3) including mixtures of an arbitrary number of any components. To provide a simultaneous discussion of these situations – multilines, multiclonal varieties, mixtures of pure lines and mixtures of any other components – we use the general terms 'mixture' and 'components'.

In the three preceding papers of this series the problems of necessary numbers of components in mixtures have been discussed with respect to 1) phenotypic yield stability (Hühn 1985a), 2) yielding ability (Hühn 1986a) and 3) risk considerations (Hühn 1986b). The aim of this paper is to give some statistical approaches and numerical results concerning the necessary number of components in mixtures with regard to juvenile-mature correlations. In these studies we don't consider successive generations. Only one period of rotation from the initial composition of the mixture until the final harvest shall be analysed.

Juvenile-mature correlations, of course, are of particular interest for such long-living organisms as forest trees. Here, the effectiveness of an early selection will be mainly determined by the numerical value of the juvenile-mature correlation. Therefore, numerous investigations have been performed to estimate these juvenile-mature correlations for many species, for several characters, for different environmental conditions etc. (see, for example: Sziklai 1974; Nanson 1967, 1968, 1974, 1976).

In this paper we are interested in the relations between the juvenile-mature correlations of mixtures and their number of components. We denote the juvenile-mature correlation based upon the values of the single components (= component means at juvenile and mature ages) by $r_E$. Then, we may ask whether or not a low $r_E$ can be increased by using mixtures of components. Introducing $r_M$ = juvenile-mature correlation based upon the means of mixtures of different components we usually will have $r_M > r_E$. Some more precise conditions for $r_M > r_E$ are given in the next chapter "Theoretical investigations ...". Furthermore, $r_M$ will increase with an increasing number of components in the mixtures.

If we have $r_M > r_E$ an increase of the juvenile-mature correlation and, therefore, an improvement of the effectiveness of early testing can be obtained by using mixtures instead of single components.

But what are the necessary numbers of components so that $r_M$ will be sufficiently high to enable an early selection of such mixtures with a sufficient precision and effectiveness? Conclusions on the necessary number of components in mixtures can be derived, given the desired $r_M$. For different plant species, especially for long-living organisms like forest trees, prediction formulae for the calculation of juvenile-mature correlations have been developed. For example:

The phenotypic correlation r for the character height between two different ages in pine trees can be estimated by the formula of Lambeth (1980):

$$r = 1.02 + 0.308 \cdot \ln Q\,, \tag{1}$$

where Q = ratio of early (= juvenile) and late (= mature) age (from planting date).

Experimental results for juvenile-mature correlations of mixtures with respect to the number of components in the mixture have been given by Kleinschmit (1983) for older Norway spruce clonal tests. To see how the correlations change, groups of clones of different size have been formed (1 – 100 clones). By this procedure, the genetic variance slowly increases within the groups, and the expected gain between groups decreases. The correlations between ages 3 and 12 of 0.30 using single clones increase very rapidly if the group size increases: 0.80 with groups of 10 clones, 0.90 with groups of 20 clones, 0.95 with groups of 50 clones and 0.99 – 1.00 for groups of 65 – 100 clones. Simultaneously, the differences in height between the group means, and therefore the gain decrease. However, this gain can be realized at an early stage if the correlations are sufficiently high. In this experimental study the juvenile-mature correlation reaches an approximate constant value for groups of 20 – 30 clones. Increasing this number results in no further significant changes of the correlation.

The aim of this paper is to give some statistical approaches and numerical results on the necessary number of components in mixtures if we proceed from some initial conditions and restrictions concerning $r_M$ and the relation between $r_M$ and $r_E$.

## Theoretical investigations and some numerical results

We introduce the following denotions:

m = number of mixtures

n = number of components in each mixture

$t_{ij}$ = mean of component j in mixture i (i = 1, 2, ..., m; j = 1, 2, ..., n)

$\sigma_T^2$ = total variability of all components – estimated by

$$\hat{\sigma}_T^2 = \frac{1}{mn-1} \sum_{i=1}^{m} \sum_{j=1}^{n} (t_{ij} - \bar{t}_{..})^2$$

$\sigma_B^2$ = variability between mixtures (= groups of components) – estimated by

$$\hat{\sigma}_B^2 = \frac{1}{m-1} \sum_{i=1}^{m} (\bar{t}_{i.} - \bar{t}_{..})^2$$

$\sigma_I^2$ = variability within mixtures – estimated by

$$\hat{\sigma}_I^2 = \frac{1}{m(n-1)} \sum_{i=1}^{m} \sum_{j=1}^{n} (t_{ij} - \bar{t}_{i.})^2$$

with

$$\bar{t}_{i.} = \frac{1}{n} \sum_{j=1}^{n} t_{ij} \text{ and } \bar{t}_{..} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} t_{ij}\,.$$

These parameters are defined for the juvenile age (symbols without asterisk) and the mature age (symbols with asterisk).

$r_E$ = juvenile-mature correlation based upon the single component means – estimated by

$$\hat{r}_E = \frac{\frac{1}{mn-1}\sum_{i=1}^{m}\sum_{j=1}^{n}(t_{ij}-\bar{t}_{..})(t^*_{ij}-\bar{t}^*_{..})}{\sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(t_{ij}-\bar{t}_{..})^2}{mn-1}}\cdot\sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(t^*_{ij}-\bar{t}^*_{..})^2}{mn-1}}}$$

$r_M$ = juvenile-mature correlation based upon the means of mixtures (= means of groups of components) – estimated by

$$\hat{r}_M = \frac{\frac{1}{m-1}\sum_{i=1}^{m}(\bar{t}_{i.}-\bar{t}_{..})(\bar{t}^*_{i.}-\bar{t}^*_{..})}{\sqrt{\frac{\sum_{i=1}^{m}(\bar{t}_{i.}-\bar{t}_{..})^2}{m-1}}\cdot\sqrt{\frac{\sum_{i=1}^{m}(\bar{t}^*_{i.}-\bar{t}^*_{..})^2}{m-1}}}$$

$r_I$ = juvenile-mature correlation computed within the mixtures – estimated by

$$\hat{r}_I = \frac{\frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}(t_{ij}-\bar{t}_{i.})(t^*_{ij}-\bar{t}^*_{i.})}{\sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(t_{ij}-\bar{t}_{i.})^2}{m(n-1)}}\cdot\sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(t^*_{ij}-\bar{t}^*_{i.})^2}{m(n-1)}}}$$

Each mixture shall be composed of equal proportions of its components.

We start from the following well-known expression:

$$\sum_{i=1}^{m}\sum_{j=1}^{n}(t_{ij}-\bar{t}_{..})(t^*_{ij}-\bar{t}^*_{..}) = \sum_{i=1}^{m}\sum_{j=1}^{n}(t_{ij}-\bar{t}_{i.})(t^*_{ij}-\bar{t}^*_{i.}) + \sum_{i=1}^{m}\sum_{j=1}^{n}(\bar{t}_{i.}-\bar{t}_{..})(\bar{t}^*_{i.}-\bar{t}^*_{..})\,. \tag{2}$$

Using the correlation coefficients $r_E$, $r_M$ and $r_I$ (2) can be expressed:

$$r_E = r_I\cdot\frac{\sigma_I\cdot\sigma_I^*}{\sigma_T\cdot\sigma_T^*} + r_M\cdot\frac{\sigma_B\cdot\sigma_B^*}{\sigma_T\cdot\sigma_T^*}\,. \tag{3}$$

From $\sigma_B^2+\sigma_I^2=\sigma_T^2$ and $(\sigma_B^2)^*+(\sigma_I^2)^*=(\sigma_T^2)^*$ together with (3) we obtain:

$$r_E = r_M\sqrt{\left(1-\frac{\sigma_I^2}{\sigma_T^2}\right)\left(1-\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}\right)} + r_I\sqrt{\left(\frac{\sigma_I^2}{\sigma_T^2}\right)\left(\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}\right)} \tag{4}$$

Equation (4) can be expressed as a linear relation between $r_M/r_E$ and $r_I/r_E$:

$$\frac{r_M}{r_E} = -\sqrt{\frac{\left(\frac{\sigma_I^2}{\sigma_T^2}\right)\left(\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}\right)}{\left(1-\frac{\sigma_I^2}{\sigma_T^2}\right)\left(1-\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}\right)}}\cdot\frac{r_I}{r_E} + \frac{1}{\sqrt{\left(1-\frac{\sigma_I^2}{\sigma_T^2}\right)\left(1-\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}\right)}}\,. \tag{5}$$

To improve the effectiveness of an early selection by using mixtures instead of pure stands $r_M > r_E$ would be a necessary condition. $r_M > r_E$ together with (5) gives:

$$\frac{r_I}{r_E} < \frac{1-\sqrt{\left(1-\frac{\sigma_I^2}{\sigma_T^2}\right)\left(1-\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}\right)}}{\sqrt{\left(\frac{\sigma_I^2}{\sigma_T^2}\right)\left(\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}\right)}}\,. \tag{6}$$

These upper bounds for $\frac{r_I}{r_E}$ by (6) are presented in Table 1 for different $\frac{\sigma_I^2}{\sigma_T^2}$ and $\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}$.

Only for the special case $\frac{\sigma_I^2}{\sigma_T^2}=\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}$ shall some comments be given. This assumption of equal proportions of the variability within the mixtures at the juvenile and at the mature age may be a realistic condition for practical applications.

Values for $\frac{r_M}{r_E}$ calculated by (5) are demonstrated in Fig. 1. It is evident that large $\frac{r_M}{r_E}$ can be realized with low $\frac{r_I}{r_E}$ and large $\frac{\sigma_I^2}{\sigma_T^2}$. Both conditions can be combined in a way that the ratio $\frac{\sigma_I^2}{\sigma_T^2}\Big/\frac{r_I}{r_E}$ should be large. Since we are only interested in the dependence on the number n of components in the mixture, "$\frac{\sigma_I^2}{\sigma_T^2}\Big/\frac{r_I}{r_E}$ = large" is equivalent to "$\frac{\sigma_I^2}{r_I}$ = large". Both the numerator and the denominator of this ratio depend on the number n of components in the mixture. Results for necessary numbers n can be obtained if these functions $\sigma_I^2=\sigma_I^2(n)$ and $r_I=r_I(n)$ are explicitly known. This may be possible either by theoretical assumptions or by experimental results but these investigations shall not be further elaborated in this paper.

For $\frac{\sigma_I^2}{\sigma_T^2}=\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}$ the condition $\frac{r_M}{R_E}>1$ implies (by (6)) $\frac{r_I}{r_E}<1$. Therefore, $r_I<r_E$ would be a necessary condition for $r_M>r_E$. For each $\frac{r_I}{r_E}$ with $\frac{r_I}{r_E}<1$, the $\frac{r_M}{r_E}$ increases with increasing $\frac{\sigma_I^2}{\sigma_T^2}$. With a decreasing number of components the mixture will become more homogeneous and $\sigma_I^2$ will decrease; therefore, $\frac{\sigma_I^2}{\sigma_T^2}$ will also decrease.

Thus, we may conclude that enlarged juvenile-mature correlations which are based upon the means of mixtures may be obtained by using an increased number of components in the mixture.

**Table 1.** Upper bounds for $\frac{r_I}{r_E}$ from formula (6) dependent on different $\frac{\sigma_I^2}{\sigma_T^2}$ and $\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}$

| $\frac{\sigma_I^2}{\sigma_T^2}$ \ $\frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.1 | 1.00 | 1.07 | 1.19 | 1.33 | 1.47 | 1.63 | 1.82 | 2.04 | 2.33 | 3.16 |
| 0.2 | 1.07 | 1.00 | 1.03 | 1.09 | 1.16 | 1.25 | 1.36 | 1.50 | 1.69 | 2.24 |
| 0.3 | 1.19 | 1.03 | 1.00 | 1.02 | 1.05 | 1.11 | 1.18 | 1.28 | 1.42 | 1.83 |
| 0.4 | 1.33 | 1.09 | 1.02 | 1.00 | 1.01 | 1.04 | 1.09 | 1.16 | 1.26 | 1.58 |
| 0.5 | 1.47 | 1.16 | 1.05 | 1.01 | 1.00 | 1.01 | 1.04 | 1.08 | 1.16 | 1.41 |
| 0.6 | 1.63 | 1.25 | 1.11 | 1.04 | 1.01 | 1.00 | 1.01 | 1.04 | 1.09 | 1.29 |
| 0.7 | 1.82 | 1.36 | 1.18 | 1.09 | 1.04 | 1.01 | 1.00 | 1.01 | 1.04 | 1.26 |
| 0.8 | 2.04 | 1.50 | 1.28 | 1.16 | 1.08 | 1.04 | 1.01 | 1.00 | 1.01 | 1.12 |
| 0.9 | 2.33 | 1.69 | 1.42 | 1.26 | 1.16 | 1.09 | 1.04 | 1.01 | 1.00 | 1.05 |
| 1.0 | 3.16 | 2.24 | 1.83 | 1.58 | 1.41 | 1.29 | 1.26 | 1.12 | 1.05 | 1.00 |

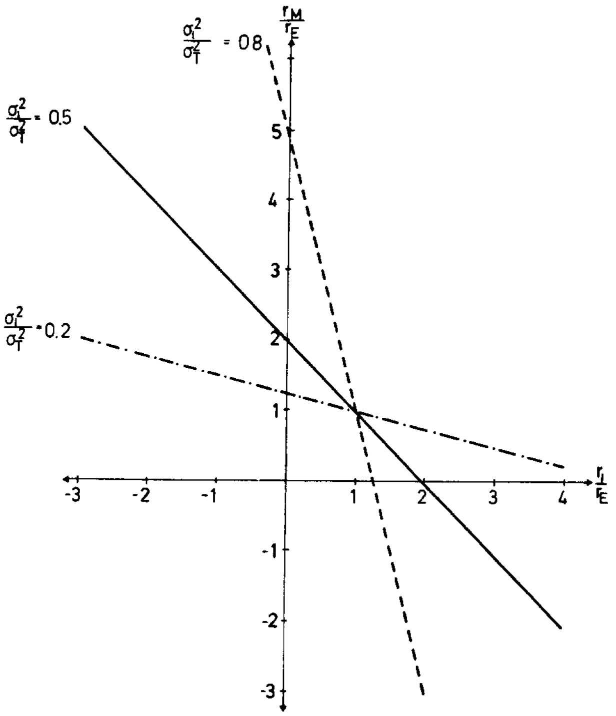

**Fig. 1.** Dependence of $\frac{r_M}{r_E}$ on $\frac{r_I}{r_E}$ for different numerical values of $\frac{\sigma_I^2}{\sigma_T^2}$ $\left(\text{in the special case } \frac{\sigma_I^2}{\sigma_T^2} = \frac{(\sigma_I^2)^*}{(\sigma_T^2)^*}\right)$

Finally, the previously discussed relations and comparison between $r_E$ and $r_M$ shall be considered statistically. That means: we may ask for the conditions for a statistically significant difference between $r_E$ and $r_M$. We denote:

N = total number of components, which are available for the composition of mixtures,

n = number of components in each mixture.

If all existing components shall be used once we have $\frac{N}{n}$ different mixtures.

If we consider $r_M$ and $r_E$ as estimates of correlation coefficients we may ask for the conditions that $r_E$ and $r_M$ differ significantly. This approach seems to be obvious because the juvenile-mature correlation $r_E$ based upon the single component means shall be increased to $r_M$ by using mixtures = groups of components to enable an effective early selection already at the moment when the mixture has to be established.

Testing the significance of a difference between two correlation coefficients $r_1$ and $r_2$ is based on the Fisher-transformation $z = \ln \sqrt{(1+r)/(1-r)}$ where the test-statistic

$$\frac{|z_1 - z_2|}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}} \tag{7}$$

can be considered as a standard normal variate (under the null hypothesis) (Fryer 1966). $n_1$ and $n_2$ are the numbers of observations which are available for the calculation of $r_1$ and $r_2$ respectively. If we identify $r_1 = r_M$ and $r_2 = r_E$ the corresponding numbers $n_1$ and $n_2$ are: $n_1 = \frac{N}{n}$ and $n_2 = N$ and (7) gives:

$$\frac{\left|\frac{1}{2} \ln \frac{(1 + r_M)(1 - r_E)}{(1 - r_M)(1 + r_E)}\right|}{\sqrt{\frac{N(n+1) - 6n}{(N - 3n)(N - 3)}}} . \tag{8}$$

For a given type I error $\alpha$ with truncation point $t_\alpha$ from the standard normal distribution we ask for the number n leading to a significant difference between $r_E$ and $r_M$.

Equating expression (8) with $t_\alpha$ and solving for n gives the explicit expression of $n = n(N, r_M, r_E, \alpha)$:

$$n = \frac{N^2\varphi - N(1+3\varphi)}{N(1+3\varphi) - (6+9\varphi)} \tag{9}$$

with

$$\varphi = \left( \frac{\left| \frac{1}{2} \ln \frac{(1+r_M)(1-r_E)}{(1-r_M)(1+r_E)} \right|}{t_\alpha} \right)^2 . \tag{10}$$

Finally, some applications to height growth in pines shall be given. If we use Lambeth's formula (1) to estimate $r_E$, the term $\varphi$ can be rewritten as:

$$\varphi = \left( \frac{\left| \frac{1}{2} \ln \frac{(1+r_M)(-0.02 - 0.308 \ln Q)}{(1-r_M)(2.02 + 0.308 \ln Q)} \right|}{t_\alpha} \right)^2 , \tag{11}$$

where Q = ratio of early and late age (from planting date).

If we use: "early (= juvenile) age" = 1 year after planting and T = time (in years) from planting date until the "late (= mature) age" we have:

$$\ln Q = \ln 1 - \ln T = -\ln T . \tag{12}$$

Combining (12) and (11) gives:

$$\varphi = \left( \frac{\left| \frac{1}{2} \ln \frac{(1+r_M)(0.308 \ln T - 0.02)}{(1-r_M)(2.02 - 0.308 \ln T)} \right|}{t_\alpha} \right)^2 . \tag{13}$$

Now, n from (9) can be calculated numerically dependent on N, $r_M$, $\alpha$ and T:

$$n = n(T, N, r_M, \alpha) . \tag{14}$$

If we require a juvenile-mature correlation $r_M$ of 0.90 and a type I error $\alpha$ of $\alpha = 0.05$ the number n of components can be calculated dependent on N and T. These results are presented in Table 2.

A numerical example can be given, which may be relevant for applications is clonal forestry – for example, in developing multiclonal varieties of poplars or aspen. Using only tested and approved clonal material for the composition of such clonal mixtures and with respect to the relatively short rotation time of these species we may use: $N \leqq 200$ and $T \leqq 10$ and from Table 2 we obtain: $n \leqq 34$. For $N \leqq 500$ and $T \leqq 20$ we have $n \leqq 100$.

For T = 5, T = 10 and T = 20 the necessary numbers n are: 6, 9 and 10 (for N = 50); 13, 17 and 20 (for N = 100); 26, 34 and 40 (for N = 200); 38, 51 and 60 (for N = 300); 64, 85 and 100 (for N = 500) and 128, 171 and 199 (for N = 1,000). Thus, the number n of components decreases considerably with decreasing N.

## Discussion

No successive generations were analysed. Only one rotation time from the initial composition of the mixture (= juvenile age) until the final harvest (= ma-

**Table 2.** Necessary numbers n of components in mixtures dependent on N (total number of components) and T (rotation time (in years) from planting date) (for $\alpha = 0.05$ and $r_M = 0.90$)

| T \ N | 30 | 50 | 100 | 150 | 200 | 250 | 300 | 400 | 500 | 700 | 1,000 | 2,000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 2 | 4 | 8 | 12 | 16 | 20 | 24 | 32 | 40 | 56 | 81 | 162 |
| 4 | 3 | 5 | 11 | 16 | 22 | 27 | 33 | 44 | 55 | 76 | 109 | 218 |
| 5 | 4 | 6 | 13 | 19 | 26 | 32 | 38 | 51 | 64 | 89 | 128 | 255 |
| 6 | 4 | 7 | 14 | 21 | 28 | 35 | 42 | 56 | 70 | 98 | 140 | 281 |
| 7 | 5 | 8 | 15 | 23 | 30 | 38 | 45 | 60 | 75 | 105 | 150 | 300 |
| 8 | 5 | 8 | 16 | 24 | 32 | 40 | 48 | 63 | 79 | 111 | 158 | 316 |
| 9 | 5 | 8 | 17 | 25 | 33 | 41 | 49 | 66 | 82 | 115 | 164 | 328 |
| 10 | 5 | 9 | 17 | 26 | 34 | 43 | 51 | 68 | 85 | 119 | 171 | 337 |
| 13 | 6 | 9 | 18 | 28 | 37 | 46 | 55 | 73 | 91 | 127 | 182 | 364 |
| 15 | 6 | 10 | 19 | 28 | 38 | 47 | 57 | 75 | 94 | 132 | 187 | 376 |
| 20 | 6 | 10 | 20 | 30 | 40 | 50 | 60 | 80 | 100 | 139 | 199 | 398 |
| 25 | 7 | 11 | 21 | 31 | 42 | 52 | 62 | 83 | 104 | 145 | 207 | 413 |
| 30 | 7 | 11 | 22 | 32 | 43 | 53 | 64 | 85 | 107 | 149 | 213 | 425 |
| 40 | 7 | 11 | 22 | 34 | 45 | 57 | 67 | 89 | 111 | 155 | 222 | 443 |
| 50 | 7 | 12 | 23 | 35 | 46 | 57 | 69 | 91 | 114 | 160 | 228 | 455 |
| 80 | 8 | 12 | 24 | 36 | 48 | 60 | 72 | 96 | 120 | 168 | 240 | 480 |
| 100 | 8 | 13 | 25 | 37 | 50 | 62 | 74 | 99 | 123 | 172 | 246 | 492 |

ture age) has been considered. Both ages are related to each other and this relation has been described quantitatively by $r_E$ (= juvenile-mature correlation based upon the means of the single components) and $r_M$ (= juvenile-mature correlation based upon the means of mixtures of different components). The effectiveness of an early selection will be mainly determined by the magnitude of the juvenile-mature correlation. If we have $r_M > r_E$ an improvement of early testing can be realized by using mixtures instead of single components. All the considerations and conclusions of this paper are based upon quantitative relations between $r_E$ and $r_M$. They are, therefore, motivated only statistically and not genetically or by breeding methodology.

A critical discussion of the numerical results on necessary numbers of components primarily should be deal with an investigation of the numerous simplifying assumptions and their possible importance and resulting restrictions. For example, if we want to generalize the simplification of equal proportions of the components in the mixtures: unequal proportions will lead to different weights of the $t_{ij}$ (= mean of component j in mixture i). Consequently, the correlation coefficients as well as the standard deviations in (3) will be influenced by these different weights. Although some of these effects may be cancelled out in computing the ratios of (3) (correlation coefficients = ratios of covariances and products of standard deviations and ratios of standard deviations) biasing effects with respect to (3) and the following conclusions must be expected. Their importance and magnitude cannot be estimated without postulating further assumptions.

The previously used approach "significant difference between $r_E$ and $r_M$ for a given numerical value of $r_M$" can be strengthened in the following way: "significant difference between $r_E$ and $r_M$ for a given numerical value of $r_M$ where this difference exceeds a required numerical value". There are several arguments indicating that such an approach would be more realistic than the previous procedure. However, what are the appropriate numerical values for this difference $r_M - r_E$ which should be proposed?

The same argument is valid with respect to the required numerical level of $r_M$. We used $r_M = 0.90$. Why just this value? No substantial arguments are given and 0.90 seems to be an arbitrary supposition.

Just identical considerations will be valid for the arbitrarily chosen numerical value of 0.05 for the type I error $\alpha$.

To investigate the dependence of n on varying $\alpha$ and varying $r_M$ we may apply elementary calculus to formula (9). But here, no separate derivatives of n with respect to $\alpha$ and $r_M$ shall be given and discussed explicitly. The following simple approach will be sufficient to demonstrate the considerable sensitivity of n according to varying $\alpha$ and varying $r_M$: For each given age T the expression $\varphi$ depends on both $\alpha$ and $r_M$. A simultaneous variation of $r_M$ and $\alpha$ causes a varying $\varphi$ (by formula (13)). If we consider n as a function of only $\varphi$ (for fixed N and T) we obtain for this derivative:

$$\frac{\partial n}{\partial \varphi} = \left( \frac{1}{3\varphi + \frac{N-6}{N-3}} \right)^2 \cdot N . \tag{15}$$

For sufficiently large N this reduces to the approximate result:

$$\frac{\partial n}{\partial \varphi} = (1 + 3\varphi)^{-2} \cdot N . \tag{16}$$

Therefore, we obtain considerable changes of n for varying $\varphi$ and, consequently, for varying $\alpha$ and varying $r_M$.

To provide some further insight into these dependencies and results numerical calculations of $n = n(T, N, r_M, \alpha)$ have been performed and summarized in Table 3. These results confirm the previous statements and conclusions completely.

Another criticism may be the application of Lambeth's formula as an estimate for $r_E$. The Lambeth-formula (1) has been developed for height growth in pines. Therefore, applications should be restricted to height in pines. However, several investigations are known where the Lambeth-formula fits very well to experimental data of other tree species too (see, for example Hühn 1984). Therefore, numerical results obtained by applying the Lambeth-formula may be of an extended validity and interest.

In the preceding correlation-study (= comparison between $r_M$ and $r_E$) we assumed that all existing components should be used once for the composition of mixtures. This leads to $\frac{N}{n}$ different mixtures. In this context some other situations may be possible and even realistic:

1. Not all the N components should be used for the composition of mixtures, but only a selected part of them – for example the highest-yielding.
2. Each of the N components or at least some of these components could be used several times. Therefore, the same component may be included in several mixtures.

In both situations the number of mixtures whereon the correlation $r_M$ has been based will be different from the previously used number $\frac{N}{n}$. Here, no further discussion of 1. and 2. shall be given.

Sometimes the main approach of this paper 'Increase of the juvenile-mature correlation by using mix-

**Table 3.** Necessary numbers n of components in mixtures dependent on N (total number of components), T (rotation time (in years) from planting date), $\alpha$ (type I error) and $r_M$ (juvenile-mature correlation based on the means of mixtures)

| T | | *N* = 100 | | | *N* = 200 | | | *N* = 300 | | | *N* = 500 | | | *N* = 1,000 | | | *N* = 1,500 | | | *N* = 2,000 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha$ \ $r_M$ | 0.90 | 0.80 | 0.70 | 0.90 | 0.80 | 0.70 | 0.90 | 0.80 | 0.70 | 0.90 | 0.80 | 0.70 | 0.90 | 0.80 | 0.70 | 0.90 | 0.80 | 0.70 | 0.90 | 0.80 | 0.70 |
| 5 | 0.10 | 16 | 8 | 2 | 31 | 15 | 5 | 47 | 23 | 8 | 78 | 38 | 14 | 157 | 76 | 27 | 235 | 115 | 41 | 313 | 153 | 55 |
| | 0.05 | 13 | 6 | 2 | 26 | 11 | 4 | 38 | 17 | 6 | 64 | 29 | 10 | 128 | 57 | 20 | 191 | 86 | 29 | 255 | 115 | 39 |
| | 0.01 | 9 | 3 | 1 | 17 | 7 | 2 | 26 | 10 | 3 | 44 | 18 | 5 | 88 | 35 | 11 | 131 | 53 | 17 | 175 | 71 | 23 |
| 10 | 0.10 | 20 | 14 | 8 | 40 | 27 | 17 | 60 | 40 | 25 | 100 | 67 | 42 | 199 | 134 | 83 | 298 | 201 | 125 | 398 | 268 | 166 |
| | 0.05 | 17 | 11 | 6 | 34 | 21 | 12 | 51 | 32 | 19 | 85 | 53 | 31 | 170 | 107 | 63 | 254 | 160 | 94 | 339 | 214 | 126 |
| | 0.01 | 13 | 7 | 4 | 25 | 14 | 8 | 37 | 21 | 12 | 62 | 36 | 19 | 125 | 71 | 39 | 187 | 107 | 59 | 249 | 143 | 79 |
| 15 | 0.10 | 22 | 16 | 11 | 44 | 32 | 23 | 65 | 48 | 34 | 108 | 80 | 57 | 216 | 160 | 113 | 324 | 240 | 170 | 432 | 321 | 227 |
| | 0.05 | 19 | 13 | 9 | 38 | 26 | 18 | 57 | 39 | 26 | 94 | 66 | 44 | 188 | 131 | 88 | 282 | 197 | 132 | 376 | 262 | 177 |
| | 0.01 | 14 | 9 | 6 | 29 | 18 | 11 | 43 | 27 | 17 | 71 | 45 | 28 | 142 | 91 | 57 | 213 | 136 | 86 | 284 | 181 | 115 |
| 20 | 0.10 | 23 | 18 | 13 | 46 | 35 | 27 | 68 | 53 | 40 | 113 | 88 | 66 | 226 | 176 | 133 | 339 | 264 | 199 | 453 | 352 | 265 |
| | 0.05 | 20 | 15 | 11 | 40 | 29 | 21 | 60 | 44 | 32 | 100 | 73 | 53 | 199 | 147 | 105 | 298 | 220 | 158 | 398 | 293 | 211 |
| | 0.01 | 16 | 10 | 7 | 31 | 21 | 14 | 46 | 31 | 21 | 77 | 52 | 35 | 154 | 104 | 70 | 230 | 156 | 105 | 307 | 207 | 140 |
| 25 | 0.10 | 24 | 19 | 15 | 47 | 38 | 29 | 70 | 56 | 44 | 117 | 94 | 73 | 233 | 187 | 147 | 350 | 281 | 220 | 467 | 374 | 293 |
| | 0.05 | 21 | 16 | 12 | 42 | 32 | 24 | 62 | 47 | 36 | 104 | 79 | 59 | 207 | 158 | 118 | 310 | 236 | 177 | 413 | 315 | 236 |
| | 0.01 | 16 | 11 | 8 | 33 | 23 | 16 | 49 | 34 | 24 | 81 | 57 | 40 | 162 | 114 | 80 | 242 | 170 | 120 | 323 | 227 | 160 |
| 30 | 0.10 | 24 | 20 | 16 | 48 | 39 | 32 | 72 | 59 | 47 | 120 | 98 | 79 | 239 | 196 | 157 | 358 | 293 | 236 | 477 | 391 | 315 |
| | 0.05 | 22 | 17 | 13 | 43 | 33 | 26 | 64 | 50 | 39 | 107 | 83 | 64 | 213 | 166 | 128 | 319 | 249 | 192 | 425 | 332 | 257 |
| | 0.01 | 17 | 12 | 9 | 34 | 24 | 18 | 51 | 36 | 26 | 84 | 61 | 44 | 168 | 121 | 88 | 252 | 182 | 133 | 336 | 243 | 177 |
| 35 | 0.10 | 25 | 21 | 17 | 49 | 41 | 33 | 73 | 61 | 50 | 122 | 101 | 83 | 243 | 202 | 166 | 364 | 303 | 249 | 485 | 404 | 332 |
| | 0.05 | 22 | 18 | 14 | 44 | 35 | 27 | 66 | 52 | 41 | 109 | 87 | 68 | 218 | 173 | 137 | 326 | 259 | 205 | 435 | 346 | 273 |
| | 0.01 | 18 | 13 | 9 | 35 | 26 | 19 | 52 | 38 | 29 | 87 | 64 | 48 | 173 | 128 | 95 | 260 | 192 | 143 | 346 | 256 | 191 |
| 40 | 0.10 | 25 | 21 | 18 | 50 | 42 | 35 | 74 | 63 | 52 | 123 | 104 | 87 | 246 | 208 | 173 | 369 | 311 | 260 | 492 | 415 | 346 |
| | 0.05 | 22 | 18 | 15 | 45 | 36 | 29 | 67 | 54 | 43 | 111 | 90 | 72 | 222 | 179 | 144 | 332 | 268 | 215 | 443 | 357 | 287 |
| | 0.01 | 18 | 13 | 10 | 36 | 27 | 20 | 54 | 40 | 30 | 89 | 67 | 51 | 178 | 133 | 101 | 266 | 200 | 152 | 355 | 267 | 203 |
| 50 | 0.10 | 26 | 22 | 19 | 51 | 44 | 37 | 76 | 65 | 56 | 126 | 108 | 93 | 252 | 216 | 185 | 378 | 325 | 277 | 503 | 433 | 369 |
| | 0.05 | 23 | 19 | 16 | 46 | 38 | 31 | 69 | 57 | 47 | 114 | 94 | 78 | 228 | 188 | 155 | 342 | 282 | 233 | 455 | 376 | 310 |
| | 0.01 | 19 | 14 | 11 | 37 | 29 | 22 | 56 | 43 | 33 | 93 | 71 | 56 | 185 | 142 | 111 | 277 | 214 | 167 | 369 | 285 | 223 |
| 100 | 0.10 | 27 | 24 | 22 | 54 | 48 | 44 | 80 | 72 | 65 | 134 | 121 | 109 | 267 | 241 | 217 | 400 | 361 | 325 | 534 | 481 | 433 |
| | 0.05 | 25 | 22 | 19 | 50 | 43 | 38 | 74 | 65 | 57 | 123 | 108 | 94 | 246 | 215 | 189 | 369 | 322 | 283 | 492 | 429 | 377 |
| | 0.01 | 21 | 17 | 14 | 42 | 34 | 29 | 62 | 51 | 43 | 103 | 85 | 72 | 206 | 170 | 143 | 309 | 255 | 214 | 412 | 340 | 286 |
| 150 | 0.10 | 28 | 26 | 24 | 55 | 51 | 47 | 83 | 76 | 70 | 138 | 127 | 117 | 275 | 253 | 234 | 413 | 380 | 351 | 550 | 506 | 467 |
| | 0.05 | 26 | 23 | 21 | 52 | 46 | 42 | 77 | 69 | 62 | 128 | 115 | 104 | 256 | 230 | 207 | 384 | 344 | 311 | 512 | 459 | 414 |
| | 0.01 | 22 | 19 | 16 | 44 | 38 | 33 | 66 | 56 | 49 | 110 | 94 | 81 | 219 | 187 | 162 | 328 | 280 | 243 | 437 | 373 | 324 |

tures instead of single components' may be useless: For example, if already $r_E$ shows such an extreme numerical level that a further increase of this correlation would be unnecessary for the purpose of improving an early selection. Just this situation of large $r_E$ may be often realized in practical applications, for example in breeding work with fast-growing tree species like poplars and aspen (Hühn 1984). These applications are characterized by:

1. Relatively short time-interval between juvenile and mature ages.
2. Small disconnected areas.
3. Short rotation time.
4. Homogeneous environments.
5. Use of tested components and mixtures.
6. Reduced effects of genotype-environment interactions.

All these factors 1 – 6 lead to enlarged $r_E$-values.

In many other situations the correlations $r_E$ will be much lower. Here, the effectiveness of an early selection can be improved by using mixtures instead of single components.

All the correlations introduced in the preceding chapter have been considered as phenotypic juvenile-mature correlations. But, for example, if the components are clones and the mixtures, therefore, multiclonal varieties the component means are estimates for the genotypic values of the clones. Thus, the correlation $r_E$ can be interpreted as a genetic correlation. The genetic correlation is an estimate based on genetic effects. Therefore, as a general comment, we think that further improvements of the previous results on necessary numbers of components may be obtained by using quantitative genetic models and procedures.

## References

Fryer HC (1966) Concepts and methods of experimental statistics. Allyn and Bacon, Boston

Hühn M (1984) Theoretische Untersuchungen zur Problematik der Mindestklonzahl in Mehrklonsorten. 3. Untersuchungen über Jugend-Alters-Korrelationen und erforderliche Klonanzahlen in Mehrklonsorten. Hessisches Ministerium für Landesentwicklung, Umwelt, Landwirtschaft und Forsten Wiesbaden, pp 1 – 30

Hühn M (1985a) Theoretical studies on the necessary number of components in mixtures. 1. Number of components and yield stability. Theor Appl Genet 70: 383 – 389

Hühn M (1986a) Theoretical studies on the necessary number of components in mixtures. 2. Number of components and yielding ability. Theor Appl Genet 71:622 – 630

Hühn M (1986b) Theoretical studies on the necessary number of components in mixtures. 3. Number of components and risk considerations. Theor Appl Genet 72:211 – 218

Kleinschmit J (1983) Concepts and experiences in clonal plantations of conifers. In: Can Tree Impr Ass, 19th Bienn Meet Toronto, pp 26–56

Lambeth CC (1980) Juvenile-mature correlations in Pinaceae and implications for early selection. For Sci 26:571 – 580

Nanson A (1967) Modèle théorique pour l'étude des tests précoces. Biom Praxim 8:84 – 107

Nanson A (1968) La valeur des tests précoces dans la sélection des arbres forestiers en particulier au point de vue de la croissance. Diss Doct Gembloux, Stat Rech Eaux et Forêts, Groenendaal, HS 38, pp 242

Nanson A (1974) Some problems of application of quantitative genetics to forest tree breeding. In: Proc IUFRO-Joint Meeting Working Parties Popul Ecol Genet, Breed Theory Prog Testing. Dep For Genet R Coll For Stockholm, pp 1 – 15

Nanson A (1976) Juvenile-mature relationships mainly in provenance and progeny tests. In: IUFRO-Joint Meeting Adv Gener Breed, Bordeaux, pp 99 – 119

Sziklai O (1974) Juvenile-mature correlation. In: Proc IUFRO-Joint Meeting Working Parties Popul Ecol Genet, Breed Theory Prog Testing. Dep For Genet R Coll For, Stockholm, pp 217 – 235