The coefficient C represents the total number of environments present. Assuming random sampling of environments, and that C is infinite so that $1-c/C = 1$, environments represent a random effect. At the other extreme, with $c = C$ so that $1-c/C = 0$ environments represent a fixed effect. The implications of the magnitude of C in the estimation of σ^2_g are obvious.

In practice the fixed effect situation is often approached. Environments can often represent more or less discrete physiographic entities, geochemical formations, or climatic regions, for example. A limited number of discrete environments may also result from a discontinuous distribution of forests. Moreover, it is normally impracticable to include more than a very limited number of environments in a genetic experiment with forest trees, and a common practice is to plant on sites which are believed to be representative of particular categories of environments. Hence C is typically finite without really being determinate.

For the overall analysis of variance to be strictly valid the residual variation should be reasonably homogeneous among environments. This is often far from the case *(e.g.,* BURDON, 1975), even with transformation of data. Another feature of this analysis is that differences between environments in the degree of expression of genetic variation will generate statistical interaction even if there is no tendency for group ranking to differ between sites (ROBERTSON, 1959, p. 478). Variation in expression of genetic differences at different environments may or may not be a scalar effect but, if scalar, it can only be fully overcome as such if the residual variation shows similar scalar relationships. Henceforth I will use the term *true interactions* for the component of statistical interaction which tends to alter the ranking of genotypes among environments. Only the true interactions represent any basis for dividing environments into entities for separate breeding programmes.

Regarding the role of environments in generating interactions, WRICKE (1962) *(see* MORGENSTERN and TEICH, 1969) developed a procedure for calculating the contribution to the interaction sum of squares of each environment (or, for that matter, of any genetic group), its main use being in identifying particular environments which are anomalous. The principle of orthogonal contrasts can be applied to partitioning the interaction sum of squares so as to identify groups of environments between which the major interactions occur, a refinement of this approach being the use of distance coefficients combined with cluster analysis (ABOU-EL-FITTOUH *et al.*, 1969). But since these methods all belong within the framework of conventional analysis of variance they share its limitations.

Expected genetic gains, if environments represent a fully random effect, appear to be independent of the extent to which interactions are "true" or not. In this situation the expected genetic gain (ΔG) from phenotypic selection of individuals is given by:

$$\Delta G = i\sigma_P h^2 = i\,\frac{\sigma^2_A}{\sigma_P} = i.\,\frac{\sigma^2_A}{\sqrt{\sigma^2_G + \sigma^2_{GE} + \sigma^2_e}} \qquad (1)$$

where: σ^2_e = micro-environmental variance

σ^2_{GE} = total genotype-environment interaction variance

= σ^2_{AE} (additive genetic × environment interaction variance) + σ^2_{DE} (genetic dominance × environment interaction variance)

σ^2_P = total phenotypic variance

h^2 = heritability

i = intensity of selection in standard deviations.

The expected gain from selecting parents on the basis of half-sib progeny test performances (assuming the same genetic parameters as above) is given by:

$$\Delta G = \frac{2^*i\sigma^2_g}{\sigma_{\bar{g}}} = \frac{2i^1/_4\sigma^2_A}{\sqrt{(\frac{1}{4}\sigma^2_A + \frac{1}{4}\sigma^2_{AE}/c + \sigma^2_w/nc)}} \qquad (2)$$

where:

$\sigma^2_{\bar{g}}$ = variance of means of groups (half-sib families) = $\sigma^2_g + \sigma^2_w/n$

$\sigma^2_g/\sigma^2_{\bar{g}}$ = heritability (or repeatability) of group means = $^1/_4\sigma^2_A/\sigma^2_{\bar{g}}$

σ^2_w = $\frac{3}{4}\sigma^2_A + \sigma^2_D + \sigma^2_e$

As the fixed effect situation is approached however, the prediction of genetic gain from conventional analysis of variance can be subject to bias, particularly if the expression of genetic differences varies between environments. For example, in selecting parents on the basis of progeny test performance with environments as a truly fixed effect, an unbiased gain prediction will be given by:

$$\Delta G = \frac{2i\sigma^2_g}{\sqrt{\sigma^2_g + (\sigma^2_w/\Sigma_c n)}} \qquad (3)$$

only if n varies between environments according to the representation of each environment in the forest estate.

Genetic correlation

Historical Background

Working with a forest tree species, BURDON (BURDON and LOW, 1973; BURDON, 1975, 1976; *see* SHELBOURNE, 1972) has used simple correlations between clonal means at pairs of sites for roughly evaluating the extent of clone-site interactions and the roles of individual environments. ABOU-EL-FITTOUH *et al.* (1969) had used the same procedure in a study of alternative methods of delimiting testing regions for cotton cultivars. However, FALCONER (1952, 1961) who worked with mice, had earlier propounded the concept of genetic correlation between environments in connection with defining an environment which gave optimal screening of genotypes; this was based on regarding the performance in two different environments as two distinct traits.

Genetic correlations have generally been estimated when both traits have been measured on the same individuals. Correlations obtained thus will be designated *Type A* genetic correlations. Where the two traits are measured on different individuals within genetic groups, a special case being genetic correlation between environments, the correlations will be designated *Type B* genetic correlations.

ROBERTSON (1959) developed some statistical aspects of the Type B genetic correlation, using the framework of conventional analysis of variance, *i.e.*, handling the different "traits" as a main effect. DICKERSON (1962) and YAMADA (1962) have pursued the development further. YAMADA showed how the Type B genetic correlation could be derived both from conventional analysis of variance and from using separate analysis of variance within each environment. YAMADA also showed how, in conventional analysis of variance, an adjustment could be made for differences between environments in the expression of genetic variation, in order to obtain unbiased estimates of genetic correlations. Notably, the value of a *Type* B genetic correlation is independent of whether environments represent a fixed or a

* The coefficients of 2, which will recur later, relates to the fact that the select parents would be intermated, instead of being mated with the base population in producing the families.

random effect. YAMADA did not, however, cover conventional analysis of variance with block replicates within environments, or with more than two environments at once with residual variance differing between environments.

Application of Type B genetic correlations (*see e.g.*, SCHEINBERG, 1973) appears to have been mainly in animal breeding work. Recently, however, TAI (1974) has noted that the Type B genetic correlation is applicable to the problems arising with a clonally propagated crop, potatoes, where successive generations of multiplication show differing experimental errors and differing expressions of genotypic variation.

Calculation of Genetic Correlation

This will be confined to separate analyses of variance for each environment, which for most purposes is more general, more flexible, and often simpler to compute than a single analysis of variance covering all environments. The estimation of genetic parameters within environments is not considered here.

The genetic correlation between environments x and y ($r_{g_{xy}}$) is of the form

$$r_{g_{xy}} = Cov_{g_{xy}} / (\sigma_{g_x} \cdot \sigma_{g_y}) \quad (4)$$

where: $Cov_{g_{xy}}$ = covariance for groups between the trait as it is expressed at environments x and y respectively

$\sigma^2_{g_x}$ and $\sigma^2_{g_y}$ are variances between groups at environments x and y respectively.

If the groups are half-sib families, $Cov_{g_{xy}} = {}^1/_4 Cov_{A_{xy}}$ and $r_{g_{xy}} = r_{A_{xy}}$, $Cov_{A_{xy}}$ and $r_{A_{xy}}$ being respectively the additive genetic covariance and correlation between the environments. $Cov_{g_{xy}}$ is estimated directly as the mean cross-product between environments of group means, and the estimation is independent of experimental layout parameters in the respective environments. This procedure is analogous to using the mean cross-products of progeny means on parental values for estimating σ^2_A. The only expected source of bias in estimating $Cov_{g_{xy}}$ will be the carryover of non-genetic effects common to members of a group, which arise prior to planting in different environments.

An alternative formulation of $r_{g_{xy}}$ is:

$$r_{g_{xy}} = r_{xy} / (h_{\bar{g}_x} \cdot h_{\bar{g}_y}) \quad (5)$$

where: r_{xy} = correlation between group means at x and y $= Cov_{g_{xy}} / (\sigma_{\bar{g}_x} \rho_{\bar{g}_y})$, where $\sigma^2_{\bar{g}_x}$ and $\sigma^2_{\bar{g}_y}$ are variances of group means at x and y respectively.

$h^2_{\bar{g}_x}$ and $h^2_{\bar{g}_y}$ are heritabilities of group means at x and y respectively (cf. Equation 2).

Relating to conventional analysis of variance, if $\sigma^2_{\bar{g}_x} = \sigma^2_{\bar{g}_y}$ then (cf. YAMADA, 1962, Formula 22)

$$r_{g_{xy}} = {}^2_g / (\sigma^2_g + \sigma^2_{gE})$$

and $Cov_{A_{xy}} = \sigma^2_A$

For illustration I will use results kindly provided by WILCOX. These are from a 5-year-old progeny planting (*assuming* $\sigma^2_g = {}^1/_4\ \sigma^2_A$) with 40 progenies, planted in 10-tree plots in five randomised complete blocks at each of three sites. For Sites 1 and 2 the data could be analysed, with only minor approximations, as for a fully balanced classification. At Site 3, however, 25 % of trees were missing, and the imbalance led to these data being analysed separately. Best estimates of progeny means there were obtained by adjusting the means for replicate effects using a least-squares procedure, but it was necessary to ignore interaction (or plot) effects. Variances at Site 3 were estimated by Henderson's Method I. Sites varied in the amount of experimental error, and traits varied in heritability and the amount of genotype-site interaction. Estimates of relevant within-site parameters are shown in *Table 2.*

Table 2. — Estimates of genetic parameters within progeny test sites (units immaterial), from WILCOX *et al.* (1974)

Trait		Site 1	Site 2	Site 3
Height	$\hat{\sigma}^2_g$	3.5488	7.4969	3.1128
	$\hat{\sigma}^2_{\bar{g}}$	3.5545[1]	9.4228[1]	8.6436[2]
	$\hat{h}^2_{\bar{g}}$ $(= \hat{\sigma}^2_g / \hat{\sigma}^2_{\bar{g}})$	0.6389	0.7956	0.3601[3]
	$\hat{h}^2$ $(= \frac{4\hat{\sigma}^2_g}{\hat{\sigma}^2_g + \hat{\sigma}^2_{gr} + \hat{\sigma}^2_{w'}})$	0.2523	0.4491	0.2025
Straightness	$\hat{\sigma}^2_g$	0.2133	0.2607	0.0686
	$\hat{\sigma}^2_{\bar{g}}$	0.32382[1]	0.37705[1]	0.29358[2]
	$\hat{h}^2_{\bar{g}}$ $(= \hat{\sigma}^2_g / \hat{\sigma}^2_{\bar{g}})$	0.6587	0.6914	0.2337[3]
	$\hat{h}^2$ $(= \frac{4\hat{\sigma}^2_g}{\hat{\sigma}^2_g + \hat{\sigma}^2_{gr} + \hat{\sigma}^2_{w'}})$	0.2646	0.2893	0.0662
Wood density	$\hat{\sigma}^2_g$	81.3057	67.0243	83.4902
	$\hat{\sigma}^2_{\bar{g}}$	91.6364[1]	79.0664[1]	110.8674[2]
	$\hat{h}^2_{\bar{g}}$ $(= \hat{\sigma}^2_g / \hat{\sigma}^2_{\bar{g}})$	0.8873	0.8477	0.7531[3]
	$\hat{h}^2$ $(= \frac{4\hat{\sigma}^2_g}{\hat{\sigma}^2_g + \hat{\sigma}^2_{gr} + \hat{\sigma}^2_{w'}})$	0.7716	0.5833	0.3990

σ^2_{gr} = replicates x progenies interaction (or plots) variance

$\sigma^2_{w'}$ = within-plots variance = σ^2_w

1 $\sigma^2_{\bar{g}} = \sigma^2_g + \sigma^2_{gr}/10 + \sigma^2_w/50$

2 $\sigma^2_{\bar{g}}$ = variance of unweighted family means adjusted for replicates ignoring interaction

3 Because of 2, not an exact solution

For the example the estimates of genetic correlations between sites (Eqution 4 or 5) are shown in *Table 3,* together with estimates of correlations between the progeny means. Wood density shows genetic correlations between environments in the region of unity (sampling errors can lead to estimates exceeding this value). In comparison, conventional analysis of variance showed significant interaction ($F_{39,312} = 1.74$, $P = 0.006$) between Sites 1 and 2, but this was very small ($\hat{\sigma}^2_{gE} = 0.002\ \hat{\sigma}^2_g$). Hence, with density as the criterion for selection, all sites would fall into a single breeding or seed collection zone. For height and straightness the estimates of genetic correlations are appreciably less than unity, and insofar as these departures are not attributable to sampling error they reflect true interactions. Conventional analysis of variance showed important progeny-site interactions for these traits over Sites 1 and 2. For straightness the relatively low genetic correlations between Site 3 and the others suggest that Site 3 would belong in a separate breeding zone if straightness was the main criterion for selection.

Table 3. — Estimates of convariances and correlations between environments for different traits

Parameter	Trait	Site pair 1/2	1/3	2/3
$Cov_{g_{xy}}$	Height	3.7385	1.9204	4.0501
	Straightness	0.19846	0.06857	0.04688
	Wood density	74.0056	92.1113	80.0787
r_{xy}	Height	0.5168***	0.2771 NS	0.4488**
	Straightness	0.5680***	0.2222 NS	0.1407 NS
	Wood density	0.8694***	0.9138***	0.8552***
$r_{g_{xy}}$ (= $r_{\hat{A}_{xy}}$)	Height	0.725	0.578	0.838
	Straightness	0.842	0.567	0.351
	Wood density	1.002	1.118	1.070

NS = P>0.05; ** = P< 0.01; *** = P<0.001

Genetic Gain Prediction

The following paragraphs illustrate some of the uses of genetic correlation for deriving gain expectations under alternative procedures.

Gain ($\Delta G_{y \cdot x}$) at environment y from selecting phenotypically in environment x can be predicted from the formula for correlated response to selection (cf. FALCONER, 1961).

$$\Delta G_{y \cdot x} = i\, h_x\, h_y\, r_{A_{xy}}\, \sigma^2_{P_y} = i\, h_x\, \sigma_{A_y} r_{A_{xy}} \qquad (6)$$

where: $\sigma^2_{P_y}$ = phenotypic variance at environment y

h^2_x = heritability at environment x.

Where selection of parents is done on the basis of a half-sib progeny test, h^2_x and h^2_y will represent heritabilities of progeny means ($h^2_{\bar{g}_x}$ and $h^2_{\bar{g}_y}$) and $\sigma^2_{P_y}$ will represent the variance of progeny means ($\sigma^2_{\bar{g}_y}$). For reasons given earlier $\sigma^2_{\bar{g}_y}$ in Equation 6 must be multiplied by 2 in this case.

The efficiency ($\Delta G_{y \cdot x}/\Delta G_{y \cdot y}$) of phenotypic selection at site x for planting at environment y relative to both selecting at and planting at y, with the same intensity of selection at the two sites (i.e., $i_x = i_y$) is given by (cf. FALCONER, 1961, 19.7).

$$\Delta G_{y \cdot x}/\Delta G_{y \cdot y} = r_{A_{xy}} h_x/h_y \qquad (7)$$

The conditions for indirect selection to be efficient are self-evident, but they are treated quantitatively by SEARLE (1965).

Applying Equation 6, *Table 4* shows the matrices of expected gains per unit i for selecting and planting in different combinations of sites. With selecting and planting at the same site x = y and $r_g = 1$, so the simplification of the formula is self-evident. For height and straightness greatest gains are expected when screening is done at the planting sites, and selections at Site 3 appear to be very unsuitable for planting elsewhere. Yet despite the poor genetic correlation with these sites Site 3 has given only marginally, if at all, the best screening of parents for planting locally, because of the poor resolution of progeny differences. For wood density, however, expected gains are marginally greatest from screening at Site 1, irrespective of the planting site, because of the better resolution of genetic differences at Site 1. The general pattern appears to hold for selecting parents on the basis of progeny performance or for phenotypic selection in the field.

One can readily compare the practical value of different environments for screening, by computing the expected gains if each individual environment was used for all screening. The weighted average genetic gain (ΔG_j) from screening on the basis of half-sib progeny tests at the j^{th} environment is given by:

$$\Delta G_j = 2 \sum_k i_j h_j \sigma_g\, r_{g_{jk}}\, q_k \qquad (8)$$

where: q_k denotes the proportion of the total planting to be done at the k^{th} planting environment, etc.

In the example it is reasonable to assume $q_1 = q_2 = 0.45$, $q_3 = 0.1$. The gains predicted by using each site for all screening are shown in *Table 5*. This underlines how Site 3 has been generally unsuitable for progeny testing for height and straightness because not only are heritabilities of progeny means lower there, but also it shows poor genetic

Table 5. — Predicted average genetic gains, per unit intensity of selection, from screening on progeny performance entirely at single sites

Trait	Site 1	2	3
Height	2.95	3.56	2.04
Straightness	0.676	0.688	0.217
Wood density	16.3	15.9	15.0

Table 4. — Predicted genetic gains, per unit intensity of selection, from all combinations of screening and planting sites

Trait	Planting site	A. Selection of seed parents on progeny performance — Selection site 1	2	3	B. Phenotypic mass selection — Selection site 1	2	3
Height	1	3.01	2.44	1.31	1.89	1.83	0.98
	2	3.17	4.88	2.76	1.99	3.67	2.07
	3	1.63	2.64	2.18	1.02	1.37	1.59
Straightness	1	0.75	0.65	0.25	0.48	0.42	0.14
	2	0.70	0.85	0.17	0.44	0.55	0.10
	3	0.24	0.15	0.25	0.15	0.10	0.14
Wood density	1	17.0	16.6	15.6	15.8	13.8	11.4
(Assuming $r_{xy} \neq 1$)	2	15.4	15.1	14.2	14.4	12.5	10.3
	3	17.2	16.8	15.6	16.1	14.0	11.5

correlations with the important planting sites. A good site for screening must also allow the wherewithal for a sufficient intensity of selection; in phenotypic selection this would demand adequate trees to select from. Heritability would probably be critical more in phenotypic selection, since the heritability of group means in a genetic experiment can be increased by enlarging families.

Correlation and Regression

So far, both the inter-relationships between environments and gain expectations have been considered in terms of genetic correlation, but, of course, regression analysis is also applicable. Both $Cov_{A_{xy}}/\sigma^2_{p_x}$ and $Cov_{g_{xy}}/\sigma^2_{\bar{g}_x}$ for instance, represent linear regressions. Hence, in selecting parents on the basis of half-sib progeny performance at environment x the expected response at environment y can be written:

$$\Delta G_{y \cdot x} = 2ib\sigma_{\bar{g}_x} \qquad (9)$$

where: $b = Cov_{g_{xy}}/\sigma^2_{\bar{g}_x}$

Valid regression estimates do not depend on the sample of genotypes being random, and the greater the within-sample variation the more efficient the regression estimate should be. Genetic correlation estimates tend to be biased with non-random genetic samples, but this should not be important in comparing genetic correlations among different pairs of environments.

Irrespective of whether regression or correlation analysis is used, the presence of curvilinearity can be important *(Fig. 1)*. Environment y would give better resolution of the very best genotypes than environment x, even though it may not give better resolution of all genotypic differences. For practical screening, environment y would normally be preferable.

It is a matter of convention whether a curvilinear component of the sort represented in *Fig. 1* is regarded as representing true interaction.

Sampling Errors and Significance Tests

Sampling errors of genetic correlations are notoriously difficult to specify, even approximately. The "standard error" of an estimate of a genetic correlation is strictly valid only when the correlation is normally distributed, a situation which is unlikely to be approached unless the experimental population is very large. Hence estimates of standard errors must be interpreted with caution, although they indicate the comparative reliability of different estimates of genetic correlations.

Formulae for standard errors of Type B genetic correlations have been derived by ROBERTSON (1959) for what corresponds to a half-sib progeny test, fully randomised within environments, but these formulae apply to two special cases which represent, in effect,

(i) same heritability for both traits and perfect genetic correlation;

(ii)* same heritability for both traits.

However, ROBERTSON'S formulae could readily be applied to a comparable clonal experiment, for instance. It is clear from ROBERTSON'S paper, and from the work on Type A genetic correlations *(e.g.,* TALLIS, 1959; MODE and ROBINSON, 1959; HAMMOND and NICHOLAS, 1972), that any more generalised formulations would be very cumbersome, for what they would be worth. Overall, it appears that the direct study of standard errors is unrewarding.

Some other approaches to sampling errors with Type B genetic correlations are simpler and appear to be at least as informative, although further statistical development is needed.

The test for whether there is a correlation at all, that is against the null hypothesis that $r_g = 0$, is provided simply by the significance of the regression (or correlation) of group means at one environment on means at the other.

A more relevant test would generally be for departures from perfect genetic correlation (null hypothesis, $r_g = 1$), in other words for true interactions. One approach to this, although not a direct solution, is testing for lack of fit in the regression (be it linear or curvilinear) of group means at one environment (y) on means at the other (x). It will be necessary to allow for the fact that both variables are subject to measurement error (sampling error), although estimates for these errors will be available.

Where the regression (or correlation) is linear and positive, a suggested test is as follows (cf. FREUND, 1971):

Source	Degrees of freedom	Expected mean square	F
(i) Lack of fit	a — 2	$\sigma^2_{\varepsilon',y} + \sigma^2_L$	(i)/(ii)
(ii) Residual sampling error	N	$\sigma^2_{\varepsilon',y}$	

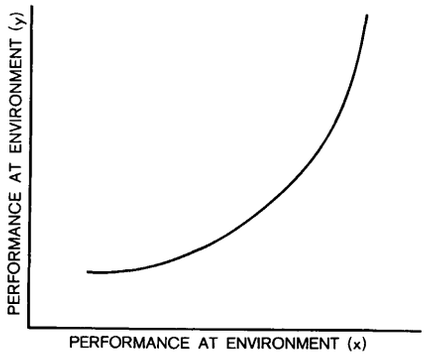

Fig. 1. — Hypothetical case of curvilinear relationship between group means at two environments

where lack of fit m.s. = error m.s. after fitting regression of group means

σ^2_L = variance deriving from true departures from the regression

$\sigma^2_{\varepsilon',y}$ = component of the variance of group means for the "dependent" (y) variable which derives purely from sampling errors in the group means at both environments

N = (conservatively) denominator degrees of freedom in test of group differences at x or y, whichever number is smaller. The value, however, would normally be large.

A formula for $\sigma^2_{\varepsilon',y}$ can be derived (solution — I.A. Andrew) from Formula 6.17.1 of SNEDECOR and COCHRAN (1967) after substituting $(\sigma^2_{Y'} - \sigma^2_{\varepsilon',y})$ for σ^2_Y. The formula is:

$$\sigma^2_{\varepsilon y} = \frac{\sigma^2_{\varepsilon x}}{(\sigma^2_{X'} - \sigma^2_{\varepsilon x})} \cdot (\sigma^2_{Y'} - \sigma^2_{Y'.X}) + \sigma^2_{\varepsilon y} \qquad (10)$$

* The formula as published (p. 479) is actually incorrect. The numerator of the second term of the right-hand side should read $2r^2_g(1-t)^2$ (A. ROBERTSON, pers. comm.).

where σ^2_X, = variance of group means at environment x ($= \sigma^2_{\bar{g}_x}$)

σ^2_Y, = variance of group means at environment y ($= \sigma^2_{\bar{g}_y}$)

$\sigma^2_{\varepsilon_x}$ = sampling variance of group means at x ($= \sigma^2_{\bar{g}_x} - \sigma^2_{g_x}$)

$\sigma^2_{\varepsilon_x}$ = sampling variance of group means at y ($= \sigma^2_{\bar{g}_y} - \sigma^2_{g_y}$)

$\sigma^2_{Y'.X'}$ = m.s. of departures from regression of group means ($=\sigma^2_{\bar{g}}$ (a — 1)/(a — 2))

$$\text{Hence } \sigma^2_{\varepsilon'_y} = \frac{\sigma^2_{\bar{g}_x} - \sigma^2_{g_x}}{\sigma^2_{g_x}} \left\{ \sigma^2_{\bar{g}_y} - \sigma^2_{\bar{g}_y} \left(1 - r^2 \left(\frac{a-1}{a-2}\right)\right) \right\} + \left(\sigma^2_{\bar{g}_y} - \sigma^2_{g_y}\right) \quad (11)$$

The power of this test will vary according to which of two environments is designated x and which y. What is really being tested is whether screening at x for merit at y is subject to significant error, and the significance of such error will depend *inter alia* on the significance of group differences at the designated y.

A measure as distinct from a test of the lack of fit would be the ratio $\sigma^2_L/\sigma^2_{g_y}$, which appears to be essentially independent of which of two environments is designated x or y.

Table 6. — Tests for lack of fit in regressions of progeny means at another, and estimates of the degree of lack of fit

	Trait	Site pair 1/2	2/1	1/3	3/1	2/3	3/2
$F_{30,156}$	Height	1.771**	2.184***	2.085***	1.415	1.561*	1.194
	Straightness	1.456	1.490*	2.514***	1.252	3.568***	1.308
	Wood density	1.015	1.012	0.459	0.480	0.707	0.731
$\sigma^2_L/\sigma^2_{g_y}$	Height	0.515		0.772		0.370	
	Straightness	0.751		0.841		1.01	
	Wood density	0.0055		<0		<0	

* P<0.05; ** P<0.01; *** P<0.001

Table 6 shows results of the above test and estimates of $\sigma^2_L/\sigma^2_{g_y}$ for the example given. For wood density the results of the tests merit comment. The F ratios of $\geqq 1$ for Sites 1 and 2 are not actually consistent with $r_g \leqq 1$, but the discrepancy is small and is readily attributed to minor approximations in the original analyses. For Sites 2 and 3 the F ratios are significantly ($P < 0.01$) less than unity, indicating that r_g exceeds unity by more than sampling error. However, this is not considered important, because r_g is not much greater than unity, while estimation procedures for Site 3 were only approximate. In one case $\hat{\sigma}^2_L/\hat{\sigma}^2_{g_y} > 1$, even though $\hat{r}_g > 0$. This reflects $F < 1$ for the regression mean square versus the mean square for departures from the regression.

A different approach again is to estimate the standard error (σ_b) of the regression (b) of group means at one environment on means at another. A linear regression is symmetrically distributed and its standard error is easily estimated. A useful measure of precision is b/σ_b (= Student's t).

The expected precision of estimating regressions, in relation to progeny size, is shown in *Fig. 2* for two special cases. Both assume half-sib families, $h^2 = 0.2$ at both environments, and a total of 2000 fully randomised individuals at each environment, but in one case $r^2_A = 1$ is assumed, in the other $r^2_A = 0.5$. *Figure 2* can be compared with Fig. 1 of ROBERTSON (1959) which shows the sampling variance of the Type B genetic correlation for different family sizes under the same general conditions. Optimal progeny sizes (reflected in minimum σ_b/b) are similar to those for estimating the correlations, but the precision of regression estimates is reduced less by sub-optimal progeny size than is the precision of correlation estimates. In practice this means that gain predictions appear to be less affected by families being too small than is the detection of true interactions.

Extension to the Multi-Trait Situation

So far only single metric traits have been considered. Yet the genetic correlation approach could be extended to the multi-trait situation by using a selection index which combines several traits into a single composite trait. This is because it is possible to formulate a heritability for the index rating, either at the level of individuals or of group means (CUNNINGHAM, 1972), although it is likely that only with group means will the estimate of the "heritability" be high enough to give reliable estimates of the correlation or expected gains at other sites.

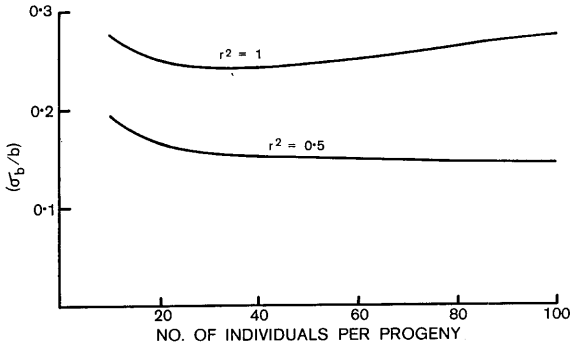

Fig. 2. — Expected precision of estimating regressions of half-sib family means at one environment on corresponding means at another environment, for differing family sizes ($h^2 = 0.2$ at both environments, σ^2_P the same at both, 2000 individuals per environment).

A problem does arise in deciding what index should be applied. Within each environment a somewhat different index would be optimal, although the construction of one index using information from all environments should not depend on conventional analysis of variance. An empirical approach is suggested, trying several alternative indices. It may be appropriate to compare gain from selection in any one environment using the locally optimal index with gain from selecting in other environments using indices which are optimal at the latter (cf. SINGH and BELLMANN, 1972).

Complex Mating Designs

Where complex mating designs (e.g., North Carolina II, diallel, half diallel, partial diallel) are used to study the interaction of non-additive as well as additive gene effects with environment, the situation is more complicated, yet the principles still appear to apply. The correlations and regressions between environments can be calculated for both the full-sib family means and the half-sib family (array) means, although the computation of repeatability for the latter would be more elaborate. Any general differences in correlations or regressions between full-sib and half-sib means would mean that non-additive gene effects interact with environments in a different way from

additive effects. Alternatively, one could estimate the specific combining ability (s.c.a.) effect for each cross, together with the repeatability at each site of such effects; this would permit the estimation of s.c.a. correlations between environments in a manner analogous to that for the estimation of genetic correlations.

With complex mating designs the conventional analysis of variance would probably remain the method of choice for evaluating interactions of s.c.a. with environment. Yet the genetic correlation approach appears feasible if the data rendered the overall analysis of variance unsatisfactory.

Further Applications

Here the Type B genetic correlation has been applied specifically to where the same trait is expressed in different environments, but the potential application is much wider. It can clearly be used when one trait is measured in one environment and another trait in a different environment. Even within a single environment it may not be practicable to measure the same groups of individuals for each trait; for instance, if one trait must be assessed destructively on young trees while the other trait must be measured on older trees. The Type B genetic correlation has been applied with animals largely because certain pairs of traits are necessarily expressed in different sexes and therefore in different individuals. Where vegetative propagation is practised over several cycles the clonal correlation (of Type B) between repropagation cycles will provide a check on whether repropagation affects the long-term validity of initial selection. Still other applications doubtless exist.

Advantages of the genetic correlation concept

The concept of genetic correlation between environments has major advantages in research strategy. It is readily directed at the question of the role of environments in generating interactions. As such it can be applied not only in defining the regions or site categories which warrant separate breeding programmes, but also in defining which sites within a particular grouping are optimal for phenotypic selection or progeny testing. For predicting genetic gains it can be more satisfactory than conventional estimates of genotype-environment interaction variance. It also avoids certain problems posed by the question of whether environments conform to a fixed or random effect.

In the actual analysis of data there can be some important advantages. Differences between sites in the amount of experimental error present no problems in correlation analysis, although they could bias analysis of variance involving all sites. With correlation analysis it is only necessary to conduct analysis of variance with one site at a time, which can be an important advantage with very large and cumbersome experiments. The more complex an experimental design the more troublesome any imbalance of classification tends to be, so this problem should also be mitigated by analysing sites singly. Where traits have been scored or measured on different bases at different sites, the genetic correlation is an attractive method of linking the information. Provided the same genetic groups are represented there are no problems in linking information from different sites where the layout parameters are totally different, yet with conventional analysis of variance this could be extremely difficult, if not impossible.

This last factor has a very important corollary. When a genetic experiment is planted at several sites there is no need to use the same sort of layout at all sites. Sites can differ in the pattern and texture of microsite variation so that widely differing layouts can be appropriate, depending on the site. One can lay the experiment out according to the requirements at each site rather than having to compromise.

Even where conventional analysis of variance can be used to detect interactions, genetic correlation analysis could be a very useful followup for detailed interpretation.

In favourable circumstances the conventional analysis of variance and estimation of genotype interaction variance will be the simplest approach, particularly with complex mating designs. But the situation is often far from ideal, so it appears overall that genetic correlation is the more powerful and flexible concept.

Acknowledgements

Thanks are due to Drs. C. J. A. Shelbourne and M. D. Wilcox for reviewing a draft, and also to Drs G. Namkoong and A. C. Matheson for some suggestions.

Summary

Variation in the relative performance of genotypes according to environment has been traditionally formulated as genotype-environment interaction, which has normally been studied quantitatively in analysis of variance with environments as a main effect.

As a statistical concept, however, genotype-environment interaction variance can have important disadvantages. Its application depends largely on the overall analysis of variance being valid and readily carried out, two conditions which do not always apply. It also has other limitations for characterising the role of different environments in influencing phenotypes.

For forest tree breeding, it is contended that primary attention should normally be given to the role of environments rather than of individual genotypes in generating interactions. This has a bearing on the appropriate method of analysis.

An alternative concept for dealing with such interaction derives from regarding performance in each environment as a distinct trait: the relative performance of genotypes for growth rate, for instance, at two different environments can be expressed as a genetic correlation between their growth rates at the two environments. Such genetic correlations are easily calculated, and their estimation, some of their statistical properties, and their applications are presented. Genotype-environment interactions of the type which could alter rankings of genotypes among environments are reflected in departures from perfect genetic correlation.

Genetic correlations can be used in various situations where analysis of variance involving environments is unsatisfactory, and remove certain constraints in experimental layout. They have advantages in characterising environments, and in predicting genetic gains as a basis for decision making, which can help not only in regionalising breeding programmes, but in deciding which particular environments give the most efficient screening of genotypes. The approach is especially useful in the absence of any obvious framework of geographic adaptation, as could be the case with exotic species. Overall, genetic correlation between environments represents a powerful and flexible statistical concept.

Key words: Genotype-environment interaction, genetic correlation, gain prediction, progeny testing, regional breeding programmes.

Zusammenfassung

Als Alternative zur herkömmlichen Schätzung von Genotyp × Umwelt — Interaktionen mittels Varianzanalyse wird ein Konzept vorgestellt, bei dem die relative Merkmals-

prägung von Genotypen unter verschiedenen Umwelteinflüssen als genetische Korrelation aufgefaßt wird. Im Vergleich zur varianzanalytischen Methode bietet dieses Konzept viele Vorteile: Es ist leichter anwendbar, erfordert weniger Voraussetzungen und ist vielseitiger verwendbar.

Literature Cited

Abou-el-Fittouh, H. A., Rawlings, J. O. and Miller, P. A.: Classification of environments to control genotype by environmental interactions with an application to cotton. Crop Sci. 9: 135—40 (1969). — Breese, E. L.: The measurement and significance of genotype-environment interactions in grasses. Heredity 24: 27—44 (1969). — Burdon, R. D.: Compression wood in *Pinus radiata* clones on four different sites. New Zealand J. For. Sci. 5 (2): 152—64 (1975). — Burdon, R. D.: Foliar macronutrient concentrations and foliage retention in radiata pine clones on four sites. New Zealand J. For. Sci. 5 (3): 250—59 (1976). — Burdon, R. D. and Low, C. B.: Effects of site on expression of cone characters in radiata pine. New Zealand J. For. Sci. 3: 110—19 (1973). — Cunningham, E. P.: Theory and application of statistical selection methods. XIVth Brit. Poultry Breeders Round Table, Birmingham, pp. 1—38 (1972). — Dickerson, G. E.: Implications of genetic environmental interaction in animal breeding. Animal Production 4: 47—64 (1962). — Eberhart, S. A. and Russell, W. A.: Stability parameters for comparing varieties. Crop Sci. 6: 36—40 (1966). — Falconer, D. S.: The problem of environment and selection. American Naturalist 86: 293—98 (1952). — Falconer, D. S.: An Introduction to Quantitative Genetics (amended 1st Ed). Oliver and Boyd, Edinburgh (1961). — Finlay, K. W. and Wilkinson, G. N.: The analysis of adaptation in a plant breeding programme. Austral. J. Agric. Res. 14: 742—54 (1963). — Freeman, G. H.: Statistical methods for the analysis of genotype-environment interactions. Heredity 31: 339—54 (1973). — Freeman, G. H. and Perkins, J. M.: Environmental and genotype-environmental components of variability. VIII. Relations between genotypes grown in different environments and measures of these environments. Heredity 27: 15—23 (1971). — Freund, R. J.: Some observations on regressions with grouped data, Amer. Statistician 25: 29—30 (1971). — Hammond, K. and Nicholas, F. W.: The sampling variance of the correlation coefficients estimated from two-fold nested and off-spring-parent regression analyses. Theoret. Appl. Genet. 42: 99—100 (1972). — Perkins, J. M. and Jinks, J. L.: Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. Heredity 23: 339—56 (1968 a). — Perkins, J. M. and Jinks, J. L.: Environmental and genotype-environmental components of variability. IV. Non-linear interactions for multiple inbred lines. Heredity 23: 525—35 (1968 b). — Mode, C. G. and Robinson, H. F.: Pleiotropism and the genetic variance and covariance. Biometrics 15: 518—37 (1959). — Morgenstern, E. K. and Teich, A. H.: Phenotypic stability of height growth of Jack pine provenances. Canad. J. Genet. Cytol. 11: 110—17 (1969). — Pederson, D. G.: The stability of varietal performance over years. 2. Analysing variety trials. Heredity 33: 217—28 (1974). — Robertson, A.: The sampling variance of the genetic correlation coefficient. Biometrics 15: 469—85 (1959). — Scheinberg, E.: The design and analysis of genetic experiments to modify genotype-environment interactions. I. Single pair matings and selection using one simple quantitative attribute measured in two environments. Canad. J. Genet. Cytol. 15: 635—45 (1973). — Searle, S. R.: The value of indirect selection: 1. Mass selection. Biometrics 21: 682—707 (1965). — Shelbourne, C. J. A.: Genotype-environment interaction: its study and its implications in forest tree improvement. IUFRO Genet. — SABRAO Joint Symp., Tokyo, October 1972 (1972). — Singh, R. K. and Bellmann, K.: Problems of generalisation of selection indices. Theoret. Appl. Genet. 42: 331—34 (1972). — Snedecor, G. W. and Cochran, W. G.: Statistical Methods, 6th Ed. Iowa Sta. Univ. Press, Ames (1967). — Tallis, G. M.: Sampling errors of genetic correlation coefficients calculated from analyses of variance and covariance. Austral. J. Stats 1: 35—43 (1959). — Tai, G. C. C.: A method for quantitative genetic analysis of early clonal generations of an asexual crop with special application to breeding populations of the potato. Theoret. Appl. Genet. 45: 150—6 (1974). — Wilcox, M. D., Firth, A., Birt, D. V. and Low, C. B.: Preliminary assessment at age 5 years of the 1968 open-pollinated progeny test of *P. radiata* select parents (268 series). N.Z. For. Serv., For. Res. Inst., Genet. & Tree Impr. Internal Rep. 54 (unpubl.) (1974). — Wricke, G.: Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen. Z. Pflanzenzucht. 47: 92—96 (1962). — Yamada, Y.: Genotype by environment interaction and genetic correlation of the same trait under different environments. Japan. J. Genet. 37: 498—509 (1962).