Theoretical and Applied Genetics 46, 87-99 (1975)

Estimation of Broad Sense Heritability in Plant Populations: an Improved Method

M. Hühn
Institut für Pflanzenbau und Pflanzenzüchtung der Universität Kiel, Kiel (BRD)

Summary. In the present paper an improved method for estimating broad sense heritability is developed by generalization and improvement of the method of Sakai and Mukaide (1967); this itself is a generalization and improvement of Sakai and Hatakeyama's (1963) modification of Shrikhande's (1957) method to separately estimate the genetic variance, competitional variance and environmental variance of a plant population. Some of the assumptions postulated by these former authors - especially the assumption that the covariance between the genetic and competitional effects equals zero - were omitted. If competitive effects are genetically caused and controlled - and numerous experimental results leave no doubt about this - then this assumption can not be right. Our proposed improvement and generalization avoids this difficulty by considering competition as a usual quantitative character (for each genotype we introduce two quantitative competition-characters: competitive ability and competitive influence). The expected values of the various terms in the variance of plot means (for plots of different sizes) were derived to give a system of simultaneous non-linear equations from which the unknown parameters, genetic variance resp. broad sense heritability, competitional variance, environmental variance etc., can be estimated using least squares methods or direct search methods. The estimates for broad sense heritability are probably more correct and realistic (because of the altered assumptions proposed in this method) than the results of the former authors. The application of the proposed estimation-procedure is demonstrated using Norway-spruce data from Slovakia: Norway-spruce stands (80-90 years old) were investigated for four characters: height, diameter, crown percentage and taper, measured for each single tree.

Introduction

The broad sense heritability ($h^2_{b.s.}$) of a character, that is the proportion of the genetic variance to the phenotypic total variance, is an important parameter in breeding and genetics, because a knowledge of the numerical magnitude of this heritability is of special importance for planning breeding programs (for example for the prediction of the gain of selection) and for the examination of experimental results. The estimation of heritabilities therefore is an essential task. The computation of $h^2_{b.s.}$ assumes knowledge of the phenotypic total variance and of its genetically caused component.

To compute this component there are different possibilities, which differ in relation to the special object of the experimental study and depend on the special breeding program.

Some general possibilities are, for example, :

1) Using the covariances between relatives.

2) Using regression techniques, for example the parent-offspring-regression.

Using each of these procedures the estimates of $h^2_{b.s.}$ are biased by not considering the competitive effects between neighbouring plants (Hühn a-c).

With all these estimation procedures several generations are necessary.

When working with plant species having very large generation intervals, for example with perennial plant species or even forest trees, these estimation procedures are not sufficient because of the long time involved. Mostly it is necessary to get heritability estimates in the early stages of breeding work - and this would be the case even if it were only possible to get some rough estimates or approximate results on broad sense heritability.

Methods to estimate $h^2_{b.s.}$, which can be applied without needing progeny tests or crosses, and which lead to estimates with satisfactory properties are therefore of special importance for breeding work with such plant species. The proposed method is a generalization and improvement of the original method of Shrikhande (1957), which has been further developed by Sakai and Hatakeyama (1963) and by Sakai and Mukaide (1967); for this purpose (estimation of $h^2_{b.s.}$) this method has been used frequently in the literature (Sakai, Hayashi and Mukaide (1966), Sakai, Mukaide and Tomita (1968), Kedharnath, Chetty and Rawat (1969), Morgner and Horn (1970), Namkoong and Squillace (1970) et al.). In two former publications (Hühn a and Hühn b) we have developed a generalization of the method of Shrikhande/Sakai/Hatakeyama/Mukaide to estimate separately the genetic variance, competitional variance and environmental variance of a plant population. In this generalization and improvement some of the assumptions of these authors - especially the assumption that the covariance between the genetic and competitional effects equals zero - were omitted, because experimental investigations have shown the opposite. This generalization of permitting covariances not equal to zero was possible under the assumption that competitive effects are genetically caused and controlled. We assumed that, for the description of the competition properties of a population, each genotype X can be assigned two characters, competitive ability F_X and competitive influence W_X, which are both hereditary like the usual quantitative characters.

This generalization and improvement of the method of Shrikhande/Sakai/Hatakeyama/Mukaide (published in Hühn a and Hühn b) has been developed in these papers for the following special and simple genetic model: random mating, one locus with two alleles in equilibrium, pleiotropic gene action of the two alleles upon the two characters F and W, equal degrees of dominance in F and W.

For practical applications, however, this is an unrealistic case. In the present paper the underlying theory and the estimation-procedure derived have been generalized further in such a way that it is not now necessary to agree on a definite special genetic model. The theory, and therefore the estimation-procedure, is also now valid and applicable to a plant population consisting of an arbitrary number of different genotypes with arbitrary frequencies, without making any assumptions about the number of the existing loci, about the number of alleles, about the mode of gene action etc.

Theoretical Investigations

We propose a population which fulfils the following simplifying assumptions:

1) The plants are regularly dispersed - that is with equal distances - over the area. Hence it follows that density is no source of variation - but the genetical and competitive correlations and variations exist.

2) The plants are of the same age.

3) The considered plant stand is 'sufficiently large'.

4) We give no regard to the fact that the reciprocal influence of neighbouring individuals is different at different ages and different stages of plant growth. We assume that it is sufficient to consider the situation at only one certain stage of life.

5) A plant is only influenced by its four direct neighbours and the effects of all other neighbouring plants are negligible.

6) The effect of one plant on another shall be independent of the directional position of its competitors. Therefore, only the genotypical composition of the group of competitors has any bearing on the resulting effects.

7) The effects of the four neighbouring plants add up linearly, and we get the phenotypic value P_X of plant X from:

$$\underset{\substack{\downarrow\\ \text{competitive ability term}}}{P_X = F_X} + \underset{\substack{\downarrow\\ \text{total competitive influence term}}}{\sum_{i(X)} W_{i(X)}} + \underset{\substack{\downarrow\\ \text{environmental term}}}{e_X} \qquad (1)$$

with:

P_X = phenotypic value of plant X.

F_X = competitive ability of plant X.

$W_{i(X)}$ = competitive influences of the four neighbours of plant X, where the summation-index i(X) runs over the four neighbours of plant X.

e_X = environmental deviation of plant X - defined as the deviation of the phenotypic value from its expected value based on the F-effects and W-effects.

Thus, using model (1), the phenotypic value of a plant is dependent only on the genotype of this plant and on the genotypical composition of its neighbourhood.

Finally we will give some further notations:

ρ = correlation coefficient between the F-values and the W-values of all plants of the total considered plant stand, with $-1 \leqslant \rho \leqslant +1$.

G = genetic variance

C = competitional variance

E = environmental variance

x = number of plants per plot (for different reasons we only consider throughout all following investigations square-plots, e.g. $x = n^2 = n \times n$).

b = parameter from F. Smith's empirical law, whose validity and applicability for the environmental component is postulated in this study

$V_{\bar{x}}$ = variance of the means of plots with x plants per plot

$\mathcal{E}$ = symbol for expectation

$h^2_{b.s.}$ = broad sense heritability

Historical excursion

Sakai and Hatakeyama (1963) tried to estimate the genetic variance (G) and environmental variance (E) using Shrikhande's equation:

$$V_{\bar{x}} = G/x + E/x^b \tag{2}$$

To estimate competitional variances and to take competitive effects (which are a main cause for the bias of the estimates) into account explicitly, this model (2) was generalized by Sakai and Mukaide (1967) by adding a competitional term C in the following form:

$$V_{\bar{x}} = G/x + E/x^b + T_x C/x \tag{3}$$

or

$$x V_{\bar{x}} = G + x^B E + T_x C \tag{4}$$

with B = 1-b and the coefficient T_x of proportionality, which follows from the assumptions of Sakai's model of competition. After applying equation (4) on various plot sizes one gets a system of simultaneous equations, which can be solved for different values of B - for example using the method of least squares. For each such chosen B one obtains estimates of the parameters G, C and E. Those values G_0, E_0, C_0 and B_0 are taken as estimates of the unknown parameters, which best fit with the empirical values $xV_{\bar{x}}$. This procedure of estimating genetic variances without performing progeny tests has been used very frequently in recent years: Sakai and Hatakeyama (1963); Sakai, Hayashi and Mukaide (1966); Sakai and Mukaide (1967); Sakai, Mukaide and Tomita (1968); Kedharnath, Chetty and Rawat (1969); Morgner and Horn (1970); et al. A critical discussion of this estimation procedure and of its weak points (for example the case b = 1) is given by Namkoong and Squillace (1970).

In order to apply this estimation-procedure it is necessary to divide the considered plant stand into plots of various sizes; for various reasons (see, for example, Hühn a) we will only consider square plots in all following investigations. After dividing the plant stand into plots of various sizes the means of these plots and their variances are calculated for each plot size. These variance-values are the empirical data which are necessary for applying this estimation-procedure.

The purpose of the following studies is to obtain estimates for G, C and E (which includes a broad-sense-heritability-estimate) making use only of these phenotypic variances of plot means for various plot sizes. For this purpose we need the expectation of the variance of plot means - expressed as a function of G, C, E and the plot size $x = n^2$.

All following investigations are based upon the following assumptions:

1) The genotypes of the considered plant stand are randomly distributed over the field.

2) Independence of the sources of variation from (1).

3) Each plot is a random sample of all the genotypes of the considered plant stand.

Derivation of the system of non-linear equations for estimation of the broad sense heritability

For this purpose we change the formulation of model (1) applying the commonly used symbols and we get:

$$y_{ij} = g_{ij} + c_{ij} + e_{ij} \tag{5}$$

with

y_{ij} = observed phenotypic value of the jth plant in the ith plot; $i=1,2,\ldots,d$; $j=1,2,\ldots,n^2$.

g_{ij} = competitive ability of the jth plant in the ith plot;

$c_{ij} = \sum_{v(ij)} W_{v(ij)}$ where the summation $v(ij)$ runs over the four neighbours of the jth plant in the ith plot;

e_{ij} = environmental deviation of the jth plant in the ith plot (defined as the deviation of the observed phenotypic value y_{ij} of this plant from its expected value based on the c-effects and g-effects, e.g. $e_{ij}=y_{ij}-(g_{ij}+c_{ij})$).

Some further notations used in the following text:

d_n = number of plots of size $n^2 = n \times n$ into which the considered plant population is divided.

n^2 = number of plants per plot (we always assume a plant population with regularly dispersed plants).

$y_{i.}$ = mean of the ith plot

$\bar{y}$ = total mean

$g_{i.}$, $\bar{g}$, $c_{i.}$, $\bar{c}$, $e_{i.}$ and $\bar{e}$ denote the analogous definitions for the different components.

Using model (5) we obtain the following analysis of the phenotypic variance of the plot means:

$$V_{\overline{n^2}} = \sum_{i=1}^{d_n} \frac{(y_{i.} - \bar{y})^2}{d_n - 1} \tag{6}$$

$$= \sum_{i=1}^{d_n} \frac{[(g_{i.} + c_{i.} + e_{i.}) - (\bar{g} + \bar{c} + \bar{e})]^2}{d_n - 1}$$

$$= \sum_{i=1}^{d_n} \frac{(g_{i.} - \bar{g})^2}{d_n - 1} + \sum_{i=1}^{d_n} \frac{(c_{i.} - \bar{c})^2}{d_n - 1} +$$

$$+ \sum_{i=1}^{d_n} \frac{(e_{i.} - \bar{e})^2}{d_n - 1} + 2 \sum_{i=1}^{d_n} \frac{(g_{i.} - \bar{g})(c_{i.} - \bar{c})}{d_n - 1}$$

$$+ 2 \sum_{i=1}^{d_n} \frac{(g_{i.} - \bar{g})(e_{i.} - \bar{e})}{d_n - 1} + 2 \sum_{i=1}^{d_n} \frac{(e_{i.} - \bar{e})(c_{i.} - \bar{c})}{d_n - 1} .$$

Going over to the expectations of the different terms in this analysis of the variance of the plot means in (6) we obtain (under the assumptions mentioned above) the following results:

$$\mathcal{E} \left\{ \sum_{i=1}^{d_n} \frac{(g_{i.} - \bar{g})^2}{d_n - 1} \right\} = \frac{G}{n^2} \tag{7}$$

$$\mathcal{E} \left\{ \sum_{i=1}^{d_n} \frac{(e_{i.} - \bar{e})^2}{d_n - 1} \right\} = \frac{E}{n^{2b}} \tag{8}$$

(F. Smith's empirical law for the environmental component)

$$\mathcal{E} \left\{ 2 \sum_{i=1}^{d_n} \frac{(g_{i.} - \bar{g})(e_{i.} - \bar{e})}{d_n - 1} \right\}$$

$$= \mathcal{E} \left\{ 2 \sum_{i=1}^{d_n} \frac{(c_{i.} - \bar{c})(e_{i.} - \bar{e})}{d_n - 1} \right\} = 0 \tag{9}$$

(because of the independence assumptions).

For the two remaining expectations we obtain after some theoretical considerations and computations the formulas (Proof see: Appendix 1):

$$\mathcal{E} \left\{ 2 \sum_{i=1}^{d_n} \frac{(g_{i.} - \bar{g})(c_{i.} - \bar{c})}{d_n - 1} \right\} = \frac{4(n-1)\rho\sqrt{GC}}{n^3} \tag{10}$$

$$\mathcal{E} \left\{ \sum_{i=1}^{d_n} \frac{(c_{i.} - \bar{c})^2}{d_n - 1} \right\} = \frac{4n^2 - 6n + 2}{n^4} C . \tag{11}$$

Therefore we get the following result for the expectation of the variance of plot means:

$$\mathcal{E} \left\{ V_{\overline{n^2}} \right\} = \frac{G}{n^2} + \frac{4n^2 - 6n + 2}{n^4} C + \frac{E}{n^{2b}} + \frac{4(n-1)\rho\sqrt{GC}}{n^3} \tag{12}$$

Thus we have expressed the expectation of the variance of plot means as a function of G, C, E, n, ρ and b.

This relation (12) now can serve for an estimation of the unknown parameters G, C, E, ρ and b by setting

$$V_{\overline{n^2}} = \mathcal{E} \left\{ V_{\overline{n^2}} \right\}$$

for different plot sizes n^2 using equation (12). From this one obtains a system of simultaneous non-linear equations:

$$\begin{array}{l} V_{\overline{1^2}} = G + C + E \\ V_{\overline{2^2}} = \frac{G}{4} + \frac{3}{8} C + \frac{E}{4^b} + \frac{\rho\sqrt{GC}}{2} \\ \hdashline V_{\overline{n^2}} = \frac{G}{n^2} + \frac{4n^2 - 6n + 2}{n^4} C + \frac{E}{n^{2b}} + \frac{4(n-1)\rho\sqrt{GC}}{n^3} . \end{array} \tag{13}$$

To estimate the unknown parameters, for example the genetic variance G, for a certain plant population using this proposed method we can proceed, starting from this system of equations (13), in several different ways:

A) It is possible to consider all the parameters in (13), namely G, C, E, ρ and b, as unknowns and to estimate them from equations (13), for example by least squares method.

B) Based upon biological considerations it would probably be possible to assign a special definite numerical value to ρ. Then the unknown parameters can be estimated from (13) - for example in the manner described above: the system (13) of simultaneous non-linear equations can be solved for different values of b - for example using the method of least squares. For each such chosen b one obtains estimates of the parameters G, C and E. Then those values G_0, C_0, E_0 and b_0 are taken as estimates of the unknown parameters which best fit with the empirical values $V_{\overline{n^2}}$. This expression 'best fit' menas precisely:

$$SQ_{\text{Residual}} = \sum_n \left[V_{\overline{n^2}} - \mathcal{E} \left\{ V_{\overline{n^2}} \right\} \right]^2 = \text{Minimum} .$$

With this procedure b varies within a certain interval by increments of a certain length - both interval length and increment length, must be given and defined before.

C) Another possibility is to apply the estimation procedure (B), described above, not only once to a certain (from biological considerations) given ρ, but to many ρ-values; for this ρ varies in the interval $-1 \leqslant \rho \leqslant +1$, with a suitable increment length. For each of these ρ-values we apply (B) and obtain estimates for G, C, E and b for each ρ-value - with a certain corresponding SQ-Residual. Finally, those values G_0, C_0, E_0, b_0 and ρ_0 could be taken for estimates of the unknown parameters for which this SQ-Residual is a minimum.

By some further considerations it is possible to reduce the number of unknown parameters which must be estimated; therefore the computational procedure can be reduced and the estimation-procedure be improved by this modification.

If we introduce the following notations

$$G^* = G/G+C+E; \quad C^* = C/G+C+E; \quad E^* = E/G+C+E \qquad (14)$$

then we have $h^2_{b.s.} = G^*$ and $G^* + C^* + E^* = 1$ with $0 \leqslant G^* \leqslant 1$, $0 \leqslant C^* \leqslant 1$ and $0 \leqslant E^* \leqslant 1$.

Because of $\mathcal{E}\left\{V_{1^2}\right\} = G + C + E$ we obtain from (12) with (14):

$$\frac{\mathcal{E}\left\{V_{n^2}\right\}}{\mathcal{E}\left\{V_{1^2}\right\}} = \frac{G^*}{n^2} + \frac{4n^2-6n+2}{n^4}C^* + \frac{E^*}{n^{2b}} + \frac{4(n-1)\rho\sqrt{G^*C^*}}{n^3}. \qquad (15)$$

If the total number N of plants in the considered plant population is high enough, and if there exists no positive and too large a kurtosis γ of the frequency distribution of the phenotypic values y_{ij} of this plant stand, then the following approximate result is valid:

$$\mathcal{E}\left\{\frac{V_{n^2}}{V_{1^2}}\right\} \cong \frac{\mathcal{E}\left\{V_{n^2}\right\}}{\mathcal{E}\left\{V_{1^2}\right\}}. \qquad (16)$$

This result (16) follows from the unequality, which can be derived relatively simply from some theoretical investigations:

$$\mathcal{E}\left|\left\{\frac{V_{n^2}}{V_{1^2}}\right\} - \frac{\mathcal{E}\left\{V_{n^2}\right\}}{\mathcal{E}\left\{V_{1^2}\right\}}\right| \leqslant \sqrt{\frac{2}{N-1} + \frac{\gamma}{N}}. \qquad (17)$$

For proof, see: Appendix 2.

After applying (16) and putting the relation $E^* = 1 - G^* - C^*$ into (15) we obtain the equation:

$$\mathcal{E}\left\{\frac{V_{n^2}}{V_{1^2}}\right\} = \frac{G^*}{n^2} + \frac{4n^2-6n+2}{n^4}C^* + \frac{1-G^*-C^*}{n^{2b}} + \frac{4(n-1)\rho\sqrt{G^*C^*}}{n^3} \qquad (18)$$

If we now equate for different plot sizes n^2:

$$\mathcal{E}\left\{\frac{V_{n^2}}{V_{1^2}}\right\} = \frac{V_{n^2}}{V_{1^2}}$$

we get a system of simultaneous equations, which is analogous to the system of equations (13)

$$\frac{V_{2^2}}{V_{1^2}} = \frac{G^*}{4} + \frac{3}{8}C^* + \frac{1-G^*-C^*}{4^b} + \frac{\rho\sqrt{G^*C^*}}{2}$$

$$\frac{V_{3^2}}{V_{1^2}} = \frac{G^*}{9} + \frac{20}{81}C^* + \frac{1-G^*-C^*}{9^b} + \frac{8\rho\sqrt{G^*C^*}}{27}$$

- -

$$\frac{V_{n^2}}{V_{1^2}} = \frac{G^*}{n^2} + \frac{4n^2-6n+2}{n^4}C^* + \frac{1-G^*-C^*}{n^{2b}} + \frac{4(n-1)\rho\sqrt{G^*C^*}}{n^3}. \qquad (19)$$

This system of equations (19) can now serve for an estimation of the unknown parameters.

Compared with the system of equations (13), this system of equations (19) shows some essential advantages. Firstly, the number of parameters is reduced. Furthermore, the parameters G* and C*, which have to be estimated, vary now only within the small intervals $0 \leqslant G^* \leqslant 1$ and $0 \leqslant C^* \leqslant 1$, while in applying the system of equations (13) the parameters G and C, which have to be estimated there, can vary within intervals which are much larger. This final argument (namely the smaller intervals, in which the parameters vary by using the system of equations (19)) is of particular importance in relation to the possible application of 'direct search methods' or 'iteration methods' for estimating the parameters from such systems of equations.

To also estimate the parameters G*, C*, E*, ρ and b here in the case of system (19), it is possible to apply the possibilities A), B) and C) for parameter-estimation previously discussed for the system (13).

In the present study we have used B) and C) in the following form. When estimating the parameters by least squares method a certain SQ-Residual has to be minimized. We have evaluated this minimum in the present formulation of the computer program not by using mathematical procedures of differential calculus (SQ-Residual ' = 0), but by applying 'direct search methods'. With this, G* and C* run from 0 to 1 in increments of the length 0.02 and b runs from 0 to 3 in increments of the length 0.02. Finally, the estimate for E* is calculated with $\hat{E}^* = 1 - \hat{G}^* - \hat{C}^*$.

If estimates for G, C and E are also needed one can obtain such estimates in the following manner:

$$\hat{G} = \hat{G}^* \cdot V_{1^2} = \hat{G}^* \cdot \text{total variance}$$

$$\hat{C} = \hat{C}^* \cdot V_{1^2} = \hat{C}^* \cdot \text{total variance}$$

$$\hat{E} = \hat{E}^* \cdot V_{1^2} = \hat{E}^* \cdot \text{total variance} \qquad (20)$$

In the theory of this estimation procedure the accuracy of the estimates, for example the variances of the estimates for G*, C*, E* and b, are not investigated. Closely connected with this is the question of the stability of the solutions of the system (19), that is the problem of the stability of the estimates: how do the estimates change if one changes the original experimental data, i.e. the empirical variances of the plot means.

It is also possible that another method for solving the system (19) of non-linear equations might result in better estimates. All these questions concerning the mathematical and statistical aspect of the proposed estimation-procedure must still be investigated and clarified. The correctness and applicability of this estimation-procedure is dependent upon a lot of assumptions, such as: equal distances between the single plants, equal age of all individuals, random distribution of the genotypes of the plant population, special assumptions referred to the model of our study (model of competition (1)), validity of F. Smith's empirical law etc.

In practical applications of the proposed method of parameter-estimation using real plant populations these assumptions would hardly ever be realised.

The estimates obtained with this procedure in the present form can, therefore, be looked upon for the time being only as approximate results. But as we have seen in the Introduction even such estimates are of special importance for breeding work with perennial plant species.

Material and Results

The estimation-procedure described in the previous chapter had been applied to an extensive collection of Norway-spruce-data from Slovakia. Eleven Norway-spruce stands (80-90 years old) were each investigated by considering four characters, measured for each single tree: height, diameter, crown percentage and taper. This extensive data on Norway-spruce had obligingly been given to me by Dr. M. Holubcik, Zvolen, Czechoslovakia.

Because Dr. Holubcik and I intend to prepare another publication, in which we will give and discuss the complete results of the computations of the heritability-estimates for the eleven different plant stands and for each of the four characters, which are measured in this study. Therefore, to demonstrate the application of the proposed estimation-procedure for broad sense heritability, I will give here in this paper the results for only two of these plant stands. For these reasons I will not enter into a detailed description of the plant stands and the plant material which we used in this study because the main purpose of this paper is to present the theory and the statistical estimation-procedure, not to give a detailed discussion of the numerical results of the broad-sense-heritability estimates of the different Norway-spruce stands.

Example 1

Table 1a. Plant stand No. 51c (that is the original notation according with Dr. Holubcik's data). Variance of plot means with different plot sizes for the four characters

plot size	height	diameter	crown percentage	taper
1 x 1 = 1	9.77	47.00	148.15	142.06
2 x 2 = 4	3.65	16.29	51.33	44.72
3 x 3 = 9	2.15	8.59	27.78	23.90
4 x 4 = 16	1.46	6.24	18.96	15.10
5 x 5 = 25	1.23	6.25	14.28	13.57
6 x 6 = 36	1.12	5.36	10.65	14.52

With this data of the variances of plot menas for different plot sizes we get the following results (the parametertupel with minimal SQ-Residual has been underlined).

Table 1b. Plant stand No. 51c. Estimates of the parameters ρ, b, G* = $h^2_{b.s.}$, C* and E* and SQ-Residual for each of the four characters height, diameter, crown percentage and taper

		character 'height'			
ρ	b	G*	C*	E*	SQ-Residual
-1.00	+0.66	0.02	0.10	0.88	$1.6583 \cdot 10^{-6}$
-0.80	+0.66	0.04	0.06	0.90	$2.1488 \cdot 10^{-6}$
-0.60	+0.66	0.04	0.10	0.86	$6.6306 \cdot 10^{-7}$
-0.40	+0.66	0.06	0.10	0.84	$1.2159 \cdot 10^{-6}$
-0.20	+0.66	0.08	0.16	0.76	$5.4589 \cdot 10^{-7}$
0.00	+0.60	0.28	0.16	0.56	$6.8186 \cdot 10^{-7}$
+0.20	+0.32	0.62	0.24	0.14	$8.3330 \cdot 10^{-7}$
+0.40	+0.46	0.62	0.02	0.36	$9.7286 \cdot 10^{-6}$
+0.60	+0.58	0.36	0.00	0.64	$1.2775 \cdot 10^{-5}$
+0.80	+0.58	0.36	0.00	0.64	$1.2775 \cdot 10^{-5}$
+1.00	+0.58	0.36	0.00	0.64	$1.2775 \cdot 10^{-5}$

Using the possibility (C) for parameter-estimation (that is, selection of that parameter-tupel from the computations and results of case (B) which shows a minimum of SQ-Residual) - we get the following estimates of the unknown parameters:

$\hat{h}^2_{b.s.} = 0.08$ $\hat{C}^* = 0.16$ $\hat{E}^* = 0.76$ $\hat{\rho} = -0.20$ $\hat{b} = +0.66$

		character 'diameter'			
ρ	b	G*	C*	E*	SQ-Residual
-1.00	+0.56	0.56	0.00	0.44	$1.0447 \cdot 10^{-4}$
-0.80	+0.56	0.56	0.00	0.44	$1.0447 \cdot 10^{-4}$
-0.60	+0.56	0.56	0.00	0.44	$1.0447 \cdot 10^{-4}$
-0.40	+0.56	0.56	0.00	0.44	$1.0447 \cdot 10^{-4}$
-0.20	+0.56	0.56	0.00	0.44	$1.0447 \cdot 10^{-4}$
0.00	+0.56	0.56	0.00	0.44	$1.0447 \cdot 10^{-4}$
+0.20	+0.52	0.64	0.02	0.34	$1.0202 \cdot 10^{-4}$
+0.40	0.00	0.88	0.08	0.04	$7.4463 \cdot 10^{-5}$
+0.60	0.00	0.92	0.04	0.04	$8.0534 \cdot 10^{-5}$
+0.80	+0.08	0.92	0.02	0.06	$6.1610 \cdot 10^{-5}$
+1.00	+0.02	0.94	0.02	0.04	$9.3778 \cdot 10^{-5}$

Using the possibility (C) for parameter-estimation (that is, selection of that parameter-tupel from the computations and results of case (B) which shows a minimum of SQ-Residual) - we get the following estimates of the unknown parameters:

$\hat{h}^2_{b.s.} = 0.92$ $\hat{C}^* = 0.02$ $\hat{E}^* = 0.06$ $\hat{\rho} = +0.80$ $\hat{b} = +0.08$

		character 'crown percentage'			
ρ	b	G*	C*	E*	SQ-Residual
-1.00	+0.66	0.36	0.00	0.64	$1.4506 \cdot 10^{-5}$
-0.80	+0.66	0.36	0.00	0.64	$1.4506 \cdot 10^{-5}$
-0.60	+0.66	0.36	0.00	0.64	$1.4506 \cdot 10^{-5}$
-0.40	+0.66	0.36	0.00	0.64	$1.4506 \cdot 10^{-5}$
-0.20	+0.66	0.36	0.00	0.64	$1.4506 \cdot 10^{-5}$
0.00	+0.66	0.36	0.00	0.64	$1.4506 \cdot 10^{-5}$
+0.20	+0.58	0.56	0.02	0.42	$1.3980 \cdot 10^{-5}$
+0.40	+0.14	0.86	0.08	0.06	$9.5929 \cdot 10^{-6}$
+0.60	+0.12	0.90	0.04	0.06	$6.9019 \cdot 10^{-6}$
+0.80	+0.20	0.90	0.02	0.08	$6.2969 \cdot 10^{-6}$
+1.00	+0.66	0.36	0.00	0.64	$1.4506 \cdot 10^{-5}$

Using the possibility (C) for parameter-estimation (that is, selection of that parameter-tupel from the computations and results of case (B) which shows a minimum of SQ-Residual) - we get the following estimates of the unknown parameters:

$\hat{h}^2_{b.s.} = 0.90$ $\hat{C}^* = 0.02$ $\hat{E}^* = 0.08$ $\hat{\rho} = +0.80$ $\hat{b} = +0.20$

		character 'taper'			
ρ	b	G*	C*	E*	SQ-Residual
-1.00	+0.88	0.00	0.24	0.76	$2.7633 \cdot 10^{-6}$
-0.80	+0.80	0.02	0.10	0.88	$1.1708 \cdot 10^{-6}$
-0.60	+0.88	0.00	0.24	0.76	$2.7633 \cdot 10^{-6}$
-0.40	+0.78	0.10	0.08	0.82	$1.6270 \cdot 10^{-6}$
-0.20	+0.82	0.06	0.14	0.80	$1.7788 \cdot 10^{-6}$
0.00	+0.72	0.46	0.14	0.40	$1.2219 \cdot 10^{-6}$
+0.20	+0.88	0.00	0.24	0.76	$2.7633 \cdot 10^{-6}$
+0.40	+0.88	0.00	0.24	0.76	$2.7633 \cdot 10^{-6}$
+0.60	+0.88	0.00	0.24	0.76	$2.7633 \cdot 10^{-6}$
+0.80	+0.88	0.00	0.24	0.76	$2.7633 \cdot 10^{-6}$
+1.00	+0.88	0.00	0.24	0.76	$2.7633 \cdot 10^{-6}$

Using the possibility (C) for parameter-estimation (that is, selection of that parameter-tupel from the computations and results of case (B) which shows a minimum of SQ-Residual) - we get the following estimates of the unknown parameters:

$\hat{h}^2_{b.s.} = 0.02$ $\hat{C}^* = 0.10$ $\hat{E}^* = 0.88$ $\hat{\rho} = -0.80$ $\hat{b} = +0.80$

Example 2

Table 2a. Plant stand No. 49b (the original notation according with Dr. Holubcik's data). Variance of plot means with different plot sizes for the four characters

plot size	height	diameter	crown percentage	taper
1 x 1 = 1	17.20	38.35	165.04	126.09
2 x 2 = 4	8.54	14.75	71.21	47.21
3 x 3 = 9	7.10	9.76	42.80	27.20
4 x 4 = 16	5.59	8.23	33.15	18.15
5 x 5 = 25	5.84	6.72	32.91	16.07
6 x 6 = 36	5.57	8.14	23.58	6.43

With this data of the variances of plot means for different plot sizes we get the following results:

Table 2b. Plant stand No. 49b. Estimates of the parameters ρ, b, $G^* = h^2_{b.s.}$, C^* and E^* and SQ-Residual for each of the four characters height, diameter, crown percentage and taper

		character 'height'			
ρ	b	G*	C*	E*	SQ-Residual
-1.00	+0.00	0.00	0.80	0.20	$2.9220 \cdot 10^{-4}$
-0.80	+0.18	0.02	0.56	0.42	$2.8744 \cdot 10^{-4}$
-0.60	+0.14	0.02	0.62	0.36	$2.7703 \cdot 10^{-4}$
-0.40	+0.14	0.04	0.60	0.36	$2.9039 \cdot 10^{-4}$
-0.20	+0.06	0.02	0.72	0.26	$2.8537 \cdot 10^{-4}$
0.00	0.00	0.00	0.80	0.20	$2.9220 \cdot 10^{-4}$
+0.20	0.00	0.00	0.80	0.20	$2.9220 \cdot 10^{-4}$
+0.40	0.00	0.00	0.80	0.20	$2.9220 \cdot 10^{-4}$
+0.60	0.00	0.00	0.80	0.20	$2.9220 \cdot 10^{-4}$
+0.80	0.00	0.00	0.80	0.20	$2.9220 \cdot 10^{-4}$
+1.00	0.00	0.00	0.80	0.20	$2.9220 \cdot 10^{-4}$

Using the possibility (C) for parameter-estimation (that is, selection of that parameter-tupel from the computations and results of case (B) which shows a minimum of SQ-Residual) - we get the following estimates of the unknown parameters:

$\hat{h}^2_{b.s.} = 0.02$ $\hat{C}^* = 0.62$ $\hat{E}^* = 0.36$ $\hat{\rho} = -0.60$ $\hat{b} = +0.14$

		character 'diameter'			
ρ	b	G*	C*	E*	SQ-Residual
-1.00	+0.08	0.80	0.00	0.20	$5.5913 \cdot 10^{-5}$
-0.80	+0.08	0.80	0.00	0.20	$5.5913 \cdot 10^{-5}$
-0.60	+0.08	0.80	0.00	0.20	$5.5913 \cdot 10^{-5}$
-0.40	+0.08	0.80	0.00	0.20	$5.5913 \cdot 10^{-5}$
-0.20	+0.08	0.80	0.00	0.20	$5.5913 \cdot 10^{-5}$
0.00	+0.08	0.78	0.02	0.20	$3.9976 \cdot 10^{-5}$
+0.20	+0.02	0.82	0.02	0.16	$3.4539 \cdot 10^{-5}$
+0.40	+0.08	0.80	0.00	0.20	$5.5913 \cdot 10^{-5}$
+0.60	+0.08	0.80	0.00	0.20	$5.5913 \cdot 10^{-5}$
+0.80	+0.08	0.80	0.00	0.20	$5.5913 \cdot 10^{-5}$
+1.00	+0.08	0.80	0.00	0.20	$5.5913 \cdot 10^{-5}$

Using the possibility (C) for parameter-estimation (that is, selection of that parameter-tupel from the computations and results of case (B) which shows a minimum of SQ-Residual) - we get the following estimates of the unknown parameters:

$\hat{h}^2_{b.s.} = 0.82$ $\hat{C}^* = 0.02$ $\hat{E}^* = 0.16$ $\hat{\rho} = +0.20$ $\hat{b} = +0.02$

		character 'crown percentage'			
ρ	b	G*	C*	E*	SQ-Residual
-1.00	+0.52	0.24	0.00	0.76	$1.3660 \cdot 10^{-4}$
-0.80	+0.52	0.24	0.00	0.76	$1.3660 \cdot 10^{-4}$
-0.60	+0.52	0.24	0.00	0.76	$1.3660 \cdot 10^{-4}$
-0.40	+0.52	0.24	0.00	0.76	$1.3660 \cdot 10^{-4}$
-0.20	+0.52	0.24	0.00	0.76	$1.3660 \cdot 10^{-4}$
0.00	+0.52	0.24	0.00	0.76	$1.3660 \cdot 10^{-4}$
+0.20	+0.52	0.24	0.00	0.76	$1.3660 \cdot 10^{-4}$
+0.40	+0.30	0.62	0.10	0.28	$1.1994 \cdot 10^{-4}$
+0.60	+0.02	0.78	0.12	0.10	$6.6807 \cdot 10^{-5}$
+0.80	+0.06	0.82	0.06	0.12	$6.5645 \cdot 10^{-5}$
+1.00	+0.06	0.84	0.04	0.12	$6.4147 \cdot 10^{-5}$

Using the possibility (C) for parameter-estimation (that is, selection of that parameter-tupel from the computations and results of case (B) which shows a minimum of SQ-Residual) - we get the following estimates of the unknown parameters:

$\hat{h}^2_{b.s.} = 0.84$ $\hat{C}^* = 0.04$ $\hat{E}^* = 0.12$ $\hat{\rho} = +1.00$ $\hat{b} = +0.06$

character 'taper'

ρ	b	G*	C*	E*	SQ-Residual
-1.00	+0.66	0.16	0.00	0.84	$6.1488 \cdot 10^{-6}$
-0.80	+0.66	0.16	0.00	0.84	$6.1488 \cdot 10^{-6}$
-0.60	+0.68	0.04	0.02	0.94	$2.9946 \cdot 10^{-6}$
-0.40	+0.68	0.04	0.06	0.90	$3.4428 \cdot 10^{-6}$
-0.20	+0.68	0.06	0.06	0.88	$1.4566 \cdot 10^{-6}$
0.00	+0.68	0.10	0.10	0.80	$6.6942 \cdot 10^{-7}$
+0.20	+0.54	0.50	0.16	0.34	$6.1855 \cdot 10^{-7}$
+0.40	+0.30	0.76	0.10	0.14	$5.5223 \cdot 10^{-6}$
+0.60	+0.66	0.16	0.00	0.84	$6.1488 \cdot 10^{-6}$
+0.80	+0.66	0.16	0.00	0.84	$6.1488 \cdot 10^{-6}$
+1.00	+0.66	0.16	0.00	0.84	$6.1488 \cdot 10^{-6}$

Using the possibility (C) for parameter-estimation (that is, selection of that parameter-tupel from the computations and results of case (B) which shows a minimum of SQ-Residual) - we get the following estimates of the unknown parameters:

$$\hat{h}^2_{b.s.} = 0.50 \quad \hat{C}^* = 0.16 \quad \hat{E}^* = 0.34 \quad \hat{\rho} = +0.20 \quad \hat{b} = +0.54$$

Conclusions

In the present paper a new method for estimating broad sense heritability is developed as a generalization and improvement of the method of Sakai and Mukaide (1967); this itself is a generalization and improvement of Sakai and Hatakeyama's (1963) modification of Shrikhande's (1957) method to separately estimate the genetic variance, competitional variance and environmental variance of a plant population. Using this proposed new estimation-procedure one obtains estimates for broad sense heritability which are more correct and realistic - because of the more realistic assumptions proposed in this method - than the results of former authors who used Shrikhande's method or some modification of that method.

In the theory of this estimation-procedure the accuracy of the estimates as well as the problem of the stability of the solutions of the system of non-linear equations, which must be solved to obtain estimates of the unknown parameters, have not been investigated and computed until now.

This proposed method for estimating broad sense heritability has the advantage that it is possible to obtain estimates for genetic variance resp. broad sense heritability without performing any crosses. This is of special importance in working with plant species of large generation intervals, for example with perennnial plant species or even forest trees, where the usual estimation-procedure (covariances between relatives, regression techniques like the parent-offspring-regression) are often not sufficient. Mostly it is necessary or important to get heritability estimates in the early stages of breeding work - and this would be the case even if it were possible to obtain only some rough estimates or approximate results of broad sense heritability, because it is not possible to wait for the results from appropriate progeny tests.

For this situation, Shrikhande's method and therefore our proposed method have been developed. Because the proposed method depends on many simplifying assumptions, it stands to reason that such estimates can only be considered as approximate results.

Of course one obtains better and more exact estimates for genetic variance resp. broad sense heritability by performing crosses, i.e. by using progeny tests and analysing them including the competitive effects. But our estimation-procedure gives the estimates immediately - and that is a very decisive point when working with forest trees.

For the reasons mentioned above, this proposed estimation-procedure on no account can or should replace the performance of appropriate progeny tests, which give better and more accurate estimates of genetic variance resp. broad sense heritability.

The estimation-procedure theoretically derived in this paper had been appliced by us to an extensive collection of Norway-spruce data from Slovakia. Eleven Norway-spruce stands (80-90 years old) were each investigated by considering four characters, measured for each single tree,: height, diameter, crown percentage and taper. The most uniform results were found for the character 'crown percentage', where the broad sense heritability for all eleven stands was uniformly high (0.62-0.94). The character 'diameter' too shows uniformly high estimates for broad sense heritability (0.64-0.94) with the exception of three stands where $h^2_{b.s.}$ was smaller than 10 % and one stand with $h^2_{b.s.} = 0.40$. In the two remaining characters, 'height' and 'taper', we found high values as well as very low values for the broad sense heritability - but the lower values for broad sense heritability prevailed in both characters ('height':

four stands with 0.68-0.90, five stands with 0.00-0.16 and two stands with 0.22 resp. 0.48; 'taper': four stands with 0.84-0.92, five stands with 0.02-0.12 and two stands with 0.34 resp. 0.50).

Acknowledgement

I am indebted to Mr. D. Krusche for performing the numerical calculations at the computing center of the University of Kiel. I am also grateful to Dr. M. Holubcik for giving me the Norway-spruce data and finally I gratefully acknowledge support by the 'Institute of Forest Genetics and Forest Tree Breeding in Schmalenbeck'.

Literature

Hühn, M. (a): Untersuchungen zur Konkurrenz zwischen verschiedenen Genotypen in Pflanzenbeständen. I-V Silvae Genetica 18 (5/6), 186-192 (1969); 19 (1), 22-31 (1970); 19 (2/3), 77-89 (1970); 19 (5/6), 151-164 (1970); 20 (5/6), 218-220 (1971)

Hühn, M. (b): The competitive environment and its genetic reaction variations. Papers presented at the second meeting of the working group on quantitative genetics (IUFRO-section 22) August 18-19, 1969, Raleigh/North Carolina, USA; published by the Southern Forest Experiment Station, Forest Service, U.S. Department of Agriculture, New Orleans, Louisiana, 62-86 (1970)

Hühn, M. (c): Untersuchungen zur Konkurrenz in Mischbeständen aus n Komponenten. I-III Acker-Pflanzenbau 135 (1), 57-77 (1972); 135 (2), 85-106 (1972); 136 (1), 14-24 (1972)

Kedharnath, S.; Chetty, C.K.R.; Rawat, M.S.: Estimation of genetic parameters in teak (*Tectona grandis*) without raising progeny. Ind. Forest. 95, 238-245 (1969)

Morgenstern, D.: Einführung in die Wahrscheinlichkeitsrechnung und mathematische Statistik. (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen Band 124). Berlin, Heidelberg, New York: Springer-Verlag 1968

Morgner, F.; Horn, W.: Biometrisch-genetische Untersuchungen an Klonen und Nachkommenschaften bei *Phlox paniculata* L. Theor. Appl. Genet. 40, 130-137 (1970)

Namkoong, G.; Squillace, A.E.: Problems in Estimating Genetic Variance by Shrikhande's Method. Silvae Genetica 19 (2/3), 74-77 (1970)

Sakai, K.I.; Hatakeyama, S.: Estimation of genetic parameters in forest trees without raising progenies. Silvae Genetica 12, 152-157 (1963)

Sakai, K.I.; Hayashi, S.; Mukaide, H.: Estimation of Genetic Parameters in Chamaecyparis Forests. Ann. Rep., Nat. Inst. of Genetics, Japan, No. 17, 27-28 (1966)

Sakai, K.I.; Mukaide, H.: Estimation of genetic, environmental and competitional variances in standing forests. Silvae Genetica 16 (5/6), 149-152 (1967)

Sakai, K.I.; Mukaide, H.; Tomita, K.: Intraspecific Competition in Forest Trees. Silvae Genetica 17, 1-5 (1968)

Shrikhande, V.J.: Some considerations in designing experiments on coconut trees. J. Ind. Soc. Agr. Stat. 9, 82-99 (1957)

Schmetterer, L.: Einführung in die mathematische Statistik, 2. Auflage. Wien/New York: Springer-Verlag 1966

Appendix 1

Proof of formulas (10) and (11)

1a) Proof of formula (10)

$$\mathcal{E}\left\{2\sum_{i=1}^{d_n}\frac{(g_{i.}-\bar{g})(c_{i.}-\bar{c})}{d_n-1}\right\} = \frac{4(n-1)\rho\sqrt{GC}}{n^3} \qquad (10)$$

The definition of 'covariance' gives:

$$\mathcal{E}\left\{2\sum_{i=1}^{d_n}\frac{(g_{i.}-\bar{g})(c_{i.}-\bar{c})}{d_n-1}\right\} = 2\,\mathrm{Kov}(g_{i.}, c_{i.}) \qquad (21)$$

and for this covariance we obtain the following expression:

$$\mathrm{Kov}(g_{i.}, c_{i.}) = \mathrm{Kov}\left(\frac{\sum_{j=1}^{n^2} g_{ij}}{n^2}, \frac{\sum_{j=1}^{n^2} c_{ij}}{n^2}\right) \qquad (22)$$

$$= \frac{1}{n^4}\,\mathrm{Kov}\left(\sum_{j=1}^{n^2} g_{ij}, \sum_{j=1}^{n^2} c_{ij}\right) \qquad (23)$$

Using the original meaning of g_{ij} and c_{ij} we get the formula:

$$\mathrm{Kov}(g_{i.}, c_{i.}) = \frac{A(n)}{n^4}\,\mathrm{Kov}(F_X, W_X) \qquad (24)$$

where: $A(n)$ = number of all possible pairs of two plants, which are direct neighbours (i.e. neighbours in the same row or neighbours in the same column in the regularly dispersed plant stand) in plots with n^2 plants per plot.

After applying simple rules of combinatorics we obtain for $A(n)$ the following explicit expression:

$$A(n) = 4n(n-1) \qquad (25)$$

and therefore with (24) we get:

$$\mathrm{Kov}(g_{i.}, c_{i.}) = \frac{4(n-1)}{n^3}\,\mathrm{Kov}(F_X, W_X) \qquad (26)$$

Because ρ denotes the correlation coefficient between the F-values and the W-values of all plants of the total considered plant stand and $V(F_X) = G$ and $4V(W_X) = C$ we obtain the following formula:

$$\rho = \frac{\mathrm{Kov}(F_X, W_X)}{\sqrt{V(F_X) \cdot V(W_X)}} = \frac{\mathrm{Kov}(F_X, W_X)}{\frac{1}{2}\sqrt{G \cdot C}} \tag{27}$$

and we get:

$$\mathrm{Kov}(F_X, W_X) = \frac{1}{2}\rho\sqrt{GC} \tag{28}$$

and - together with (26) - it follows:

$$2\,\mathrm{Kov}(g_{i.}, c_{i.}) = \frac{4(n-1)\rho\sqrt{GC}}{n^3}\,. \tag{29}$$

Formulas (21) and (29) together conclude the proof of formula (10).

1b) Proof of formula (11)

$$\mathcal{E}\left\{\sum_{i=1}^{d_n} \frac{(c_{i.} - \bar{c})^2}{d_n - 1}\right\} = \frac{4n^2 - 6n + 2}{n^4}\,C \tag{11}$$

The definition of 'variance' gives:

$$\mathcal{E}\left\{\sum_{i=1}^{d_n} \frac{(c_{i.} - \bar{c})^2}{d_n - 1}\right\} = V(c_{i.}) \tag{30}$$

and for this variance we obtain the following expression:

$$V(c_{i.}) = V\left(\frac{\sum_{j=1}^{n^2} c_{ij}}{n^2}\right) = \frac{1}{n^4}\,V\left(\sum_{j=1}^{n^2} c_{ij}\right). \tag{31}$$

The variance of the sum of these c_{ij}-values can be computed in the following way using well-known mathematical rules as well as the simplifying assumptions mentioned in the chapter Theoretical Investigations:

$$V(c_{i.}) = \frac{1}{n^4}\left\{n^2 V(c_X) + D(n)\mathrm{Kov}(c_Y, c_Z) + B(n)\mathrm{Kov}(c_U, c_M)\right\} \tag{32}$$

where:

c_X = c-value of an arbitrary plant X.

$D(n)$ = number of all possible pairs of two plants, which are direct neighbours and which are standing diagonally to one another in the regularly dispersed plant stand in plots with n^2 plants per plot (see figure 1).

c_Y, c_Z = c-values of two arbitrary plants Y and Z, which are neighbours and which are standing diagonally against one another in the regularly dispersed plant stand (see figure 1).

$B(n)$ = number of all possible pairs of two plants - both in the same row or both in the same column in the regularly dispersed plant stand - which are separated from each other by one other plant standing between these two plants in plots with n^2 plants per plot (see figure 2).

c_U, c_M = c-values of two arbitrary plants U and M, which are both standing in the same row or in the same column in the regularly dispersed plant stand and which are separated from each other by one other plant standing between these two plants (see figure 2).

After applying simple rules of combinatorics we obtain for D(n) and B(n) the following explicit expressions:

$$D(n) = 4\,(n-1)^2 \tag{33}$$

$$B(n) = 4n(n-2) \tag{34}$$

Now we need explicit expressions for the covariances $\mathrm{Kov}(c_Y, c_Z)$ and $\mathrm{Kov}(c_U, c_M)$:

x	x	x	x	x	x
x	x	x	xQ	x	x
x	x	xO	xZ	xT	x
x	xN	xY	xR	x	x
x	x	xS	x	x	x
x	x	x	x	x	x

Fig. 1 Spatial arrangement of diagonal neighbours Y and Z'

$$c_Y = W_N + W_O + W_R + W_S$$
$$c_Z = W_O + W_Q + W_T + W_R \tag{35}$$

where W_X denotes the W-value of a plant X. With (35) we get:

$$\begin{aligned}\mathrm{Kov}(c_Y, c_Z) &= \mathrm{Kov}(W_N + W_O + W_R + W_S,\\ &\qquad W_O + W_Q + W_T + W_R)\\ &= \mathrm{Kov}(W_O, W_O) + \mathrm{Kov}(W_R, W_R)\\ &= 2\,\mathrm{Kov}(W_X, W_X) = 2\,V(W_X)\end{aligned} \tag{36}$$

In an analogous manner it is possible to derive an explicit expression for the second covariance $\mathrm{Kov}(c_U, c_M)$:

x	x	x	x	x	x	x
x	x	xI	x	xL	x	x
x	xH	xU	xJ	xM	xR	x
x	x	xK	x	xS	x	x
x	x	x	x	x	x	x

Fig.2 Spatial arrangement of plants U and M, which are separated from each other by one other plant

$$c_U = W_H + W_I + W_J + W_K \qquad c_M = W_J + W_L + W_R + W_S \tag{37}$$

and with (37) we get:

$$\begin{aligned}\mathrm{Kov}(c_U, c_M) &= \mathrm{Kov}(W_H + W_I + W_J + W_K,\\ &\qquad W_J + W_L + W_R + W_S)\\ &= \mathrm{Kov}(W_J, W_J) = V(W_X)\end{aligned} \tag{38}$$

With (32), (33), (34), (36), (38) and the relation $V(c_X) = 4\,V(W_X)$ we obtain the following equation:

$$V(c_{i.}) = \frac{1}{n^4}\Big\{4n^2 V(W_X) + 4(n-1)^2 \cdot 2V(W_X) + 4n(n-2)V(W_X)\Big\} \tag{39}$$

and after some algebraic simplifications in this formula (39) we get:

$$V(c_{i.}) = \frac{4\,V(W_X)}{n^4}\,\{4n^2 - 6n + 2\} \tag{40}$$

Putting $4\,V(W_X) = C$ we obtain:

$$V(c_{i.}) = \frac{4n^2 - 6n + 2}{n^4} \cdot C \tag{41}$$

which concludes - together with (30) - the proof of formula (11).

Appendix 2

For arbitrary random variables X and Y the following unequality holds (Morgenstern 1968, p.111):

$$\left|\mathcal{E}\left\{\frac{X}{Y}\right\} - \frac{\mathcal{E}\{X\}}{\mathcal{E}\{Y\}}\right| \leqslant \frac{k\sqrt{\text{Variance } Y}}{|\mathcal{E}\{Y\}|} \tag{42}$$

where k is a constant with $|X/Y| \leqslant k$ and $|a|$ denotes the absolute value of a.

Now we put $X = V_{\overline{n^2}}$ and $Y = V_{\overline{1^2}}$.

Because of $V_{\overline{n^2}} \leqslant V_{\overline{1^2}}$ (resp. $V_{\overline{n^2}}/V_{\overline{1^2}} \leqslant 1$) and $V_{\overline{n^2}} \geqslant 0$, $V_{\overline{1^2}} \geqslant 0$ we obtain from unequality (42) using k=1:

$$\left|\mathcal{E}\left\{\frac{V_{\overline{n^2}}}{V_{\overline{1^2}}}\right\} - \frac{\mathcal{E}\{V_{\overline{n^2}}\}}{\mathcal{E}\{V_{\overline{1^2}}\}}\right| \leqslant \frac{\sqrt{\text{Variance } V_{\overline{1^2}}}}{\left|\mathcal{E}\{V_{\overline{1^2}}\}\right|}\,. \tag{43}$$

For $\mathcal{E}\{V_{\overline{1^2}}\}$ we can replace the expression

$$\mathcal{E}\{V_{\overline{1^2}}\} = G + C + E\,. \tag{44}$$

We now need an explicit expression for the variance of $V_{\overline{1^2}}$. If we denote with N the total number of plants in the considered plant population and with m_4 the fourth central moment of the frequency distribution of the phenotypic values of the total number of all plants in the considered plant population the following relation holds:

$$\text{Variance of } V_{\overline{1^2}} = \frac{m_4}{N} - \frac{N-3}{N(N-1)}\left[\mathcal{E}\{V_{\overline{1^2}}\}\right]^2 \tag{45}$$

(see: Schmetterer 1966, p.159).
From (45) we obtain with (44):

$$\text{Variance of } V_{\overline{1^2}} = \frac{m_4}{N} - \frac{(N-3)(G+C+E)^2}{N(N-1)} \tag{46}$$

If we now introduce the kurtosis γ of the frequency distribution of the phenotypic values of all plants in the considered plant stand - we have the kurtosis-definition:

$$\gamma = \frac{m_4}{(G+C+E)^2} - 3 \tag{47}$$

and from (47) we can derive the following expression:

$$m_4 = (\gamma + 3)(G+C+E)^2. \tag{48}$$

If we put (48) in (46) we obtain - after some algebraic manipulations - the following relation:

$$\text{Variance of } V_{\overline{1^2}} = \left(\frac{\gamma}{N} + \frac{2}{N-1}\right)(G+C+E)^2. \tag{49}$$

From (49), (43) and (44) we obtain the unequality:

$$\left| \mathcal{E}\left\{ \frac{V_{\overline{n^2}}}{V_{\overline{1^2}}} \right\} - \frac{\mathcal{E}\left\{ V_{\overline{n^2}} \right\}}{\mathcal{E}\left\{ V_{\overline{1^2}} \right\}} \right| \leqslant \sqrt{\frac{\gamma}{N} + \frac{2}{N-1}} \qquad (17)$$

which concludes the proof of formula (17).

If the total number N of plants in the considered plant population is high enough and if there exists no positive and too large a kurtosis γ of the frequency distribution of the phenotypic values of this plant stand, then the following approximate result is valid:

$$\mathcal{E}\left\{ \frac{V_{\overline{n^2}}}{V_{\overline{1^2}}} \right\} \simeq \frac{\mathcal{E}\left\{ V_{\overline{n^2}} \right\}}{\mathcal{E}\left\{ V_{\overline{1^2}} \right\}} \, . \qquad (16)$$

Received August 3, 1974
Communicated by E. Andersson

Doz. Dr. Manfred Hühn
Institut für Pflanzenbau und Pflanzenzüchtung der Universität Kiel
Olshausenstr. 40-60
D-23 Kiel (Germany/BRD)