

M. J. Dieters · T. L. White · R. C. Littell
G. R. Hodge

Application of approximate variances of variance components and their ratios in genetic tests

Received: 10 August 1994 / Accepted: 30 September 1994

Abstract The application and underlying assumptions of formulae used to estimate the variance of variance components and ratios of variance components are fully described for (1) variance components estimated using Henderson's Method 3 (HM3) and Restricted Maximum Likelihood (REML) and (2) ratios of variance components commonly used in genetic tests – biased and unbiased heritabilities. A first-order Taylor series approximation is often used to estimate the variance of a ratio of two random variables (e.g., heritability), however the formula is complicated, thus making calculations prone to errors. Dickerson's approximation is considerably simpler, though relatively rarely used. In case studies using data from 148 slash pine full-sib progeny tests, Dickerson's method was found to be slightly more conservative than the Taylor series approximation when estimating the variance of heritability estimates, regardless of test size, age, or the trait (volume, which is a continuous trait, and rust resistance, which is a bernoulli trait). Both the Taylor series and Dickerson approximations compared favorably with an empirical estimate of the variance of heritability estimates, however there is some evidence of small-sample bias associated with the use of the asymptotic variance-covariances from REML variance component estimation.

Key words Variance of ratios · Heritability · Taylor series · Variance components · REML

Introduction

The results of quantitative genetics tests are commonly summarized in terms of estimated variance components and ratios of these variance components, such as heritability and genetic correlation. When such statistics are examined, it is natural to inquire about their precision; hence, standard errors are customarily reported for the variance component estimates and ratios of interest. The estimation of such standard errors is not simple, and in the case of the variance of a ratio of variance components, no closed-form expression exists.

Knapp et al. (1985) present exact confidence intervals for family mean heritability from balanced data, and nonparametric confidence intervals can also be estimated (Knapp et al. 1989). However, the formulae most commonly used to estimate standard errors of ratios of variance components are based on a Taylor series approximation (Tallis 1959; Swiger et al. 1964; Kempthorne 1957; Becker 1975; Namkoong 1979; Bulmer 1980; Baker 1986; Hallauer and Miranda 1988; Falconer 1989). Although the theory of this Taylor series approximation is accessible in advanced calculus and statistics texts (Kaplan 1952; Kendall et al. 1987; Searle et al. 1992), neither the derivation nor the underlying assumptions required in the derivation and application of these formulae have been presented in either the forestry or the genetics literature. Failure to clearly understand the parameters and the assumptions implicit in these formulae can lead to erroneous and inappropriate application.

The objectives of this paper are twofold. First, to clearly lay out the theory and assumptions needed to derive approximate standard errors of ANOVA- (analysis of variance) and REML- (restricted maximum likelihood) based estimators of variance components and their ratios for both balanced and unbalanced data. The

This is Florida Agricultural Experiment Station Journal Series No. R-03964 of the Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL 32611, USA

Communicated by P. M. A. Tigerstedt

M. J. Dieters¹ (✉) · T. L. White · G. R. Hodge
Department of Forestry, University of Florida, Gainesville,
FL 32611, USA

R. C. Littell
Department of Statistics, University of Florida, Gainesville,
FL 32611, USA

Present address:

¹ Queensland Forest Research Institute, MS 483 Fraser Road, Gympie, QLD 4570, Australia

second objective is to compare the approximate variance of ratios of variance components to the empirical variance among many separate estimates of these ratios, which were obtained from real data. The ratios of variance components which will be used in this study are (1) single-site biased heritability (h_b^2) estimated using ANOVA and REML and (2) unbiased heritability (h^2) from paired sites estimated using REML. The case studies will examine two traits of slash pine (*Pinus elliottii* Engelm. var. *elliottii*) in the Cooperative Forest Genetic Research Program (CFGRP) at the University of Florida: resistance to fusiform rust (caused by the fungus *Cronartium quercuum* f. sp. *fusiforme*), which is a bernoulli trait, assessed as infected or rust-free, with moderate heritability; and tree volume, which is a continuous trait with a lower heritability.

Theory

Variance of variance components – ANOVA estimates

All ANOVA-based estimators of variance components rely on equating observed mean squares to their expectations and solving for the variance components. When data are balanced, ANOVA-based estimates of variance components are unique and unbiased, and have the minimum variance among all unbiased quadratic estimators (Swallow and Monahan 1984; Khuri and Sahai 1985; Searle et al. 1992). These optimal statistical properties usually make ANOVA-based estimators the best choice when data are balanced. Although ANOVA-based variance component estimates are derived from sums of squares in the ANOVA table, no assumption of normality is required to derive ANOVA estimators (Searle et al. 1992). However, in order to derive the usual estimator for the variance of variance component estimates, the traditional normality assumptions for analysis of variance are required (i.e., that all random effects in the model are normally distributed with a mean of zero, some finite variance, and zero covariances between effects (Searle et al. 1992; Winer et al. 1991)).

In order to apply formulae for the variance of ratios of variance components, it is necessary to know the variance and covariance among the variance component estimates. To obtain this information we must first estimate the variance of the mean squares in the ANOVA. Once the variance of the mean squares is known, it is then relatively simple to estimate the covariances of the variance component estimates.

If SS is a sum of squares with degrees of freedom f in an ANOVA table, then the variance of the corresponding mean square ($MS = SS/f$) is

$$\text{Var}(MS) = \frac{2(MS^2)}{(f+2)} \quad (1)$$

For a complete derivation of Eq. 1 refer to chapter 2 of Dieters (1994) or Appendix S.3 of Searle et al. (1992).

The variance of an ANOVA-based variance component estimate is derived by using the fact that all such variance component estimates are linear functions of the observed mean squares. Let X_1, X_2, \dots, X_n be random variables, and a_1, a_2, \dots, a_n be constants. Then for $U = \sum_i a_i X_i$,

$$\text{Var}(U) = \sum_i a_i^2 \text{Var}(X_i) + 2 \sum_{i < j} a_i a_j \text{Cov}(X_i, X_j), \quad (2)$$

where the double summation is over all pairs (i, j) with $i < j$ (Mendenhall et al. 1990).

Equations 1 and 2 can be used to estimate the variance of any ANOVA-based variance component estimate. For example, let $\hat{\sigma}_i^2$ be estimated by

$$\hat{\sigma}_i^2 = (MS_1 - MS_2)/c,$$

where MS_1 and MS_2 are mean squares from the analysis of variance with degrees of freedom f_1 and f_2 , respectively, and c is a constant. Then,

$$\text{Var}(\hat{\sigma}_i^2) = \text{Var}((MS_1 - MS_2)/c) = 1/c^2 \text{Var}(MS_1 - MS_2)$$

using Eq. 2,

$$\text{Var}(\hat{\sigma}_i^2) = 1/c^2 [\text{Var}(MS_1) + \text{Var}(MS_2) - 2\text{Cov}(MS_1, MS_2)]$$

applying Eq. 1,

$$\text{Var}(\hat{\sigma}_i^2) = 2/c^2 [(MS_1^2/(f_1 + 2)) + (MS_2^2/(f_2 + 2))],$$

if we assume that the covariance between the mean squares is zero. With balanced data the mean squares are orthogonal (Steel and Torrie 1981), which in turn means that there is no covariance between the mean squares (Namkoong 1979). Nevertheless, with the assumed variance covariance structure and the assumption of normality all the mean squares are independent (Searle et al. 1992). These assumptions are usually reasonable in most designed experiments where random samples are drawn from normal populations and treatments are assigned randomly to experimental units (Winer et al. 1991).

This procedure can be generalized. The i^{th} variance component can be estimated as a linear combination of the l mean squares in the ANOVA table, each with f_i degrees of freedom,

$$\hat{\sigma}_i^2 = (a_1 MS_1 + a_2 MS_2 + \dots + a_l MS_l),$$

or equivalently,

$$\hat{\sigma}_i^2 = \sum_i a_i MS_i,$$

where the a 's are constants, equivalent to $1/c$ and $-1/c$ in the previous example. And the variance of this linear function of the mean squares is estimated using

$$\begin{aligned}\text{Var}(\hat{\sigma}_i^2) &= \text{Var}\left(\sum_i^l a_i \text{MS}_i\right) = \sum_i^l \text{Var}(a_i \text{MS}_i) \\ &= \sum_i^l a_i^2 \text{Var}(\text{MS}_i).\end{aligned}$$

By using Eq. 1 to estimate the variance of MS_i , we obtain

$$\hat{\text{Var}}(\hat{\sigma}_i^2) = \sum_{i=1}^l \left[\frac{2a_i^2 (\text{MS}_i)^2}{f_i + 2} \right] \quad (3)$$

A matrix formulation of Eq. 3 is given by Searle et al. (1992, pp 128–138).

In summary, for balanced data it is possible to derive the variance of ANOVA-based variance component estimates, requiring only the traditional normality assumptions, plus the assumption of zero covariance between the mean squares. However, in the case of unbalanced data using a random or mixed model, in general the mean squares do not have chi-squared distributions, nor any multiple of a chi-squared distribution (Searle et al. 1992), which is an assumption required to derive Eq. 1. The validity of applying this methodology to variance components estimated from unbalanced data cannot be readily determined theoretically, but at best can be expected to provide approximate variances when the data are nearly balanced.

Variance of variance components – REML estimates

For unbalanced data, REML can offer significant advantages over ANOVA-based estimators: REML estimates are unique and nonnegative and have some useful large-sample statistical properties, while ANOVA-based estimates only retain the property of being unbiased (Swallow and Monahan 1984; Khuri and Sahai 1985; Searle et al. 1992). Hartley and Rao (1967) showed that with certain restrictions REML-based estimates are asymptotically normal, consistent, and asymptotically efficient as the design size increases. However, these properties have been verified under less restrictive assumptions (Harville 1977).

The derivation of REML estimators requires that an underlying probability distribution for all random effects in the model be assumed. Commonly, a multivariate normal distribution is assumed, because this assumption leads to solutions that can be readily handled mathematically, even when the data are unbalanced (Searle et al. 1992). Nevertheless, Banks et al. (1985) and Westfall (1987) have demonstrated that REML-based estimators are robust to violations of this assumption.

When using any maximum likelihood procedure, the asymptotic variance and covariances of the variance component estimates can be derived readily as part of the estimation process (Searle et al. 1992). However, these are asymptotic or large-sample properties, and so are only approximate when applied to variance component estimates derived from finite (small) samples. Many computer programs, such as PROC VARCOMP of the SAS system (SAS Institute 1988) and GAREML (Huber 1993) which perform REML estimation also provide the asymptotic variance-covariance matrix of the estimates. These variances and covariances are used to provide a measure of the precision of REML estimates even when sample size is small because no other estimates are available.

Variance of a ratio

Unlike linear functions of random variables, no closed-form expression exists for estimating the variance of a ratio of two random variables. Heritabilities and genetic correlations are ratios of estimated variance components and so are ratios of random variables. The variance of such ratios can, however, be approximated by making simplifying assumptions. Two approximate methods will be described here. The first is a simple method proposed by Dickerson (1969) that may have considerable merit but which is rarely used. The second method is based on a Taylor series approximation and is the method most commonly used in quantitative genetics.

Dickerson (1969) suggested a simple method for obtaining estimates of the variance of the ratio cx_1/x_2 (where x_1 and x_2 are random variables and c is a constant):

$$\text{Var}\left(\frac{cx_1}{x_2}\right) = \frac{c^2 \text{Var}(x_1)}{x_2^2} \quad (4)$$

Equation 4 is derived by treating the denominator, x_2 , as a constant rather than as a random variable. The assumption that the denominator is a constant appears to be justifiable in the case of heritability estimates. Narrow-sense heritability is the ratio between the additive variance and the total phenotypic variance. Estimates of the total phenotypic variance are usually much more precise than are estimates of additive variance. Thus, for most practical purposes the estimate of phenotypic variance could be considered as a known parameter rather than an estimate of that parameter. Dickerson (1969) states that although this method is conservative (i.e., slightly overestimates the variance), it gives results that agree closely with those obtained from the Taylor series approximation.

Standard quantitative genetics texts such as Kempthorne (1957, p 246), Becker (1975, p 45), and Namkoong (1979, p 232) present without proof or derivation

formulae for the variance of the ratio of two random variables based on a first-order Taylor series expansion (Kendall et al. 1987) that are all numerically equivalent to

$$\text{Var}(g) \cong \text{Var}(x_1)/\theta_2^2 + \theta_1^2 \text{Var}(x_2)/\theta_2^4 - 2\theta_1 \text{Cov}(x_1, x_2)/\theta_2^3, \quad (5)$$

where $g = x_1/x_2$, $E(x_1) = \theta_1$, $E(x_2) = \theta_2$, and x_1 and x_2 are random variables. The derivation of Eq. 5 is described in detail along with the assumptions required in the Appendix. However, it should be noted that implicit in the derivation of Eq. 5 is that $E(x_1) = \theta_1$, $E(x_2) = \theta_2$, $\text{Var}(x_1)$, $\text{Var}(x_2)$ and $\text{Cov}(x_1, x_2)$ are assumed known when in fact they must be estimated from the data. This point may seem trivial, but because of this assumption the validity of Eq. 5 is completely dependent on the accuracy and precision of the estimated parameters used, and hence the data.

Case study: heritability estimates

Two types of heritability estimates are commonly reported in the forestry literature: (1) biased heritability (Ades et al. 1992; Borralho et al. 1992; Carson 1989; Cotterill et al. 1987; Dean et al. 1986; Dieters et al. 1990; Johnson and Burdon 1990; King et al. 1988; Sluder 1993; Woolaston et al. 1990) and (2) unbiased heritability (Balocchi et al. 1993; Barnes et al. 1992; Hodge and White 1992; Kiss and Yeh 1988; Lowe and Greene 1990; White and Hodge 1992; Yeh and Heaman 1982). Biased estimates of heritability are obtained when variance components are estimated from data collected on single sites, while unbiased heritability estimates result from analyses combining data from paired or multiple test locations. In the analysis of single-site data the interaction between additive genetic effects and environmental effects is confounded with additive genetic effects. The inability to separate additive-by-environment effects from additive effects means that the estimated additive genetic variance and, consequently, the heritability, is biased upwards (Dickerson 1962; Comstock and Moll 1963).

Materials and methods

Experimental material

Variance components were estimated from single-site analyses of 148 separate full-sib progeny tests of slash pine establish between 1966 and 1988, on sites located primarily in Florida, Georgia, and Alabama, with a few tests in Mississippi and South Carolina. Each test was established in a randomized complete block design, with each full-sib family being represented by either a single row plot or non-contiguous plot composed of between 5 and 10 trees in each block. There were between 6 and 30 parents (mean of 16.4) in each test, crossed in either a diallel or factorial mating design, forming between 10 and 57 full-sib families (mean of 29.2). For the single-site analyses

(to estimate h^2), the tests were split into one of two size groups based on the number of parents in the test: 'small' (6–15 parents) and 'large' (16–30 parents).

Measurements of total height, diameter (4.5 ft or 1.37 m), and the presence/absence of fusiform rust infection were collected at between 4 and 12 years after planting. Measurements were grouped into three age classes centered at 5, 8 and 11 years. Tree volume (Goddard and Strickland 1968; Rockwood 1981) and a rust score, 0 for no rust and 100 if infected with rust, were subsequently calculated for each tree. The mean tree volume, across all tests, was 0.222, 1.104, and 2.012 cubic feet ($35.315 \text{ ft}^3 = 1 \text{ m}^3$) at 5, 8, and 11 years, respectively. Similarly, the mean rust infection levels were 28.8%, 33.4%, and 36.1% at 5, 8, and 11 years, respectively.

As a consequence of the tests being established over a 23-year period, not all tests were measured at all three ages. For this study, data were available for 121, 100, and 81 tests at 5, 8, and 11 years, respectively. The overall level of rust infection varied from 0% to near 100% infection. It is known that tests with low or high levels of rust provide little information on rust resistance (White and Hodge 1989); therefore, only tests with more than 20% and less than 85% rust were included. Hence, the study includes only 113 tests assessed for rust infection, with 78, 70, and 59 tests at 5, 8, and 11 years, respectively.

From these 148 full-sib slash pine tests, it was possible to identify 76 independent pairs of connected tests, with no tests represented in more than one test pair. However, only 39 of these test pairs had a rust infection level greater than 20% and less than 85% in both tests of the pair. These test pairs were used to obtain independent estimates of unbiased heritability. The much smaller number of test pairs than individual tests did not make it possible to divide the test pairs into groups based on test size, as was done in case of biased heritability.

Estimation of variance components

Variance components were estimated using DIALL (Schaffer and Usanis 1969) for single-site analyses, and GAREML (Huber 1993) for both single- and paired-site analyses. Prior to conducting the paired-site analyses, scale effects were removed by dividing the volume of each tree by the square root of the within-plot variance (estimated using REML on the individual site data). DIALL yields Henderson's Method III (HM3), ANOVA-based estimates, while GAREML provides REML estimates using Giesbrecht's (1983) algorithm. Negative estimates obtained using HM3 were accepted to ensure that these estimates were unbiased; however, REML estimates are by definition nonnegative.

The most complex model used was that for paired-site analyses of factorial tests, using GAREML, while all other models can be thought of as a subset of this model. This model was

$$y_{ijklmn} = \mu + t_i + b_{ij} + \text{set}_n + f_k + m_l + tf_{ik} + tm_{il} + fm_{kl} + tfm_{ikl} + p_{ijkl} + e_{ijklm}$$

where y_{ijklmn} is the m^{th} tree in the kl^{th} family in the n^{th} set, and j^{th} block of the i^{th} test, μ is the population mean, t_i is the random effect of the i^{th} test environment, $E(t_i) = 0$ and $\text{Var}(t_i) = \sigma_t^2$, b_{ij} is the random effect of the j^{th} block in the i^{th} test, $E(b_{ij}) = 0$ and $\text{Var}(b_{ij}) = \sigma_b^2$, set_n is the random effect of the n^{th} disconnected set, $E(\text{set}_n) = 0$ and $\text{Var}(\text{set}_n) = \sigma_s^2$, f_k is the random effect of the k^{th} female, $E(f_k) = 0$ and $\text{Var}(f_k) = \sigma_{fca}^2$, m_l is the random effect of the l^{th} male, $E(m_l) = 0$ and $\text{Var}(m_l) = \sigma_{mca}^2$, tf_{ik} is the random interaction between the i^{th} test and k^{th} female, $E(tf_{ik}) = 0$, and $\text{Var}(tf_{ik}) = \sigma_{tgc}^2$, tm_{il} is the random interaction between the i^{th} test and l^{th} male, $E(tm_{il}) = 0$, and $\text{Var}(tm_{il}) = \sigma_{tgc}^2$, tfm_{ikl} is the random interaction between the i^{th} test and kl^{th} family, $E(tfm_{ikl}) = 0$, and $\text{Var}(tfm_{ikl}) = \sigma_{tsc}^2$, p_{ijkl} is the random effect of the $ijkl^{\text{th}}$ plot, $E(p_{ijkl}) = 0$ and $\text{Var}(p_{ijkl}) = \sigma_p^2$, and e_{ijklm} is the random effect within the $ijkl^{\text{th}}$ plot, $E(e_{ijklm}) = 0$ and $\text{Var}(e_{ijklm}) = \sigma_w^2$.

In this model it was assumed that there was no covariance between the random effects in the model, that variances due to the female and male effects were equal, and that the female and male-by-environment interactions were the same. GAREML produced one estimate of σ_{fca}^2 and σ_{mca}^2 by pooling the estimates from the male and female parents. Also, when the set term is invoked, all subsequent terms in the model are nested within sets.

In the case of diallel tests, individual parents are used as both males and females in the crosses. By assuming the absence of any reciprocal effects, i.e., that it does not matter whether a parent is used as a male or as a female, it is possible to estimate the variance components for GCA and GCA-by-test location. Thus, in the above linear model f_k , m_i , tf_{ik} , and tm_{ij} were replaced by g_k , g_i , tg_{ik} , and tg_{ij} respectively, for the general combining ability of the k^{th} (or i^{th}) parent, and fm_{kl} was replaced by s_{kl} for the specific combining ability (Griffing 1956).

When conducting single-site analyses, the models used for factorial and diallel experiments are as given above, except that all terms involving effects due to the i^{th} test are dropped from the model.

Estimation of biased heritability and its variance

All formulae given in this section are for REML-based variance component estimates. DIALL did not allow the inclusion of a plot term in the model, thereby confounding any plot effect with the within-plot error. Hence, for HM3 estimates from DIALL, the expectation of the within-plot error is $E(\hat{\sigma}_w^2) = \sigma_p^2 + \sigma_w^2$. Thus, for HM3 estimates, all terms involving σ_p^2 are omitted, and σ_w^2 is replaced with σ_w^2 . Estimates of narrow sense, biased heritability were obtained using the formula

$$\hat{h}_b^2 = \frac{4\hat{\sigma}_{GCA}^2}{2\hat{\sigma}_{GCA}^2 + \hat{\sigma}_{SCA}^2 + \hat{\sigma}_p^2 + \hat{\sigma}_w^2}$$

where $\hat{\sigma}_{GCA}^2$ and $\hat{\sigma}_{SCA}^2$ when estimated from single sites are confounded with genotype-by-environment interaction, thus $E(\hat{\sigma}_{GCA}^2) = \sigma_{gca}^2 + \sigma_{igca}^2$ and $E(\hat{\sigma}_{SCA}^2) = \sigma_{sca}^2 + \sigma_{tsca}^2$ (Comstock and Moll 1963). These formulae for biased heritability are similar to those used in forest genetics (Cotterill 1987) and are appropriate for use with data corrected for set and block effects.

Estimates of the variance of biased heritability estimates were obtained for both the HM3 and REML variance component estimates using (1) Dickerson's method, (2) the Taylor series approximation, and (3) an empirical estimate of the variance. The empirical estimate of the variance of heritability was simply the variance among comparable heritability estimates. This empirical estimate therefore contained an estimate of the variance of heritability estimates due to the differences in site quality and uniformity, differences in management and maintenance of the tests, differences in the mating design and experimental design, as well as many other factors. Given the large number of different progeny tests involved in this empirical estimate, it is believed that the empirical estimate approached the 'true' variance of heritability estimates of slash pine in the south-eastern USA at these ages.

The formula used to calculate the Dickerson's approximate estimate of the variance of the biased heritability was

$$\text{Var}(\hat{h}_b^2) \cong \frac{16\text{Var}(\hat{\sigma}_{GCA}^2)}{(2\hat{\sigma}_{GCA}^2 + \hat{\sigma}_{SCA}^2 + \hat{\sigma}_p^2 + \hat{\sigma}_w^2)^2}$$

The formula used to calculate the variance of the biased heritability, estimated using the Taylor series approximation, was obtained by inserting into Eq. 5 the approximate values of $E(x_1) = \theta_1$, $E(x_2) = \theta_2$, $\text{Var}(x_1)$, $\text{Var}(x_2)$, and $\text{Cov}(x_1, x_2)$ given below:

$$E(x_1) \cong 4\hat{\sigma}_{GCA}^2$$

$$E(x_2) \cong 2\hat{\sigma}_{GCA}^2 + \hat{\sigma}_{SCA}^2 + \hat{\sigma}_p^2 + \hat{\sigma}_w^2$$

$$\text{Var}(x_1) \cong 16\text{Var}(\hat{\sigma}_{GCA}^2)$$

$$\begin{aligned} \text{Var}(x_2) \cong & 4\text{Var}(\hat{\sigma}_{GCA}^2) + \text{Var}(\hat{\sigma}_{SCA}^2) + \text{Var}(\hat{\sigma}_p^2) + \text{Var}(\hat{\sigma}_w^2) \\ & + 4\text{Cov}(\hat{\sigma}_{GCA}^2, \hat{\sigma}_{SCA}^2) + 4\text{Cov}(\hat{\sigma}_{GCA}^2, \hat{\sigma}_p^2) + 4\text{Cov}(\hat{\sigma}_{GCA}^2, \hat{\sigma}_w^2) \\ & + 2\text{Cov}(\hat{\sigma}_{SCA}^2, \hat{\sigma}_p^2) + 2\text{Cov}(\hat{\sigma}_{SCA}^2, \hat{\sigma}_w^2) + 2\text{Cov}(\hat{\sigma}_p^2, \hat{\sigma}_w^2) \end{aligned}$$

$$\begin{aligned} \text{Cov}(x_1, x_2) \cong & 8\text{Var}(\hat{\sigma}_{GCA}^2) + 4\text{Cov}(\hat{\sigma}_{GCA}^2, \hat{\sigma}_{SCA}^2) \\ & + 4\text{Cov}(\hat{\sigma}_{GCA}^2, \hat{\sigma}_p^2) + 4\text{Cov}(\hat{\sigma}_{GCA}^2, \hat{\sigma}_w^2) \end{aligned}$$

Estimation of unbiased heritability and its variance

Variance component estimates from the paired-site analyses were substituted into the following formula to calculate heritability estimates:

$$\hat{h}^2 = \frac{4\hat{\sigma}_{gca}^2}{2\hat{\sigma}_{gac}^2 + \hat{\sigma}_{sca}^2 + 2\hat{\sigma}_{tgca}^2 + \hat{\sigma}_{tsca}^2 + \hat{\sigma}_p^2 + \hat{\sigma}_w^2}$$

The formula used for calculating the Dickerson's approximate estimate of the variance of the biased heritability was

$$\text{Var}(\hat{h}^2) \cong \frac{16\text{Var}(\hat{\sigma}_{gca}^2)}{(2\hat{\sigma}_{gca}^2 + \hat{\sigma}_{sca}^2 + 2\hat{\sigma}_{tgca}^2 + \hat{\sigma}_{tsca}^2 + \hat{\sigma}_p^2 + \hat{\sigma}_w^2)^2}$$

The formula used for calculating the Taylor series approximate variance of the heritability was obtained by inserting into Eq. 5 the approximate values of $E(x_1) = \theta_1$, $E(x_2) = \theta_2$, $\text{Var}(x_1)$, $\text{Var}(x_2)$, and $\text{Cov}(x_1, x_2)$ given below:

$$E(x_1) \cong 4\sigma_{gca}^2$$

$$E(x_2) \cong 2\hat{\sigma}_{gca}^2 + \hat{\sigma}_{sca}^2 + 2\hat{\sigma}_{tgca}^2 + \hat{\sigma}_{tsca}^2 + \hat{\sigma}_p^2 + \hat{\sigma}_w^2$$

$$\text{Var}(x_1) \cong 16\text{Var}(\hat{\sigma}_{gca}^2)$$

$$\begin{aligned} \text{Var}(x_2) \cong & 4\text{Var}(\hat{\sigma}_{gca}^2) + \text{Var}(\hat{\sigma}_{sca}^2) + 4\text{Var}(\hat{\sigma}_{tgca}^2) + \text{Var}(\hat{\sigma}_{tsca}^2) \\ & + \text{Var}(\hat{\sigma}_p^2) + \text{Var}(\hat{\sigma}_w^2) + 4\text{Cov}(\hat{\sigma}_{gca}^2, \hat{\sigma}_{sca}^2) \\ & + 8\text{Cov}(\hat{\sigma}_{gca}^2, \hat{\sigma}_{tgca}^2) + 4\text{Cov}(\hat{\sigma}_{gca}^2, \hat{\sigma}_{tsca}^2) \\ & + 4\text{Cov}(\hat{\sigma}_{gca}^2, \hat{\sigma}_p^2) + 4\text{Cov}(\hat{\sigma}_{gca}^2, \hat{\sigma}_w^2) \\ & + 4\text{Cov}(\hat{\sigma}_{sca}^2, \hat{\sigma}_{tgca}^2) + 2\text{Cov}(\hat{\sigma}_{sca}^2, \hat{\sigma}_{tsca}^2) \\ & + 2\text{Cov}(\hat{\sigma}_{sca}^2, \hat{\sigma}_p^2) + 2\text{Cov}(\hat{\sigma}_{sca}^2, \hat{\sigma}_w^2) \\ & + 4\text{Cov}(\hat{\sigma}_{tgca}^2, \hat{\sigma}_{tsca}^2) + 4\text{Cov}(\hat{\sigma}_{tgca}^2, \hat{\sigma}_p^2) \\ & + 4\text{Cov}(\hat{\sigma}_{tgca}^2, \hat{\sigma}_w^2) + 2\text{Cov}(\hat{\sigma}_{tsca}^2, \hat{\sigma}_p^2) \\ & + 2\text{Cov}(\hat{\sigma}_{tsca}^2, \hat{\sigma}_w^2) + 2\text{Cov}(\hat{\sigma}_p^2, \hat{\sigma}_w^2) \end{aligned}$$

Results and discussion

The estimated mean biased heritability of volume ranged between 0.005 and 0.177, while the mean heritability of rust resistance ranged between 0.105 and 0.189 (Table 1). Estimates of unbiased heritability (Table 2) are generally smaller than comparable estimates obtained from single-sites (Tables 1 and 2). Heritability estimates obtained from HM3 and REML variance components are similar (except for volume at 5 years), though rarely the same, and neither method yields estimates consistently lower or higher than the other method. Given the imbalance in the data, the REML-based estimates can be expected to be more reliable, and these observed differences, though not large, do serve to highlight the need to use the most appropriate estimation method available.

Dickerson's versus Taylor series approximations

The variance of heritability estimated by the Dickerson and Taylor series approximations were compared using

Table 1 Overall mean single-site (biased) heritability of volume and rust resistance in slash pine, estimated in small (less than 16 parents), large (16 or more parents), or all full-sib tests. Heritability estimates were based upon variance components estimated by either Henderson's Method 3 (HM3) or Restricted Maximum Likelihood (REML)

Age class	Size class	Volume		Rust resistance				
		Number of test	Mean single-site heritability (h_b^2)			Number of tests	Mean single-site heritability (h_b^2)	
				HM3	REML		HM3	REML
Five years	Small	60	0.0048	0.1097		32	0.1131	0.1045
	Large	60	0.0075	0.1202		45	0.1896	0.1569
	All	120	0.0061	0.1150		77	0.1578	0.1351
Eight years	Small	48	0.1169	0.1397		33	0.1139	0.1004
	Large	52	0.1274	0.1407		37	0.1779	0.1677
	All	100	0.1224	0.1402		70	0.1477	0.1360
Eleven years	Small	43	0.1770	0.1709		27	0.1051	0.1268
	Large	39	0.1764	0.1396		33	0.1892	0.1424
	All	82	0.1767	0.1560		60	0.1514	0.1354

Table 2 Overall mean heritability of paired-site volume and rust in slash pine, estimated using Restricted Maximum Likelihood (REML). The mean heritability (h^2) and its mean variance (\pm standard error of mean) is estimated using Dickerson's and Taylor series approximations and compared to an empirical estimate

Trait	Age class (years)	Number of test pairs	Mean h^2	Estimated mean variance of heritability		
				Dickerson	Taylor	Empirical
Volume	5	62	0.0912	0.005181 \pm 0.000615	0.004617 \pm 0.000502	0.004849
	8	53	0.1181	0.007188 \pm 0.000741	0.006323 \pm 0.000633	0.006342
	11	36	0.1529	0.011918 \pm 0.000233	0.009759 \pm 0.001749	0.007657
Rust	5	28	0.1519	0.008621 \pm 0.002165	0.007243 \pm 0.001932	0.009479
	8	21	0.1471	0.007315 \pm 0.001875	0.005772 \pm 0.001332	0.015803
	11	18	0.1622	0.007438 \pm 0.001475	0.006041 \pm 0.001032	0.008535

simple linear regression, and the relationship was always strong (all R squares exceeded 0.92). The strength of this relationship was apparently unaffected by the method used to estimate the variance components, the trait examined, test size, or whether single or paired-site analyses are used.

The average variance for h_b^2 of volume and rust resistance, estimated by the Dickerson and Taylor series approximations are of similar magnitude (Figs. 1 and 2). The mean variance estimated by Dickerson's approximation was usually larger than, but sometimes equal to, the Taylor series estimates. Although these differences were not necessarily significant, Dickerson's method was on average consistently more conservative. These same results were also observed for the unbiased heritability of volume and rust resistance (Table 2) and were also reported by Dickerson (1969).

Empirical estimate versus Dickerson and Taylor series approximations

The mean variance of h_b^2 for both volume growth and rust, estimated using the Dickerson, Taylor series and empirical methods (Figs. 1 and 2) were of the same magnitude, regardless of the procedure used to estimate

the variance components and the size of the test. However, the empirical estimate of the variance of h_b^2 was usually larger than that estimated by either of the approximate methods. The more conservative nature of the Dickerson estimates means that these estimates were generally closer to the empirical estimates than were the Taylor series estimates.

The only consistent exception to this observed underestimation of the empirical variance was when using REML estimates from the small tests. In this case the approximate variances were nearly the same as, or larger than, the empirical estimate of variance. It is possible that the difference between HM3 and REML in small tests reflected the fact that asymptotic variances and covariances were used for REML estimates. Therefore, the Taylor series and Dickerson approximations, which use the asymptotic variances and covariances, would have been affected by sample size. Thus, it might reasonably be expected that estimates from larger tests would be more reliable than those from small tests when using REML.

Unbiased heritability estimates (h^2) also showed a good relationship between the empirical variance estimates and the Dickerson and Taylor series approximations (Table 2). However, for rust resistance the approximate methods were consistently low, but only

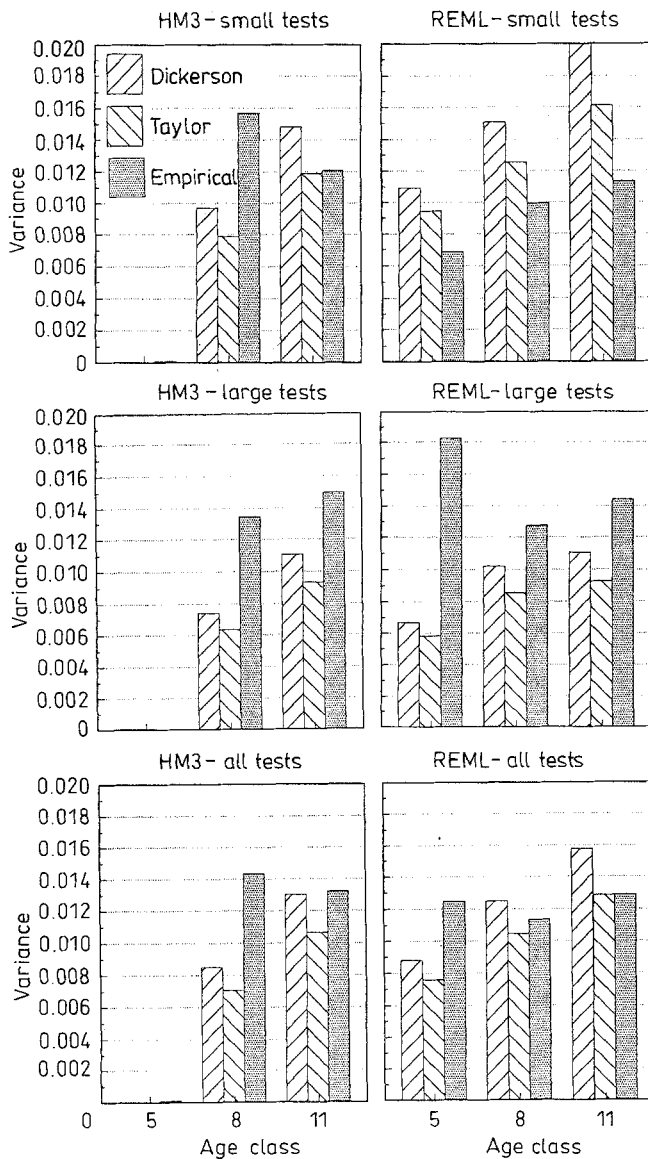


Fig. 1 Mean variance of the biased heritability (h_b^2) of tree volume in slash pine, estimated using Dickerson's approximation, and a Taylor series approximation compared to an empirical estimate of the variance. Heritability estimates are obtained from variance components estimated using Henderson's Method 3 (HM3) and Restricted Maximum Likelihood (REML), at 5, 8, and 11 years, in small (less than 16 parents), large (16 or more parents), and all full-sib tests

significantly so at 8 years. It is possible that the small number of test pairs (less than 30) involved in the rust resistance estimates of Table 2 may have been a contributing factor. Nevertheless, the estimates obtained from the two approximate methods were always of a similar magnitude to the empirical estimates, and there is no strong evidence to suggest that the approximate methods consistently over or underestimated the 'true' variance of paired test, unbiased heritability estimates (for volume the approximations were very close, but for rust the approximations were similar to, but less than the empirical estimates).

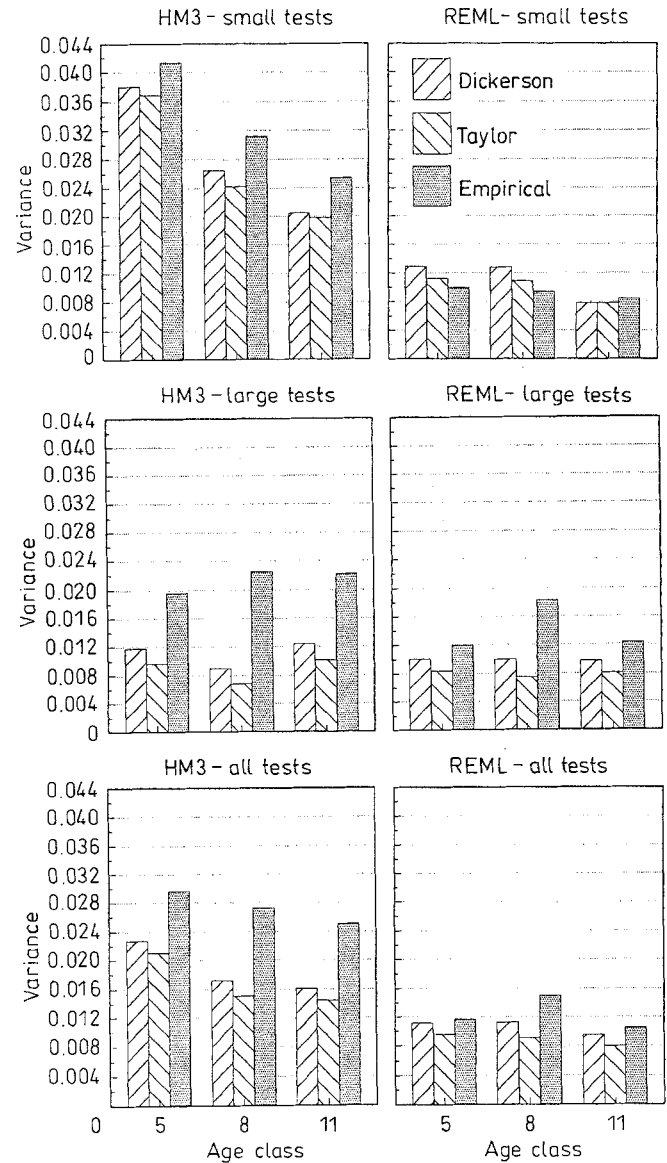


Fig. 2 Mean variance of the biased heritability (h_b^2) of rust resistance in slash pine, estimated using Dickerson's approximation, and a Taylor series approximation compared to an empirical estimate of the variance. Heritability estimates are obtained from variance components estimated using Henderson's Method 3 (HM3) and Restricted Maximum Likelihood (REML), at 5, 8, and 11 years, in small (less than 16 parents), large (16 or more parents), and all full-sib tests

It is unlikely that all the assumptions required in the development and application of the Dickerson and Taylor series approximations were fulfilled for both HM3 and REML variance component estimates. Nevertheless, these two methods provided reasonable approximations to the empirical variance of biased and unbiased heritabilities for both volume and rust resistance in slash pine. The observed deviations of both the Dickerson and Taylor series approximations from the empirical estimates of the variance only serves to highlight the fact that these are approximate and not exact methods.

Conclusions

The data presented indicate that Dickerson's method provided estimates of the variance of h_b^2 and h^2 that corresponded closely to estimates obtained from the Taylor series approximation, but which were slightly more conservative (i.e., larger). This conclusion was found to be valid for both tree volume (a continuous trait) and fusiform rust resistance (a bernoulli trait) in slash pine, and was not dependent on the method of variance component estimation or on test size.

Comparisons of the Dickerson's and Taylor series approximations with a empirical estimates of the variance of h_b^2 and h^2 indicated that these are both reasonable approximations to the 'true' variance (taken to be the empirical estimate) of h_b^2 and h^2 . However, these approximate variances tended to underestimate the 'true' variance of h_b^2 , while for h^2 there is no strong evidence to suggest that the approximate methods consistently over- or underestimated the empirical variance. When REML estimates from small tests (less than 16 parents in this case) are used, it is possible that the asymptotic variance and covariances may not be reliable, causing an over-estimation of the empirical variance of h_b^2 .

For most practical purposes when seeking to estimate the variance of biased or unbiased heritability estimates, it seems reasonable to conclude that either the Dickerson or the Taylor series approximations will provide useful estimates of the 'true' variance. However, Dickerson's method seemed to be preferable to the Taylor series approximation because (1) both methods provided essentially the same information about the variance of heritability estimates, (2) it was more conservative than the Taylor series approximation and thus perhaps closer to the presumed 'true' variance, and (3) the Dickerson method was considerably simpler to calculate.

Acknowledgements We gratefully acknowledge members of the Cooperative Forest Genetics Research Program who supplied the data that made these analyses possible. The authors also wish to thank Drs. D. L. Rockwood and M. DeLorenzo for their critical review of this paper.

Appendix

Derivation of the first-order Taylor series approximation of the variance of a ratio of two random variables

Kempthorne (1957, p 246), Becker (1975, p 45), and Namkoong (1979, p. 232) present without proof or derivation formulae for the variance of the ratio of two random variables based on a first-order Taylor series expansion that are numerically equivalent to

$$\text{Var}(g) \cong \text{Var}(x_1)/\theta_1^2 + \theta_1^2 \text{Var}(x_2)/\theta_2^4 - 2\theta_1 \text{Cov}(x_1, x_2)/\theta_2^3, \quad (\text{A1})$$

where $g = x_1/x_2$, $E(x_1) = \theta_1$, $E(x_2) = \theta_2$, and x_1 and x_2 are random variables.

In order to derive this formula let us first consider the multivariate function $g(x_1, x_2, \dots, x_k)$. If $g'_i(\theta)$ is $\partial g(x_1, x_2, \dots, x_k)/\partial x_i$ evaluated at

$\theta_1, \theta_2, \dots, \theta_k$, from Kendall et al. (1987) the first-order Taylor series expansion of $g(x_1, x_2, \dots, x_k)$ is

$$g(x_1, x_2, \dots, x_k) = g(\theta) + \sum_i g'_i(\theta)(x_i - \theta_i) + O(n^{-r}) \quad (\text{A2})$$

where θ is the vector $\theta_1, \theta_2, \dots, \theta_k$, $g(\theta)$ is $g(x_1, x_2, \dots, x_k)$ evaluated at θ , and $O(n^{-r})$ is a remainder term.

In the following discussion Eq. A1 is derived from the first-order expansion of the Taylor series given in Eq. A2. For brevity $g(x_1, x_2, \dots, x_k)$ will be referred to simply as $g(x)$. In order to do this it is necessary to derive (1) the expected value of $g(x)$ and (2) the variance of $g(x)$, which are in turn used to derive the formula for the variance of a ratio given in Eq. A1.

The expected value of $g(x)$ is derived as follows:

$$\begin{aligned} E[g(x)] &= E[g(\theta) + \sum_i g'_i(\theta)(x_i - \theta_i) + O(n^{-r})] \\ &= E[g(\theta)] + E[\sum_i g'_i(\theta)(x_i - \theta_i)] + E[O(n^{-r})] \quad \text{Step 1} \\ &= g(\theta) + \sum_i E[g'_i(\theta)(x_i - \theta_i)] + O(n^{-r}) \quad \text{Step 2} \\ &= g(\theta) + \sum_i g'_i(\theta)E[(x_i - \theta_i)] + O(n^{-r}) \quad \text{Step 3} \\ &= g(\theta) + \sum_i g'_i(\theta)[E(x_i) - \theta_i] + O(n^{-r}) \quad \text{Step 4} \\ &= g(\theta) + O(n^{-r}) \quad (\text{A3}) \end{aligned}$$

Initially, Eq. A2 is substituted for $g(x)$, and Step 1 proceeds by distributing the expectation operator with respect to addition. In Step 2, $g(\theta)$ and $O(n^{-r})$ are constants and so are equal to their own expectations, and in a similar manner to Step 1 the expectation operator can be placed inside the summation over i . Steps 3 and 4 use the fact that $g'_i(\theta)$ and θ_i are constants. And in the final step, the expression $E(x_i) - \theta_i = 0$, because $E(x_i) = \theta_i$.

The variance of $g(x)$ is derived as follows:

$$\begin{aligned} \text{Var}[g(x)] &= \text{Var}[g(\theta) + \sum_i g'_i(\theta)(x_i - \theta_i) + O(n^{-r})] \\ &= \text{Var}[g(\theta)] + \text{Var}[\sum_i g'_i(\theta)(x_i - \theta_i)] \\ &\quad + \text{Var}[O(n^{-r})] + \text{Covariances} \quad \text{Step 1} \\ &= \text{Var}[\sum_i g'_i(\theta)(x_i - \theta_i)] + O(n^{-r})^* \quad \text{Step 2} \\ &= \text{Var}[g'_1(\theta)(x_1 - \theta_1) + g'_2(\theta)(x_2 - \theta_2) + \dots + g'_i(\theta)(x_i - \theta_i) \\ &\quad + \dots + g'_k(\theta)(x_k - \theta_k)] + O(n^{-r})^* \quad \text{Step 3} \\ &= \text{Var}[g'_1(\theta)(x_1 - \theta_1)] + \text{Var}[g'_2(\theta)(x_2 - \theta_2)] + \dots \\ &\quad + \text{Var}[g'_i(\theta)(x_i - \theta_i)] + \dots + \text{Var}[g'_k(\theta)(x_k - \theta_k)] \\ &\quad + 2\text{Cov}[g'_1(\theta)(x_1 - \theta_1), g'_2(\theta)(x_2 - \theta_2)] \\ &\quad + 2\text{Cov}[g'_1(\theta)(x_1 - \theta_1), g'_3(\theta)(x_3 - \theta_3)] + \dots \\ &\quad + 2\text{Cov}[g'_1(\theta)(x_1 - \theta_1), g'_i(\theta)(x_i - \theta_i)] + \dots \\ &\quad + O(n^{-r})^* \quad \text{Step 4} \\ &= \sum_{i=1}^k \text{Var}[g'_i(\theta)(x_i - \theta_i)] \\ &\quad + \sum_{i \neq j=1}^k \text{Cov}[g'_i(\theta)(x_i - \theta_i), g'_j(\theta)(x_j - \theta_j)] \\ &\quad + O(n^{-r})^* \quad \text{Step 5} \\ &= \sum_{i=1}^k \{g'_i(\theta)\}^2 \text{Var}[(x_i - \theta_i)] \\ &\quad + \sum_{i \neq j=1}^k g'_i(\theta) g'_j(\theta) \text{Cov}[(x_i - \theta_i), (x_j - \theta_j)] \\ &\quad + O(n^{-r})^* \quad \text{Step 6} \\ &= \sum_{i=1}^k \{g'_i(\theta)\}^2 \text{Var}(x_i) \\ &\quad + \sum_{i \neq j=1}^k g'_i(\theta) g'_j(\theta) \text{Cov}(x_i, x_j) + O(n^{-r})^* \quad (\text{A4}) \end{aligned}$$

The initial step in this derivation involves substitution of Eq. A2 for $g(x)$, and Step 1 proceeds by distributing the variance operator. Step 2 is based on the fact that $g(\theta)$ is a constant, and so has zero variance. In Step 2 the variance of $O(n^{-r})$ as well as the covariances between the x_i 's and $O(n^{-r})$ are now represented by the single term,

$O(n^{-r})^*$. The third step simply involves an expansion of the summation operator. Step 4 utilizes the theorem presented in Eq. 2 of the main paper. Step 5 is an alternative expression for Step 4 using summation notation. In Step 6 $g'_i(\theta)$ and $g'_j(\theta)$ are constants, and so they can be taken outside the variance and covariance operators. The final step involves that fact that the θ'_i s are constants and so contribute no variance or covariance.

It is now possible to derive a formula for the variance of a ratio: let $g(x_1, x_2) = x_1/x_2$. Thus, from Eq. A2, $g(x_1, x_2)$ is equal to

$$g(x_1, x_2) = g(x_1, x_2)|_{\theta_1\theta_2} + \partial g(x_1, x_2)/\partial x_1|_{\theta_1\theta_2}(x_1 - \theta_1) + \partial g(x_1, x_2)/\partial x_2|_{\theta_1\theta_2}(x_2 - \theta_2) + O(n^{-r})$$

where

$$g(x_1, x_2)|_{\theta_1\theta_2} = \theta_1/\theta_2$$

$$\partial g(x_1, x_2)/\partial x_1|_{\theta_1\theta_2} = 1/x_2|_{\theta_1\theta_2} = 1/\theta_2, \text{ and}$$

$$\partial g(x_1, x_2)/\partial x_2|_{\theta_1\theta_2} = -x_1/(x_2)^2|_{\theta_1\theta_2} = -\theta_1/(\theta_2)^2.$$

Therefore,

$$g(x_1, x_2) = \theta_1/\theta_2 + (1/\theta_2)(x_1 - \theta_1) + (-\theta_1/(\theta_2)^2)(x_2 - \theta_2) + O(n^{-r}) \quad (\text{A5})$$

From Eqs. A4 and A5 it is possible to estimate the variance of x_1/x_2 in the following manner

$$\begin{aligned} \text{Var}(x_1/x_2) &= \Sigma_{i=1}^2 \{g'_i(\theta)\}^2 \text{Var}(x_i) + \Sigma \Sigma_{i \neq j=1}^2 g'_i(\theta) g'_j(\theta) \text{Cov}(x_i, x_j) \\ &\quad + O(n^{-r}) \\ &= (1/\theta_2)^2 \text{Var}(x_1) + (-\theta_1/(\theta_2)^2)^2 \text{Var}(x_2) \\ &\quad + 2(1/\theta_2)(-\theta_1/(\theta_2)^2) \text{Cov}(x_1, x_2) + O(n^{-r}) \\ &= \text{Var}(x_1)/\theta_2^2 + \theta_1^2 \text{Var}(x_2)/\theta_2^4 - 2\theta_1 \text{Cov}(x_1, x_2)/\theta_2^3 \\ &\quad + O(n^{-r}) \end{aligned} \quad (\text{A6})$$

If we assume that $O(n^{-r})$ is negligible, then Eq. A6 is the same as Eq. A1 (Eq. 5 of the main paper). Thus, the three assumptions required are (1) all second- and higher-order terms contained in $O(n^{-r})$ are negligible, (2) all the covariances between the x_i 's and the remainder are relatively small, and (3) the values of $E(x_1) = \theta_1$, $E(x_2) = \theta_2$, $\text{Var}(x_1)$, $\text{Var}(x_2)$ and $\text{Cov}(x_1, x_2)$, are implicitly assumed to be known when in practice they are unknown and must be estimated from the data.

References

- Ades PK, Simpson JA, Eldridge KG, Eldridge RH (1992) Genetic variation in susceptibility to *Dothistroma* needle blight among provenance and families in *Pinus muricata*. Can J For Res 22: 1111–1117
- Baker RJ (1986) Selection indices in plant breeding. CRC Press, Boca Raton, Fla.
- Balocchi CE, Bridgwater FE, Zobel BJ, Jahromi S (1993) Age trends in genetic parameters for tree height in nonselected populations of loblolly pine. For Sci 39: 231–251
- Banks BD, Mao IL, Walter JP (1985) Robustness of the restricted maximum likelihood estimator derived under normality as applied to data with skewed distributions. J Dairy Sci 68: 1785–1792
- Barnes RD, Mullin LJ, Battle G (1992) Genetic control of fifth year traits in *Pinus patula* Schiede and Deppe. Silvae Genet 41: 242–248
- Becker WA (1975) Manual of quantitative genetics. Washington State University Press, Pullman, Wash.
- Borralho NMG, Cotterill PP, Kanowski PK (1992) Genetic parameters and gains expected from selection for dry weight in *Eucalyptus globulus* ssp. *globulus* in Portugal. For Sci 38: 80–94
- Bulmer MG (1980) The mathematical theory of quantitative genetics. Clarendon Press, Oxford, UK
- Carson SD (1989) Selecting *Pinus radiata* for resistance to *Dothistroma* needle blight. N Z J For Sci 19: 3–21
- Comstock RE, Moll RH (1963) Genotype-environment interactions. In: Hanson RE, Robinson HF (eds) Statistical genetics and plant breeding. NAS-NRC Pub. 982, Washington D.C., pp 164–194
- Cotterill PP (1987) Short note: on estimating heritability according to practical applications. Silvae Genet 36: 46–48
- Cotterill PP, Dean CA, van Wyk G (1987) Additive and dominance genetic effects in *Pinus pinaster*, *P. radiata* and *P. elliottii* and some implications for breeding strategy. Silvae Genet 36: 221–232
- Dean CA, Cotterill PP, Eisemann RL (1986) Genetic parameters and gain expected from selection in *Pinus caribaea* var *hondurensis* in northern Queensland, Australia. Silvae Genet 35: 229–236
- Dickerson GE (1962) Implications of genetic-environmental interaction in animal breeding. Anim Prod 4: 47–64
- Dickerson GE (1969) Techniques for research in quantitative animal genetics. In: Techniques and procedures in animal science research. Am Soc Anim Sci, Albany, N.Y., pp 36–79.
- Dieters MJJ, Woolaston RR, Nikles DG (1990) Internode length of hoop pine: genetic parameters and prospects for developing a long-internode breed. N Z J For Sci 20: 138–147
- Dieters MJJ (1994) Inheritance of volume and rust resistance in slash pine. PhD thesis, Department of Forestry, University of Florida, Fla.
- Falconer DS (1989) Introduction to quantitative genetics. Longman Scientific and Technical, Longman Group, Hong Kong
- Giesbrecht FG (1983) An efficient procedure for computing MINQUE of variance components and generalized least squares estimates of fixed effects. Commun Stat Theor Method 12: 2169–2177
- Goddard RE, Strickland RK (1968) Volume and weight tables for five-year-old plantation grown slash pine. School of Forest Resources, Rep. No. 14, University of Florida, Gainesville, Fla.
- Griffing B (1956) Concept of general and specific combining ability in relation to diallel crossing systems. Aust J Biol Sci 9: 463–493
- Hallauer AR, Miranda JB (1988) Quantitative genetics in maize breeding, 2nd edn. Iowa State University Press, Ames, Iowa
- Hartley HO, Rao JNK (1967) Maximum-likelihood estimation for the mixed analysis of variance model. Biometrika 54: 93–108
- Harville DA (1977) Maximum likelihood approaches to variance component estimation and to related problems. J Am Stat Assoc 72: 320–338
- Hodge GR, White TL (1992) Genetic parameter estimates for growth traits at different ages in slash pine and some implications for breeding strategy. Silvae Genet 41: 252–262
- Huber DA (1993) Optimal mating designs and optimal techniques for analysis of quantitative traits in forest genetics. PhD thesis, Department of Forestry, University of Florida, Fla.
- Johnson GR, Burdon RD (1990) Family-site interaction in *Pinus radiata*: implications for progeny testing strategy and regionalised breeding in New Zealand. Silvae Genet 39: 55–62
- Kaplan W (1952) Advanced calculus. Addison-Wesley Publ Co, Reading, Mass.
- Kempthorne O (1957) An introduction to genetic statistics. John Wiley and Sons, New York
- Kendall MG, Stuart A, Ord JK (1987) Kendall's advanced theory of statistics, vol 3: Design and analysis, and time series. Oxford University Press, New York
- Khuri AI, Sahai H (1985) Variance components analysis: a selective literature survey. Int Stat Rev 53: 279–300
- King JN, Yeh FC, Heaman JCH, Dancik BP (1988) Selection of wood density and diameter in controlled crosses of coastal Douglas-fir. Silvae Genet 37: 152–157
- Kiss G, Yeh FC (1988) Heritability estimates for height for young interior spruce in British Columbia. Can J For Res 18: 158–162
- Knapp SJ, Stroup WW, Ross WM (1985) Exact confidence intervals for heritability on a progeny mean basis. Crop Sci 25: 192–194
- Knapp SJ, Bridges WC, Yang M (1989) Non-parametric confidence interval estimators for heritability and expected selection response. Genetics 121: 891–898

- Lowe WJ, Greene TA (1990) Geographic variation in specific gravity and fiber length of green ash (*Fraxinus pennsylvanica* Marsh.) in east Texas. *Silvae Genet* 39:194–198
- Mendenhall W, Wackerly DD, Scheaffer RL (1990) Mathematical statistics with applications, 4th edn. PWS-Kent Publ Co, Boston, Mass.
- Namkoong G (1979) Introduction to quantitative genetics. USDA Forest Service, Technical Bulletin No. 1588
- Rockwood DL (1981) Volume prediction for genetically improved slash pine trees. Agric Exp Stn Inst Food Agric Sci, University of Florida, Gainesville, Bulletin 819 (Technical)
- SAS Institute Inc (1988) SAS/STAT user's guide, release 6.03 edition. SAS Institute Inc, Cary, N.C.
- Schaffer HE, Usanis RA (1969) General least squares analysis of diallel experiments. A computer program – DIALL. Genet Dep Res No 1, North Carolina State University, Raleigh N.C.
- Searle SR, Casella G, McCulloch CE (1992) Variance components. John Wiley and Sons, New York
- Sluder ER (1993) Results at age 15 years from a half-diallel cross among 10 loblolly pines selected for resistance to fusiform rust (*Cronartium quercuum* f. sp. *fusiforme*). *Silvae Genet* 42:223–230
- Steel GD, Torrie JH (1981) Principles and procedures of statistics: a biometrical approach, 2nd edn. McGraw-Hill Book Co, Singapore
- Swallow WH, Monahan JF (1984) Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components. *Technometrics* 26:47–57
- Swiger LA, Harvey WR, Everson DO, Gregory KE (1964) The variance of intraclass correlation involving groups with one observation. *Biometrics* 20:818–826
- Tallis GM (1959) Sampling errors of genetic correlation coefficients calculated from analyses of variance and covariance. *Aust J Stat* 1:35–43
- Westfall PH (1987) A comparison of variance component estimates of arbitrary underlying distributions. *J Am Stat Assoc* 82:866–873
- White TL, Hodge GR (1989) Predicting breeding values with applications in forest tree improvement. (Forestry Sciences Series vol 33). Kluwer Academic Publ, Dordrecht
- White TL, Hodge GR (1992) Test designs and optimum age for parental selection in advanced generation progeny test of slash pine. *Silvae Genet* 41:293–302
- Winer BJ, Brown DR, Michels KM (1991) Statistical principles in experimental design, 3rd edn. (McGraw-Hill Series in Psychology). McGraw-Hill, New York
- Woolaston RR, Kanowski PJ, Nikles DG (1990) Genetic parameter estimates for *Pinus caribaea* var 'hondurensis' in coastal Queensland, Australia. *Silvae Genet* 39:21–28
- Yeh FC, Heaman C (1982) Heritabilities and genetic and phenotypic correlations for height and diameter in coastal Douglas-fir. *Can J For Res* 12:181–185