Theor Appl Genet (1990) 80:11–16

Relationships between genotype × environment interaction and genetic correlation of the same trait measured in different environments

Y. Itoh and Y. Yamada *

Department of Animal Science, Faculty of Agriculture, Kyoto University, Kyoto 606, Japan

Received August 25, 1989; Accepted March 6, 1990
Communicated by J. S. F. Barker

Summary. Relationships between genotype × environment interaction and genetic correlation of the same trait measured in different fixed environments are derived by comparing the variance-covariance structures of observations between a one-way multiple-trait linear model and a two-way single-trait mixed linear model. In the latter model, heterogeneity of interaction variances among environments and non-zero covariances among interactions are assumed, in addition to the heterogeneity of error variances and non-zero covariances between genetic-group effects and interactions that were accommodated in earlier work. The results are applicable to more than two environments and to unbalanced data. This paper is a generalization and a correction of earlier works.

Key words: Genotype-environment interaction – Genetic correlation – Genetic parameters – Unbalanced classification – Heterogeneous variances

Introduction

An animal or plant that performs well in one environment does not necessarily perform well in another environment, if there is genotype × environment interaction. Falconer (1952) extended the concept of the genetic correlation between two traits to the genetic correlation between measurements of the same trait expressed in two different environments in the study of selection responses, and pointed out that the genetic correlation (or, rather, its departure from +1) might be an appropriate measure of genotype × environment interaction when only two environments are involved. Robertson (1959) derived the relationship between genetic correlation and genotype × environment interaction under the condition of the same heritability in two environments. Dickerson (1962) studied the selection responses involving many environments and discussed the pooled genetic correlation among many environments. Yamada (1962) presented various formulae on the relationships between genotype × environment interaction and genetic correlation in the cases of random and mixed linear models. He derived the relationships by comparing the expected mean squares and cross-products of a one-way multiple-trait model and the expected mean squares of a two-way single-trait model assuming balanced data. Later, Fernando et al. (1984) criticized his method, stating that it gave biased estimates of genetic covariances when applied to unbalanced data and, therefore, should not be used in general situations that frequently occur in animal breeding. However, Yamada et al. (1988) claimed that the criticism was inappropriate, and used an alternative method that was also applicable to unbalanced data. Their method was to compare directly the variance-covariance structures of observations between the one-way multiple-trait model and the two-way single-trait model involving two fixed environments. Further, they made complete distinction between parameters and estimates, which was ambiguous in Yamada (1962). However, their results were inconsistent with those of Yamada (1962) [e.g., Eq. (25) in 1962 versus Eq. (7) in 1988]. Thus, it was possible that their method included some errors.

The objectives of this paper are to correct the results of Yamada et al. (1988), to generalize them for the case involving more than two environments, and to compare them with results obtained in earlier works.

---

* Present address: Faculty of Veterinary Medicine and Animal Science, Universiti Pertanian Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

Models

Suppose the observations on one trait have been obtained in k different environments. Each individual can be observed only once in one environment. The individuals are divided into families or strains chosen at random which, for generality, are called genetic groups in this paper. Let p be the number of the genetic groups and n_i the number of observations in the i^{th} environment ($i=1$ to k). Two different models are considered: Model A and Model B.

Model A

First, assume that measurements on a trait in different environments represent different traits. Then a one-way multiple-trait model is assumed:

$$y=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} \mathbf{1}_1\,\mu_1 \\ \mathbf{1}_2\,\mu_2 \\ \vdots \\ \mathbf{1}_k\,\mu_k \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z}_1 & & 0 \\ & \boldsymbol{Z}_2 & \\ 0 & & \ddots \\ & & & \boldsymbol{Z}_k \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_k \end{bmatrix} + \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_k \end{bmatrix} \tag{1}$$

where $\boldsymbol{y}_i$ ($i=1$ to k) is an $n_i \times 1$ vector of observations for 'trait' i; μ_i is the expected value of trait i; $\mathbf{1}_i$ is an $n_i \times 1$ vector of all ones; $\boldsymbol{u}_i$ is a $p \times 1$ vector of random genetic-group effects for trait i; $\boldsymbol{Z}_i$ is an $n_i \times p$ incidence matrix relating $\boldsymbol{u}_i$ to $\boldsymbol{y}_i$, $\boldsymbol{e}_i$ is an $n_i \times 1$ vector of residuals for trait i. Expectations and variance-covariance matrices for Model A are:

$E(\boldsymbol{y}_i)=\mathbf{1}_i\,\mu_i$, $E(\boldsymbol{u}_i)=\mathbf{0}$, $E(\boldsymbol{e}_i)=\mathbf{0}$,

$$\operatorname{Var}\begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}\sigma_1^2 & \boldsymbol{I}\sigma_{12} & \cdots & \boldsymbol{I}\sigma_{1k} \\ \boldsymbol{I}\sigma_{21} & \boldsymbol{I}\sigma_2^2 & \cdots & \boldsymbol{I}\sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{I}\sigma_{k1} & \boldsymbol{I}\sigma_{k2} & \cdots & \boldsymbol{I}\sigma_k^2 \end{bmatrix},$$

$$\operatorname{Var}\begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_1\,\sigma_{e_1}^2 & & 0 \\ & \boldsymbol{I}_2\,\sigma_{e_2}^2 & \\ & & \ddots \\ 0 & & & \boldsymbol{I}_k\,\sigma_{e_k}^2 \end{bmatrix},$$

$\operatorname{Cov}(\boldsymbol{u}_i, \boldsymbol{e}_j)=\mathbf{0}\,(i, j=1 \text{ to } k,\ i\neq j)$

where E represents expectation, Var represents variance-covariance matrix, and Cov represent covariance matrix; σ_i^2 ($i=1$ to k) is the genetic-group variance of trait i; σ_{ij} ($i, j=1$ to k, $i\neq j$) is the genetic-group covariance between trait i and j; $\sigma_{e_i}^2$ is the error variance of trait i; $\boldsymbol{I}$ and $\boldsymbol{I}_i$ are identity matrices of appropriate order. Note that the elements of $\boldsymbol{u}_i$ are assumed to be uncorrelated. In animal breeding terms, this means that the genetic groups are not related. Consequently, we obtain:

$$\operatorname{Var}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z}_1\boldsymbol{Z}_1'\sigma_1^2 & \boldsymbol{Z}_1\boldsymbol{Z}_2'\sigma_{12} & \cdots & \boldsymbol{Z}_1\boldsymbol{Z}_k'\sigma_{1k} \\ \boldsymbol{Z}_2\boldsymbol{Z}_1'\sigma_{21} & \boldsymbol{Z}_2\boldsymbol{Z}_2'\sigma_2^2 & \cdots & \boldsymbol{Z}_2\boldsymbol{Z}_k'\sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{Z}_k\boldsymbol{Z}_1'\sigma_{k1} & \boldsymbol{Z}_k\boldsymbol{Z}_2'\sigma_{2k} & \cdots & \boldsymbol{Z}_k\boldsymbol{Z}_k'\sigma_k^2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{I}_1\,\sigma_{e_1}^2 & & 0 \\ & \boldsymbol{I}_2\,\sigma_{e_2}^2 & \\ & & \ddots \\ 0 & & & \boldsymbol{I}_k\,\sigma_{e_k}^2 \end{bmatrix}. \tag{2}$$

Thus, the variances and covariances of $\boldsymbol{y}_i$'s are:

$$\operatorname{Var}(\boldsymbol{y}_i)=\boldsymbol{Z}_i\,\boldsymbol{Z}_i'\,\sigma_i^2+\boldsymbol{I}_i\,\sigma_{e_i}^2, \tag{3}$$

$$\operatorname{Cov}(\boldsymbol{y}_i, \boldsymbol{y}_j)=\boldsymbol{Z}_i\,\boldsymbol{Z}_j'\,\sigma_{ij}. \tag{4}$$

Model B

Next, a two-way single-trait mixed model is assumed for the same observations as were used in Eq. (1). This model shall be called Model B:

$$y=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} \mathbf{1}_1\,\mu_1 \\ \mathbf{1}_2\,\mu_2 \\ \vdots \\ \mathbf{1}_k\,\mu_k \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z}_1 \\ \boldsymbol{Z}_2 \\ \vdots \\ \boldsymbol{Z}_k \end{bmatrix} \boldsymbol{u}_G + \begin{bmatrix} \boldsymbol{Z}_1 & & 0 \\ & \boldsymbol{Z}_2 & \\ 0 & & \ddots \\ & & & \boldsymbol{Z}_k \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_{I_1} \\ \boldsymbol{u}_{I_2} \\ \vdots \\ \boldsymbol{u}_{I_k} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_k \end{bmatrix} \tag{5}$$

where $\boldsymbol{y}_i$, μ_i, $\mathbf{1}_i$, and $\boldsymbol{Z}_i$ are as defined earlier; $\boldsymbol{u}_G$ is a vector of genetic-group effects averaged over environments; $\boldsymbol{u}_{I_i}$ is an $n_i \times 1$ vector of genetic-group × environment interaction in the i^{th} environment, and $\boldsymbol{\varepsilon}_i$ is an $n_i \times 1$ vector of residuals in the i^{th} environment.

Expectations and variance-covariance matrices of random variables in Model B are:

$E(\boldsymbol{y}_i)=\mathbf{1}_i\,\mu_i$, $E(\boldsymbol{u}_G)=\mathbf{0}$, $E(\boldsymbol{u}_{I_i})=\mathbf{0}$, $E(\boldsymbol{\varepsilon}_i)=\mathbf{0}$,

$\operatorname{Var}(\boldsymbol{u}_G)=\boldsymbol{I}\,\sigma_G^2$,

$$\mathrm{Var}\begin{bmatrix} \boldsymbol{u}_{I_1} \\ \boldsymbol{u}_{I_2} \\ \vdots \\ \boldsymbol{u}_{I_k} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}\,\sigma^2_{I_2} & \boldsymbol{I}\,\sigma_{I_{12}} & \cdots & \boldsymbol{I}\,\sigma_{I_{1k}} \\ \boldsymbol{I}\,\sigma_{I_{12}} & \boldsymbol{I}\,\sigma^2_{I_2} & \cdots & \boldsymbol{I}\,\sigma_{I_{2k}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{I}\,\sigma_{I_{1k}} & \boldsymbol{I}\,\sigma_{I_{2k}} & \cdots & \boldsymbol{I}\,\sigma^2_{I_k} \end{bmatrix},$$

$$\mathrm{Var}\begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_1\,\sigma^2_{\varepsilon_1} & & 0 \\ & \boldsymbol{I}_2\,\sigma^2_{\varepsilon_2} & \\ 0 & & \boldsymbol{I}_k\,\sigma^2_{\varepsilon_k} \end{bmatrix},$$

$\mathrm{Cov}(\boldsymbol{u}_G, \boldsymbol{\varepsilon}_i) = \boldsymbol{0}$, $\mathrm{Cov}(\boldsymbol{u}_{I_i}, \boldsymbol{\varepsilon}_j) = \boldsymbol{0}$,
$\mathrm{Cov}(\boldsymbol{u}_G, \boldsymbol{u}_{I_i}) = \boldsymbol{I}\,\sigma_{GI_i}$

where σ^2_G is the genetic-group variance component; $\sigma^2_{I_i}$ ($i=1$ to k) is the interaction variance component in the i^{th} environment; $\sigma_{I_{ij}}$ ($i, j=1$ to k, $i \neq j$) is the covariance component between the interactions in the i^{th} and j^{th} environments; $\sigma^2_{\varepsilon_i}$ is the error variance component in the i^{th} environment; σ_{GI_i} is the covariance component between genetic-group effects and interactions in the i^{th} environment. The characteristics of these assumptions are summarized as follows: (1) the error variances may be heterogeneous; (2) the covariances between genetic group effects and interactions exist; (3) the interaction variances may be heterogeneous; (4) the covariances among interactions exist. Assumptions (1) and (2) include a special case of homogeneous variances, and assumptions (2) and (4) include cases where some covariances are zero. Assumptions (1) and (2) were used by Yamada et al. (1988), but (4) was not. Assumption (3) is irrelevant in Yamada et al. (1988), because they assumed only two environments and thus the two interaction variance components should become identical. All four assumptions are necessary to derive the correct relationships between the parameters of the two models. In addition to these assumptions, within any environment homogeneity of error variances among genetic groups is being assumed as usual.

Using the assumptions in Model B stated above, we obtain

$$\mathrm{Var}\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z}_1\boldsymbol{Z}_1' & \boldsymbol{Z}_1\boldsymbol{Z}_2' & \cdots & \boldsymbol{Z}_1\boldsymbol{Z}_k' \\ \boldsymbol{Z}_2\boldsymbol{Z}_1' & \boldsymbol{Z}_2\boldsymbol{Z}_2' & \cdots & \boldsymbol{Z}_2\boldsymbol{Z}_k' \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{Z}_k\boldsymbol{Z}_1' & \boldsymbol{Z}_k\boldsymbol{Z}_2' & \cdots & \boldsymbol{Z}_k\boldsymbol{Z}_k' \end{bmatrix} \sigma^2_G + \begin{bmatrix} \boldsymbol{Z}_1\boldsymbol{Z}_1'\sigma^2_{I_1} & \boldsymbol{Z}_1\boldsymbol{Z}_2'\sigma_{I_{12}} \cdots & \boldsymbol{Z}_1\boldsymbol{Z}_k'\sigma_{I_{1k}} \\ \boldsymbol{Z}_2\boldsymbol{Z}_1'\sigma_{I_{12}} & \boldsymbol{Z}_2\boldsymbol{Z}_2'\sigma^2_{I_2} \cdots & \boldsymbol{Z}_2\boldsymbol{Z}_k'\sigma_{I_{2k}} \\ \vdots & \vdots & \vdots \\ \boldsymbol{Z}_k\boldsymbol{Z}_1'\sigma_{I_{1k}} & \boldsymbol{Z}_k\boldsymbol{Z}_2'\sigma_{I_{2k}} \cdots & \boldsymbol{Z}_k\boldsymbol{Z}_k'\sigma^2_{I_k} \end{bmatrix} \tag{6}$$

$$+ \begin{bmatrix} 2\boldsymbol{Z}_1\boldsymbol{Z}_1'\sigma_{GI_1} & \boldsymbol{Z}_1\boldsymbol{Z}_2'(\sigma_{GI_1}+\sigma_{GI_2}) & \cdots & \boldsymbol{Z}_1\boldsymbol{Z}_k'(\sigma_{GI_1}+\sigma_{GI_k}) \\ \boldsymbol{Z}_2\boldsymbol{Z}_1'(\sigma_{GI_2}+\sigma_{GI_1}) & 2\boldsymbol{Z}_2\boldsymbol{Z}_2'\sigma_{GI_2} & \cdots & \boldsymbol{Z}_2\boldsymbol{Z}_k'(\sigma_{GI_2}+\sigma_{GI_k}) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{Z}_k\boldsymbol{Z}_1'(\sigma_{GI_k}+\sigma_{GI_1}) & \boldsymbol{Z}_k\boldsymbol{Z}_2'(\sigma_{GI_k}+\sigma_{GI_2}) & \cdots & 2\boldsymbol{Z}_k\boldsymbol{Z}_k'\sigma_{GI_k} \end{bmatrix} + \begin{bmatrix} \boldsymbol{I}_1\,\sigma^2_{\varepsilon_1} & & 0 \\ & \boldsymbol{I}_2\,\sigma^2_{\varepsilon_2} & \\ & & \ddots & \\ 0 & & & \boldsymbol{I}_k\,\sigma^2_{\varepsilon_k} \end{bmatrix}.$$

Thus the variances and covariances of the $\boldsymbol{y}_i$'s are:

$$\mathrm{Var}(\boldsymbol{y}_i) = \boldsymbol{Z}_i\boldsymbol{Z}_i'(\sigma^2_G + \sigma^2_{I_2} + 2\sigma_{GI_i}) + \boldsymbol{I}_i\,\sigma_{\varepsilon^i}{}^2, \tag{7}$$

$$\mathrm{Cov}(\boldsymbol{y}_i, \boldsymbol{y}_j) = \boldsymbol{Z}_i\boldsymbol{Z}_j'(\sigma^2_G + \sigma_{I_{ij}} + \sigma_{GI_i} + \sigma_{GI_j}). \tag{8}$$

Relationships between parameters

Comparison of the two models

Both Models A and B are for the same observations, so that Eq. (2) and Eq. (6) [or Eqs. (3) and (7) and Eqs. (4) and (8)] should always be identical, irrespective of the choice of $\boldsymbol{Z}_i$. Thus, we obtain

$$\sigma^2_i = \sigma^2_G + \sigma^2_{I_i} + 2\sigma_{GI_i}, \tag{9}$$

$$\sigma_{ij} = \sigma^2_G + \sigma_{I_{ij}} + \sigma_{GI_i} + \sigma_{GI_j}, \tag{10}$$

$$\sigma^2_{e_i} = \sigma^2_{\varepsilon_i}. \tag{11}$$

These equations represent the basic relationships of the parameters between the two models. From these equations, useful relationships can be derived as shown in the following.

Restriction on Model B

Because $\boldsymbol{u}_{I_i}$ consists of the deviations of subclass means from the combinations of μ_i and $\boldsymbol{u}_G$, it seems reasonable to assume that the sum of $\boldsymbol{u}_{I_i}$ over k fixed environments should become zero, i.e.,

$$\sum_{i=1}^{k} \boldsymbol{u}_{I_i} = \boldsymbol{0}. \tag{12}$$

This was explicitly used also by Anderson and Bancroft (1952), p. 340), Scheffé (1959, p. 263) and Graybill (1961, p. 397). In fact, this restriction leads to the existence of covariances σ_{GI_i} and $\sigma_{I_{ij}}$ in Model B.

Now let us derive the relationship among parameters in Model B when restriction (12) is applied. Using Eq. (12),

$$\mathrm{Cov}(\boldsymbol{u}_G, \sum_{i=1}^{k} \boldsymbol{u}_{I_i}) = \boldsymbol{I} \sum_{i=1}^{k} \sigma_{GI_i} = \boldsymbol{0},$$

we thus obtain

$$\sum_{i=1}^{k} \sigma_{GI_i} = 0. \tag{13}$$

Further, from Eq. (12),

$$\mathrm{Cov}(\boldsymbol{u}_{I_i}, \sum_{j=1}^{k} \boldsymbol{u}_{I_j}) = \boldsymbol{I} \sum_{j=1}^{k} \sigma_{I_{ij}} = \boldsymbol{0},$$

thus, using $\sigma_{I_{ii}} = \sigma^2_{I_i}$,

$$\sum_{j=1}^{k} \sigma_{I_{ij}} = \sigma^2_{I_i} + \sum_{\substack{j \neq i \\ j=1}}^{k} \sigma_{I_{ij}} = 0.$$

Summing this for i gives

$$\sum_{i=1}^{k} \sigma^2_{I_i} + \sum_{i \neq j}\sum \sigma_{I_{ij}} = 0,$$

or

$$\sum_{i<j}\sum \sigma_{I_{ij}} = -\tfrac{1}{2} \sum_{i=1}^{k} \sigma^2_{I_i}. \tag{14}$$

Using Eqs. (13) and (14), Eqs. (9) and (10) can be rewritten as follows.

Summing Eq. (9) for i ($i=1$ to k) gives

$$\sum_{i=1}^{k} \sigma^2_i = k\,\sigma^2_G + \sum_{i=1}^{k} \sigma^2_{I_i} + 2 \sum_{i=1}^{k} \sigma_{GI_i} = k\,\sigma^2_G + \sum_{i=1}^{k} \sigma^2_{I_i}. \tag{15}$$

Summing Eq. (10) for all combinations of i and j gives

$$\begin{aligned} \sum_{i<j}\sum \sigma_{ij} &= \frac{k(k-1)}{2}\sigma^2_G + \sum_{i<j}\sum \sigma_{I_{ij}} + \sum_{i<j}\sum (\sigma_{GI_i} + \sigma_{GI_j}) \\ &= \frac{k(k-1)}{2}\sigma^2_G - \frac{1}{2}\sum_{i=1}^{k} \sigma^2_{I_i} + (k-1)\sum_{i=1}^{k} \sigma_{GI_i} \\ &= \frac{k(k-1)}{2}\sigma^2_G - \frac{1}{2}\sum_{i=1}^{k} \sigma^2_{I_i}. \end{aligned} \tag{16}$$

Deleting σ^2_G from Eqs. (15) and (16), we get

$$\begin{aligned} \sum_{i=1}^{k} \sigma^2_{I_i} &= \frac{k-1}{k}\sum_{i=1}^{k} \sigma^2_i - \frac{2}{k}\sum_{i<j}\sum \sigma_{ij} \\ &= \frac{1}{k}\sum_{i<j}\sum \{(\sigma_i - \sigma_j)^2 + 2\sigma_i\sigma_j\} - \frac{2}{k}\sum_{i<j}\sum \sigma_{ij} \\ &= \frac{1}{k}\sum_{i<j}\sum (\sigma_i - \sigma_j)^2 + \frac{2}{k}\sum_{i<j}\sum (\sigma_i\sigma_j - \sigma_{ij}). \end{aligned} \tag{17}$$

Let $r_{g_{ij}}$ be the genetic correlation between environments i and j defined as $r_{g_{ij}} = \sigma_{ij}/\sigma_i\sigma_j$. Then Eq. (17) is rewritten as

$$\sum_{i=1}^{k} \sigma^2_{I_i} = \frac{1}{k}\sum_{i<j}\sum (\sigma_i - \sigma_j)^2 + \frac{2}{k}\left\{\sum_{i<j}\sum (1 - r_{g_{ij}})\,\sigma_i\sigma_j\right\}. \tag{18}$$

Reparameterization

In Model A, let us define the variance components averaged over k environments and the covariance component averaged over all $k(k-1)/2$ pairs of environments as follows:

$$\overline{\sigma^2_i} = \frac{1}{k}\sum_{i=1}^{k} \sigma^2_i, \quad \overline{\sigma_{ij}} = \frac{2}{k(k-1)}\sum_{i<j}\sum \sigma_{ij}, \quad \overline{\sigma^2_e} = \frac{1}{k}\sum_{i=1}^{k} \sigma^2_{e_i}. \tag{19}$$

Similarly, let us define the average error variance component in Model B as

$$\overline{\sigma^2_\varepsilon} = \frac{1}{k}\sum_{i=1}^{k} \sigma^2_{\varepsilon_i}. \tag{20}$$

Following Scheffé (1959), p. 264, 8.1.9), let us define the average interaction variance component as

$$\overline{\sigma^2_I} = \frac{1}{k-1}\sum_{i=1}^{k} \sigma^2_{I_i}. \tag{21}$$

Note that the denominator of Eq. (21) is $k-1$ instead of k. This is because restriction (12) is imposed on the interaction. From Eqs. (14) and (21), we obtain

$$\sum_{i<j}\sum \sigma_{I_{ij}} = -\frac{k-1}{2}\overline{\sigma^2_I}. \tag{22}$$

This indicates that non-zero covariances among interactions exist whenever a non-zero interaction variance exists. Using Eqs. (19)–(22), Eqs. (15), (16), and (11) are rewritten as:

$$\overline{\sigma^2_i} = \sigma^2_G + \frac{k-1}{k}\overline{\sigma^2_I}, \tag{23}$$

$$\overline{\sigma_{ij}} = \sigma^2_G - \frac{1}{k}\overline{\sigma^2_I}, \tag{24}$$

$$\overline{\sigma^2_e} = \overline{\sigma^2_\varepsilon}. \tag{25}$$

From Eqs. (23) and (24), we obtain

$$\sigma^2_G = \frac{1}{k}\overline{\sigma^2_i} + \frac{k-1}{k}\overline{\sigma_{ij}}, \tag{26}$$

$$\overline{\sigma^2_I} = \overline{\sigma^2_i} - \overline{\sigma_{ij}}. \tag{27}$$

Formulae (25)–(27) agree completely with the results of Yamada (1962, Eq. 24).

Dividing Eq. (18) by $k-1$ and substituting Eq. (21), we get

$$\begin{aligned} \overline{\sigma^2_I} &= \frac{1}{k(k-1)}\sum_{i<j}\sum (\sigma_i - \sigma_j)^2 \\ &\quad + \frac{2}{k(k-1)}\left\{\sum_{i<j}\sum (1 - r_{g_{ij}})\,\sigma_i\sigma_j\right\}. \end{aligned} \tag{28}$$

Formula (28) expresses the relationship between the interaction variance component and the genetic correlation. When $k=2$, this agrees with the result of Robertson (1959, p. 478).

Following the notations used by Yamada (1962), let us denote the variance of the standard deviations of genetic group effects $\theta(\sigma_i)$ as:

$$\begin{aligned}\theta(\sigma_i) &= \frac{1}{k-1}\left\{\sum_{i=1}^{k}\sigma_i^2-\frac{1}{k}\left(\sum_{i=1}^{k}\sigma_i\right)^2\right\}\\ &= \frac{1}{k-1}\left\{\sum_{i=1}^{k}\sigma_i^2-\frac{1}{k}\left(\sum_{i=1}^{k}\sigma_i^2+2\sum_{i<j}\sum\sigma_i\sigma_j\right)\right\}\\ &= \frac{1}{k(k-1)}\left\{(k-1)\left(\sum_{i=1}^{k}\sigma_i^2-2\sum_{i<j}\sum\sigma_i\sigma_j\right)\right\}\\ &= \frac{1}{k(1-k)}\sum_{i<j}\sum(\sigma_i^2+\sigma_j^2-2\sigma_i\sigma_j)\\ &= \frac{1}{k(k-1)}\sum_{i<j}\sum(\sigma_i-\sigma_j)^2. \qquad (29)\end{aligned}$$

Using this, Eq. (28) can be expressed in another way as

$$\overline{\sigma_I^2}=\theta(\sigma_i)+\frac{2}{k(k-1)}\left\{\sum_{i<j}\sum(1-r_{g_{ij}})\,\sigma_i\sigma_j\right\}. \qquad (30)$$

Average genetic correlation

Now consider the average genetic correlation defined as

$$\bar{r}_g=\frac{\overline{\sigma_{ij}}}{\overline{\sigma_i\sigma_j}}=\frac{\sum_{i<j}\sum\sigma_{ij}}{\sum_{i<j}\sum\sigma_i\sigma_j}, \qquad (31)$$

where $\overline{\sigma_i\sigma_j}$ is the average of $\sigma_i\sigma_j$ for all pairs of i and j, i.e.,

$$\overline{\sigma_i\sigma_j}=\frac{2}{k(k-1)}\sum_{i<j}\sum\sigma_i\sigma_j. \qquad (32)$$

This definition of the average genetic correlation is the same as that of Yamada (1962, Eq. 35). However, it is not clear whether or not it is the same as that of Dickerson (1962), because he did not define the average genetic correlation explicitly. In practical applications, $\bar{r}_g$ may become of limited interest relative to the individual $r_{g_{ij}}$ values, but the relationship between $\bar{r}_g$ and other parameters is presented here for comparison with the results obtained in earlier work.

Using Eq. (29), it is shown that

$$\begin{aligned}\sum_{i=1}^{k}\sigma_i^2 &= \frac{1}{k-1}\left\{\sum_{i<j}\sum(\sigma_i-\sigma_j)^2+2\sum_{i<j}\sum\sigma_i\sigma_j\right\}\\ &= k\,\theta(\sigma_i)+\frac{2}{k-1}\sum_{i<j}\sum\sigma_i\sigma_j. \qquad (33)\end{aligned}$$

Equating Eqs. (15) and (33), and using Eq. (32), it can be shown that

$$\overline{\sigma_i\sigma_j}=\sigma_G^2+\frac{k-1}{k}\overline{\sigma_I^2}-\theta(\sigma_i). \qquad (34)$$

Substituting Eqs. (24) and (34) into Eq. (31), we get

$$\bar{r}_g=\frac{\sigma_G^2-\frac{1}{k}\overline{\sigma_I^2}}{\sigma_G^2+\frac{k-1}{k}\overline{\sigma_I^2}-\theta(\sigma_i)}. \qquad (35)$$

This formula agrees with the result of Yamada (1962, Eq. 41).

Discussion

Fernando et al. (1984) compared two models (Model A and Model B) and stated that when genetic and error variances are heterogeneous among environments, no meaningful relationships exist between parameters of the two models (p. 176). Their Model A is the same as that used in this paper, but their Model B is one that is customarily used for unbalanced data. Insofar as such a model is assumed, their result is appropriate. On the other hand, in this paper, a different model (Model B) was assumed and, consequently, the relationships between parameters could be derived [Eqs. (9)–(11)]. The difference between the results of Fernando et al. (1984) and those of this paper derives from the difference between the models assumed. Except for the problem of the heterogeneous variances, a major difference between these models comes from whether covariances σ_{GI_i} and $\sigma_{I_{ij}}$ are assumed or not. Their existence should be considered when restriction (12) is put on the model. Thus, we can conclude that the primary difference between Fernando et al. (1984) and this paper comes from whether restriction (12) is imposed or not. This problem pertaining to the mixed model is discussed by Hartley and Searle (1969) and Searle (1971, p. 16). The question as to whether this restriction should be used or not in the mixed model analysis generally is not yet answered. However, for the purpose of deriving the relationships of parameters between the models, this restriction is absolutely required.

Model B′ used by Yamada et al. (1988) is different from Model B in this paper, and thus their result is also different from ours. Their model incorporates σ_{GI_i} but not $\sigma_{I_{ij}}$. Both of those covariances should be incorporated, which gives rise to some changes in the coefficient matrix of σ_I in Eq. (5) of Yamada et al. (1988). Thus, their model and result seem to be inappropriate.

The basic formulae representing the relationships between parameters, Eqs. (9)–(11), were then transformed by reparameterization using average variance and covariance components. Then the result, Eqs. (25)–(27), which agrees with that of Yamada (1962) was obtained. Thus, the result of Yamada is also applicable for unbalanced data. However, it should be noted that the average variance components given in this paper, as well as those given by Yamada (1962), are different from the variance components that are given in the two-way mixed linear model used customarily for unbalanced data. Furthermore, in reparameterization (19)–(21), equal weights were given for each environment. It is the same with restriction (12) imposed on Model B. This assumes that the population size in each environment is infinite, or finite and equal. When the populations have finite and unequal sizes, it is difficult to rationalize those equal weights. In such cases, weights that are proportional to the population sizes should be used, and then our results (12)–(35) will become inappropriate.

The results derived in this study are the relationships between parameters, so that they should not be influenced by sampling or by the methods used for estimating variance components. The models in this study include only one environmental factor. However, even when multiple environmental factors are involved, the same results will be obtained.

References

Anderson RL, Bancroft TA (1952) Statistical theory in research. McGraw-Hill, New York

Dickerson GE (1962) Implications of genetic-environmental interaction in animal breeding. Anim Prod 4:47–64

Falconer DS (1952) The problem of environment and selection. Am Nat 86:293–298

Fernando RL, Knight SA, Gianola D (1984) On the method of estimating the genetic correlation between characters measured in different experimental units. Theor Appl Genet 67:175–178

Graybill FA (1961) An introduction to linear statistical models. McGraw-Hill, New York

Hartley HO, Searle SR (1969) A discontinuity in mixed model analysis. Biometrics 25:573–576

Robertson A (1959) The sampling variance of the genetic correlation coefficient. Biometrics 15:469–485

Scheffé H (1959) The analysis of variance. John Wiley, New York

Searle SR (1971) Topics in variance component estimation. Biometrics 278:1–76

Yamada Y (1962) Genotype by environment interaction and genetic correlation of the same trait under different environments. Jpn J Genet 37:498–509

Yamada Y, Itoh Y, Sugimoto I (1988) Parametric relationships between genotype × environment interaction and genetic correlation when two environments are involved. Theor Appl Genet 76:850–854