Theor Appl Genet (1994) 89:153–159

J. Zhu · B. S. Weir

Analysis of cytoplasmic and maternal effects I. A genetic model for diploid plant seeds and animals

Received: 28 July 1993 / Accepted: 27 September 1993

Abstract A genetic model for modified diallel crosses is proposed for estimating variance and covariance components of cytoplasmic, maternal additive and dominance effects, as well as direct additive and dominance effects. Monte Carlo simulations were conducted to compare the efficiencies of minimum norm quadratic unbiased estimation (MINQUE) methods. For both balanced and unbalanced mating designs, MINQUE (0/1), which has 0 for all the prior covariances and 1 for all the prior variances, has similar efficiency to MINQUE(θ), which has parameter values for the prior values. Unbiased estimates of variance and covariance components and their sampling variances could be obtained with MINQUE(0/1) and jackknifing. A t-test following jackknifing is applicable to test hypotheses for zero variance and covariance components. The genetic model is robust for estimating variance and covariance components under several situations of no specific effects. A MINQUE(0/1) procedure is suggested for unbiased estimation of covariance components between two traits with equal design matrices. Methods of unbiased prediction for random genetic effects are discussed. A linear unbiased prediction (LUP) method is shown to be efficient for the genetic model. An example is given for a demonstration of estimating variance and covariance components and predicting genetic effects.

Key words Modified diallel crosses · Monte Carlo simulation · Cytoplasmic and maternal effects · Variance and covariance components · Genetic prediction

Communicated by A. L. Kahler

J. Zhu
Department of Agronomy, Zheijiang Agricultural University, Hangzhou, Zheijiang, China

B. S. Weir (✉)
Program in Statistical Genetics, Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203, USA

Introduction

Diallel mating designs with reciprocal crosses provide a way for analyzing extranuclear effects. Cytoplasmic and maternal effects are the major sources of extranuclear effects for plant seeds (Mosjidis and Yermanos 1984), and maternal effects are an important component of extranuclear effects for mammals (Gowe and Fairfull 1982). The characterization of extranuclear effects on quantitative traits is of importance for seed quality improvement and animal breeding. Several genetic models have been proposed for analyzing extranuclear effects as well as nuclear effects. Henderson (1948) and Griffing (1956) provided diallel models including reciprocal effects that are the average extranuclear effects. Topham (1966) proposed diallel analysis for maternal and maternal interaction effects. Cockerham and Weir (1977) partitioned extranuclear effects into maternal effects and paternal effects by a bio-model for diallel crosses. Mather and Jinks (1982) suggested the use of parents, F_1, F_2, backcross, and their reciprocal crosses for estimating maternal additive and dominance effects. Eisen et al. (1983) expanded the models of Gardner and Eberhart (1966) and Vencovsky (1970) to include maternal effects for diallel crosses among random mating lines. Considering cytoplasmic genes as the only source of maternal effects, Beavis et al. (1987) proposed a model including cytoplasmic effects and nuclear-cytoplasmic interactions. Foolad and Jones (1992) introduced genetic models for analyzing quantitative seed characters. The model includes testa, cytoplasm, and embryo effects as well as endosperm effects.

Monte Carlo simulations have been used recently to evaluate procedures for estimating variance components (Tan and Shiue 1982; Swallow and Monahan 1984). Keele and Harvey (1989) compared the efficiencies of several methods for estimating variance components of direct and maternal breeding values and covariances between direct and maternal values by an additive genetic model (Quaas and Pollack 1980).

In this paper, a genetic model is proposed for quantitative traits influenced by genes with cytoplasmic effects, maternal additive and dominance effects, as well as direct additive and dominance effects. Unbiasedness and efficiency of minimum norm quadratic unbiased estimation (MINQUE) (Rao 1970, 1971) for variance and covariance components in the genetic model are tested by Monte Carlo simulations. The robustness of the model is examined under several situations of no specific variation. A MINQUE procedure is suggested for estimating covariance components between two traits with equal design matrices. A method of linear unbiased prediction (LUP) for genetic effects is compared to the best linear unbiased prediction (BLUP) procedure. Data for seed-oil content in Upland cotton (*Gossypium hirsutum* L.) from Dani and Kohel (1989) are used as an example of estimating variance and covariance components and for predicting genetic effects.

Model and methodology

Willham (1963) developed a theory for quantitative traits influenced by maternal effects. The phenotypic value of individuals can be partitioned as

$$y = \mu + G_o + G_m + E_m + E_o$$

where μ is the population mean, G_o and E_o are direct genetic and environmental effects, and G_m and E_m are maternal genetic and common environmental effects. The maternal common environmental effect can be eliminated by special experimental designs, e.g., randomized complete block design. Then E_m is replaced by block effect B:

$$y = \mu + G_o + G_m + B + E_o$$

Both inbred lines and F_1s can be used as maternal parents for producing plant seeds or animal offspring. They may have the same cytoplasmic effects but different maternal genetic effects on their offspring. Therefore, cytoplasmic effects should be distinguised from maternal genetic effects. A genetic model including cytoplasmic effects C is

$$y = \mu + G_o + C + G_m + B + E_o$$
$$= \mu + G + B + \varepsilon$$

with total genetic effect $G = G_o + C + G_m$ and residual error $\varepsilon = E_o$.

Modified diallel crosses consisting of F_1s and reciprocal F_1s from a set of completely inbred lines, and backcrosses of F_1s to their two parents are used in a genetic model with cytoplasmic and maternal effects for diploid plant seeds or animals. When Cockerham's (1980) general genetic model is expanded by adding cytoplasmic and maternal genetic effects but excluding epistatic effects, the total genetic effect G can be expressed as

$$G = \sum_i \tau_i A_i + \sum_{i \le j} \delta_{ij} D_{ij} + \sum_i \gamma_i C_i + \sum_i \tau_{m_{ij}} A_m + \sum_{i \le j} \delta_{m_{ij}} D_{m_{ij}}$$

where A_i is the cumulative additive effect of direct genes from line i, $A_i \sim (0, \sigma_A^2)$; the cumulative dominance effect of direct genes is $D_{ij} \sim (0, \sigma_D^2)$; the cytoplasmic effect is $C_i \sim (0, \sigma_C^2)$; the cumulative additive effect of maternal genes is $A_{m_i} \sim (0, \sigma_{A_m}^2)$; and the cumulative dominance effect of maternal genes is $D_{m_{ij}} \sim (0, \sigma_{D_m}^2)$. There are covariances between direct and maternal gene effects, $\mathrm{Cov}(A_i, A_{m_i}) = \sigma_{A.A_m}$ and $\mathrm{Cov}(D_{ij}, D_{m_{ij}}) = \sigma_{D.D_m}$. The genetic model assumes (1) inbred parents randomly sampled from a reference population; (2) no paternal effects; (3) no maternal interaction effects; (4) no epistatic effects; (5) no genotype-by-environment interaction; and (6) constant inheritance of cytoplasmic genes through maternal lines. If some of the first four assumptions are not valid, a more complicated model should be constructed to include the appropriate effects.

The genetic model can be rewritten as a linear model for the mean observation in the lth block of the kth type of genetic entry from lines i and j.

$$y_{ijkl} = \mu + G_{ijk} + B_l + \varepsilon_{ijkl} \tag{1}$$

where the total genetic effect G_{ijk} depends on the specific genetic entry; for $F_{1_{ij}}$ from maternal line $i \times$ paternal line j $(k = 1)$:

$$G_{ij1} = A_i + A_j + D_{ij} + C_i + 2A_{m_i} + D_{m_{ii}}$$

for backcross BC_i from maternal $F_{1_{ij}} \times$ paternal line $i (k = 2)$:

$$G_{ij2} = 1.5A_i + 0.5A_j + 0.5D_{ii} + 0.5D_{ij} + C_i + A_{m_i} + A_{m_j} + D_{m_{ij}}$$

and for backcross BC_j from maternal $F_{1_{ij}} \times$ paternal line j $(k = 3)$:

$$G_{ij3} = 0.5A_i + 1.5A_j + 0.5D_{jj} + 0.5D_{ij} + C_i + A_{m_i} + A_{m_j} + D_{m_{ij}}$$

These modified diallel crosses with F_1s, reciprocal F_1s, and their backcrosses are suitable for cross-pollinated crops or animals. For some self-pollinated crops, F_2 seed can be easily obtained from F_1 plants, and F_2 can then be included in the genetic model. The total genetic effect for $F_{2_{ij}} (k = 4)$ is

$$G_{ij4} = A_i + A_j + 0.25D_{ii} + 0.25D_{jj} + 0.5D_{ij} + C_i + A_{m_i} + A_{m_j} + D_{m_{ij}}$$

The genetic model can be written in a matrix form for all entries of the mating design,

$$\mathbf{y} = 1\mu + \mathbf{U}_A\mathbf{e}_A + \mathbf{U}_D\mathbf{e}_D + \mathbf{U}_C\mathbf{e}_C + \mathbf{U}_{A_m}\mathbf{e}_{A_m} + \mathbf{U}_{D_m}\mathbf{e}_{D_m} + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_e$$
$$= 1\mu + \sum_{u=1}^{7} \mathbf{U}_u\mathbf{e}_u \tag{2}$$

with variance-covariance matrix

$$\begin{aligned}\mathrm{Var}(\mathbf{y}) = {} & \sigma_A^2\mathbf{U}_A\mathbf{U}_A' + \sigma_D^2\mathbf{U}_D\mathbf{U}_D' + \sigma_C^2\mathbf{U}_C\mathbf{U}_C' \\ & + \sigma_{A_m}^2\mathbf{U}_{A_m}\mathbf{U}_{A_m}' + \sigma_{D_m}^2\mathbf{U}_{D_m}\mathbf{U}_{D_m}' + \sigma_B^2\mathbf{U}_B\mathbf{U}_B' \\ & + \sigma_{A.A_m}(\mathbf{U}_A\mathbf{U}_{A_m}' + \mathbf{U}_{A_m}\mathbf{U}_A') + \sigma_{D.D_m}(\mathbf{U}_D\mathbf{U}_{D_m}' + \mathbf{U}_{D_m}\mathbf{U}_D') \\ & + \sigma_e^2\mathbf{I} \\ = {} & \sum_{u=1}^{9} \theta_u\mathbf{V}_u\end{aligned}$$

where $\mathbf{U}_u$ is the known incidence matrix relating to the random vector $\mathbf{e}_u \sim (\mathbf{0}, \sigma_u^2\mathbf{I})$ for $u = 1, 2, \ldots, 7$; $\mathbf{U}_u'$ is the transpose of $\mathbf{U}_u$, $\mathbf{U}_7 = \mathbf{I}$ is an identity matrix; $\mathbf{V}_u = \mathbf{U}_u\mathbf{U}_u'$ for $u = 1, 2, \ldots 6$; $\mathbf{V}_7 = (\mathbf{U}_1\mathbf{U}_4' + \mathbf{U}_4\mathbf{U}_1')$, $\mathbf{V}_8 = (\mathbf{U}_2\mathbf{U}_5' + \mathbf{U}_5\mathbf{U}_2')$, and $\mathbf{V}_9 = \mathbf{I}$.

The MINQUE method (Rao 1970, 1971) can be used to estimate variances and covariances for the genetic model. Estimators of variance and covariance components $\hat{\theta}$ for the genetic model can be obtained by solving the following MINQUE equations for u, $v = 1, 2, \ldots, 9$;

$$[\mathrm{tr}(\mathbf{Q}_\alpha\mathbf{V}_u\mathbf{Q}_\alpha\mathbf{V}_v)][\hat{\theta}_u] = [\mathbf{y}'\mathbf{Q}_\alpha\mathbf{V}_u\mathbf{Q}_\alpha\mathbf{y}] \tag{3}$$

where

$$\mathbf{Q}_\alpha = \mathbf{V}_{(\alpha)}^{-1} - \mathbf{V}_{(\alpha)}^{-1}\mathbf{1}(\mathbf{1}'\mathbf{V}_{(\alpha)}^{-1}\mathbf{1})^{-1}\mathbf{1}'\mathbf{V}_{(\alpha)}^{-1}$$

$$\mathbf{V}_{(\alpha)} = \sum_{u=1}^{9} \alpha_u\mathbf{V}_u$$

and tr is the trace of a matrix (the sum of the diagonal terms).

The estimates are unbiased, provided the choice of prior values α_u do not depend on the data. MINQUE(θ) with the parameter values as the prior values gives the minimum variance invariant unbiased estimators for linear functions of variance components under the normality assumption (Rao 1972). MINQUE(0/1) is a MINQUE with prior values choosing 0 for covariances and 1 for variances, and with a much simpler matrix form

$$\mathbf{V}_{(0/1)} = \sum_{u=1}^{6} \mathbf{V}_u + \mathbf{I} = \sum_{u=1}^{6} \mathbf{U}_u\mathbf{U}_u' + \mathbf{I}$$

For the genetic model of diploid plant seeds and animal offspring, the phenotypic variance V_P can be partitioned as

$$V_P = V_{G_o} + V_C + V_{G_m} + 2C_{G_o \cdot G_m} + V_e$$
$$= (V_A + V_D) + V_C + (V_{A_m} + V_{D_m}) + 2(C_{A \cdot A_m} + C_{D \cdot D_m}) + V_e$$

where V_{G_o} is direct genetic variance with additive variance V_A and dominance variance V_D; V_C is cytoplasm variance; V_{G_m} is maternal genetic variance with maternal additive variance V_{A_m} and maternal dominance variance V_{D_m}; $C_{G_o \cdot G_m}$ is genetic covariance consisting additive covariance $C_{A \cdot A_m}$ and dominance covariance $C_{D \cdot D_m}$; V_e is residual variance.

There are usually several different generations (F_1, BC_i, BC_j, and/or F_2) in the mating design. Although each generation has the same kinds of variance components, phenotypic variance for each generation is calculated differently:

$$V_P(F_1) = (2\sigma_A^2 + \sigma_D^2) + \sigma_C^2 + (4\sigma_{A_m}^2 + \sigma_{D_m}^2) + 2(2\sigma_{A \cdot A_m}^2) + \sigma_e^2$$
$$V_P(F_2) = (2\sigma_A^2 + \tfrac{3}{8}\sigma_D^2) + \sigma_C^2 + (2\sigma_{A_m}^2 + \sigma_{D_m}^2) + 2(2\sigma_{A \cdot A_m}^2 + \tfrac{1}{2}\sigma_{D \cdot D_m}^2) + \sigma_e^2$$
$$V_P(BC_i) = V_P(BC_j)$$
$$= (2\tfrac{1}{2}\sigma_A^2 + \tfrac{1}{2}\sigma_D^2) + \sigma_C^2 + (2\sigma_{A_m}^2 + \sigma_{D_m}^2) + 2(2\sigma_{A \cdot A_m}^2 + \tfrac{1}{2}\sigma_{D \cdot D_m}^2) + \sigma_e^2$$

Methods of estimating covariance components with the MINQUE procedure for multiple traits have been discussed by Rao and Kleffe (1980). Those methods involve extensive computations and are limited by the number of traits involved. A much simpler MINQUE procedure for estimating covariance components can be derived for any number of traits for the genetic model. The covariance matrix of two variables $\mathbf{y}_a$ and $\mathbf{y}_b$ with equal design matrices is $\mathbf{V}_{a/b} = \sum_{v=1}^{9} \theta_{a_v/b_v} \mathbf{V}_v$ where

$\theta_{a_1/b_1} = \sigma_{A/A}$	is the additive covariance component for two traits,
$\theta_{a_2/b_2} = \sigma_{D/D}$	is the dominance covariance component,
$\theta_{a_3/b_3} = \sigma_{C/C}$	is the cytoplasmic covariance component,
$\theta_{a_4/b_4} = \sigma_{A_m/A_m}$	is the maternal additive covariance component,
$\theta_{a_5/b_5} = \sigma_{D_m/D_m}$	is the maternal dominance covariance component,
$\theta_{a_6/b_6} = \sigma_{B/B}$	is the block covariance component,
$\theta_{a_7/b_7} = \sigma_{A/A_m}$	is the average covariance component between direct and maternal additive effects,
$\theta_{a_8/b_8} = \sigma_{D/D_m}$	is the average covariance component between direct and maternal dominance effects,
$\theta_{a_9/b_9} = \sigma_{e/e}$	is the residual covariance component.

The expectation of the quadratic function $\mathbf{y}_\alpha'\mathbf{Q}_\alpha\mathbf{V}_u\mathbf{Q}_\alpha\mathbf{y}_b$ is

$$\mathrm{tr}(\mathbf{Q}_\alpha\mathbf{V}_u\mathbf{Q}_\alpha\mathbf{V}_{a/b}) = \sum_{v=1}^{9} \theta_{a_v/b_v} \mathrm{tr}(\mathbf{Q}_\alpha\mathbf{V}_u\mathbf{Q}_\alpha\mathbf{V}_v)$$

By MINQUE theory the invariant and unbiased estimators of covariance components can then be obtained by solving the following system of equations for $u, v = 1, 2, \ldots, 9$;

$$[\mathrm{tr}(\mathbf{Q}_\alpha\mathbf{V}_u\mathbf{Q}_\alpha\mathbf{V}_v)][\hat{\theta}_{a_u/b_u}] = [\mathbf{y}_a'\mathbf{Q}_\alpha\mathbf{V}_u\mathbf{Q}_\alpha\mathbf{y}_b] \quad (4)$$

The matrices $[\mathrm{tr}(\mathbf{Q}_\alpha\mathbf{V}_u\mathbf{Q}_\alpha\mathbf{V}_v)]$ and $\mathbf{Q}_\alpha\mathbf{V}_u\mathbf{Q}_\alpha$ in Eqs. 4 are the same as those in Eqs. 3. Therefore, they can be stored for later recall to estimate variances and covariances for multiple traits with the same design matrices.

For two traits, phenotypic covariance C_P can also be partitioned as

$$C_P = C_{G_o} + C_C + C_{G_m} + 2C_{G_o/G_m} + C_e$$
$$= (C_A + C_D) + C_C + (C_{A_m} + C_{D_m}) + 2(C_{A/A_m} + C_{D/D_m}) + C_e$$

where C_{G_o} is direct genetic covariance and has two components, direct additive covariance C_A and direct dominance covariance C_D; C_C is cytoplasm covariance; C_{G_m} is maternal genetic covariance with maternal additive covariance C_{A_m} and maternal dominance covariance C_{D_m}; $C_{G_o \cdot G_m}$ is genetic covariance between nuclear and maternal gene effects for two traits and has additive covariance C_{A/A_m} and dominance covariance C_{D/D_m}; C_e is residual covariance. Phenotypic covariance for each generation can be calculated by methods similar to variance calculation.

Sampling variances for estimates of variance and covariance components can be estimated by the jackknife procedure (Miller 1974; Efron 1982). If $\hat{\theta}$ is an estimate of a genetic parameter θ from a sample of L observations, and $\hat{\theta}_{(l)}$ is the estimate with the lth observation omitted, then the lth pseudovalue is

$$J_l(\hat{\theta}) = L\hat{\theta} - (L-1)\hat{\theta}_{(l)}$$

The jackknife estimator $J(\hat{\theta})$ of parameter θ is the mean of the pseudovalues. If L is not very large, $(J(\hat{\theta}) - \theta)/SE(J(\hat{\theta}))$ is approximately distributed as a t distribution with $(L-1)$ degrees of freedom.

In plant and animal breeding, breeders are sometimes interested in evaluating the genetic merits of parental lines. If the uth random vector $\mathbf{e}_u$ has independent components with $\mathrm{Var}(\mathbf{e}_u) = \sigma_u^2\mathbf{I}$, but is correlated with the vth random vector $\mathbf{e}_v$ with $\mathrm{Cov}(\mathbf{e}_u, \mathbf{e}_v) = \sigma_{u.v}\mathbf{I}$, then the best linear unbiased prediction (BLUP) (Henderson 1963) for the uth random vector of genetic effects can be obtained by

$$\hat{\mathbf{e}}_{u(\theta)} = (\sigma_u^2\mathbf{U}_u' + \sigma_{u.u}\mathbf{U}_u')\mathbf{V}_{(\theta)}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu})$$
$$= (\sigma_u^2\mathbf{U}_u' + \sigma_{u.v}\mathbf{U}_v')\mathbf{Q}_\theta\mathbf{y}$$

where

$$\hat{\mu} = (\mathbf{1}'\mathbf{V}_{(\theta)}^{-1}\mathbf{1})^{-1}\mathbf{1}'\mathbf{V}_{(\theta)}^{-1}\mathbf{y}$$
$$\mathbf{Q}_\theta = \mathbf{V}_{(\theta)}^{-1} - \mathbf{V}_{(\theta)}^{-1}(\mathbf{1}'\mathbf{V}_{(\theta)}^{-1}\mathbf{1})^{-1}\mathbf{1}'\mathbf{V}_{(\theta)}^{-1}$$

Since the true variances and covariances are unknown in practice, the unknown parameters θ can be replaced by prior values α from prior experiments, from estimates, or from reasonable guesses. Therefore, the genetic effects can be predicted by choosing prior values α as in the case of the MINQUE procedure for variance and covariance estimation,

$$\hat{\mathbf{e}}_{u(\alpha)} = (\alpha_u\mathbf{U}_u' + \alpha_{u.v}\mathbf{U}_v')\mathbf{Q}_\alpha\mathbf{y}$$

where σ_u^2 is replaced by the prior value α_u, and $\sigma_{u.v}$ by $\alpha_{u.v}$. In this study, 1 was chosen for prior variances α_u and 0 for prior covariances $\alpha_{u.v}$, to give the MINQUE(0/1) procedure of genetic effect prediction. It results in a linear unbiased prediction (LUP)

$$\hat{\mathbf{e}}_{u(0/1)} = \mathbf{U}_u'\mathbf{Q}_{(0/1)}\mathbf{y}$$

where $\mathbf{Q}_{(0/1)}$ is $\mathbf{Q}_\alpha$ with $\mathbf{V}_{(\alpha)}^{-1}$ replaced by $\mathbf{V}_{(0/1)}^{-1}$. The unbiasedness and efficiency of prediction by this method were compared to the BLUP $\hat{\mathbf{e}}_{u(\theta)}$. The distance between predictor vector $\hat{\mathbf{e}}$ and the sampling vector $\tilde{\mathbf{e}}$ is defined as

$$\|\hat{\mathbf{e}} - \tilde{\mathbf{e}}\| = \sqrt{\sum_v (\hat{e}_v - \tilde{e}_v)^2}.$$

Monte Carlo simulations were conducted in this study for estimating variance and covariance components by MINQUE(0/1) and MINQUE(θ). Covariance components for two traits were estimated by the MINQUE(0/1) procedure. Pseudo-random normal deviates with zero mean and unit variance were generated by the method of Kinderman and Monahan (1977). For each case 200 simulations were run to obtain sample means of estimates, bias, and Mean Squared Error (MSE). If the absolute value of bias is less than 10% of the parameter value, the parameter is said to be well estimated. In cases where the parameter value of variance or covariance component is zero, bias < 1% of the sum of variances and covariances is considered to be negligible. MSE is calculated by $[\mathrm{Var}(\hat{\theta}) + (\mathrm{bias})^2]$, which is usually used as a main criterion for comparing efficiency of estimation methods.

In this study, randomized complete block designs with three blocks were used with genetic entries assigned at random within each

block. The block was used as a resampling unit for the jackknife procedure. When estimated sampling variances of the estimates were obtained, the null hypotheses of no variation for random effects were tested. Power values (the probabilities of rejecting the null hypotheses) were obtained from 200 simulation runs. Since block variance is usually of not much concern to breeders, unbiased estimates for block variances are not presented in this paper.

Monte Carlo simulation results

All of the simulations were based on modified diallel crosses with three randomized complete blocks. The balanced design included F_1s, reciprocal F_1s, and their backcrosses from five inbred lines. Seven inbred lines were used for constructing unbalanced design by assuming that these lines are divided into two incompatible groups (lines one through five as group 1, and lines six and seven as group 2) within which lines could not mate to each other. Both the balanced design and the unbalanced design had the same experimental sizes.

For both balanced and unbalanced modified diallel crosses, MINQUE(0/1) and MINQUE(θ) give similar results for bias and MSE of estimated variance and covariance components when the correlation between direct genetic effects and maternal genetic effects is high ($\rho = 0.9$) or weak ($\rho = 0.1$). Therefore, MINQUE(0/1) is almost as efficient as MINQUE(θ) for estimating variance and covariance components for the genetic model with both balanced and unbalanced data.

Estimates and their sampling variances can be obtained by the MINQUE(0/1) method with the jackknife procedure. For three different levels of correlations between direct and maternal genetic effects ($\rho = 0.9$, 0.5, and 0.1) all of the variance components are well estimated with similar bias, MSE, and power value for both balanced and unbalanced mating designs. It is indicated that MINQUE(0/1) is equally efficient for estimating variance components of the genetic model under whatever correlations between direct and maternal genetic effects.

The simulation results of bias, MSE, and power value are summarized in Table 1 for the genetic model with a moderate correlation ($\rho = 0.5$) between direct and maternal effects. In this study of balanced five-parent and unbalanced seven-parent modified diallel crosses with three blocks, the significance of non-zero σ^2_A, $\sigma^2_{A_m}$, $\sigma^2_{D_m}$, and σ^2_e can be detected with a probability of over 75%. Significant σ^2_D and σ^2_C are detectable with a probability near 50%. The probability of detecting the significance of additive and dominance covariance components between direct and maternal effects is relatively low, and will increase (or decrease) when correlations become larger (or smaller). Robustness of estimation can be tested by simulation under the conditions of no specific variation. If there are no cytoplasmic effects ($\sigma^2_C = 0$), other parameters can be estimated with similar bias, MSE, and power value as when cytoplasmic effects are present. Non-significance of the cytoplasmic variance component can be detected with a probability over 96%. Without maternal genetic effects ($\sigma^2_{A_m} = \sigma^2_{D_m} = \sigma_{A.A_m} = \sigma_{D.D_m} = 0$) and/or cytoplasmic effects ($\sigma^2_C = 0$), direct additive and dominance variance components and error variance can be detected with a similar probability as when these effects exist. Conclusions of non-significance of maternal variance components and covariance components, and/or cytoplasmic variance components can be drawn with a probability around $94 \sim 99\%$. When there are no dominance variations ($\sigma^2_D = \sigma^2_{D_m} = \sigma_{D.D_m} = 0$), other parameters are well estimated. MSE is decreased and power value is increased

Table 1 Bias, MSE and power value from simulations by MINQUE(0/1) with the jackknife procedure for modified diallel crosses

Parameter	True value	Balanced design			Unbalanced design		
		Bias	MSE	Power[a]	Bias	MSE	Power[a]
One trait							
σ^2_A	25	−0.31	419.7	0.79	0.44	344.6	0.75
σ^2_D	16	−0.00	89.5	0.56	0.75	110.6	0.45
σ^2_C	10	0.49	63.4	0.47	0.25	52.1	0.43
$\sigma^2_{A_m}$	36	−2.44	908.2	0.81	−0.54	629.5	0.85
$\sigma^2_{D_m}$	25	1.34	169.8	0.84	0.84	178.9	0.84
$\sigma_{A.A_m}$	15	−0.48	304.1	0.52	0.23	255.9	0.50
$\sigma_{D.D_m}$	10	0.16	89.6	0.36	−0.05	88.5	0.28
σ^2_e	20	−0.07	6.3	0.99	0.06	5.5	0.99
Two traits							
$\sigma_{A/A}$	12.5	−0.87	287.1	0.37	0.69	196.2	0.35
$\sigma_{D/D}$	8	−0.13	54.7	0.17	−0.76	60.1	0.14
$\sigma_{C/C}$	5	0.09	41.0	0.18	−0.05	30.5	0.11
σ_{A_m/A_m}	18	−1.17	418.4	0.42	−0.21	385.5	0.46
σ_{D_m/D_m}	12.5	−0.42	108.6	0.38	−1.11	90.4	0.35
σ_{A/A_m}	15	−0.91	303.8	0.51	0.72	234.5	0.54
σ_{D/D_m}	10	−0.12	75.6	0.37	−1.30[b]	68.7	0.28
$\sigma_{e/e}$	10	0.11	3.3	0.70	0.10	3.7	0.67

[a] Probability of correctly rejecting the null hypothesis of no variation by the t-test with $\alpha = 0.05$

[b] Bias > 10% of the true value

for direct and maternal additive variance components. Non-significance of dominance variance and covariance components can be detected with a probability near 95%. The genetic model is robust for estimating variance and covariance components even though there are no cytoplasmic and maternal effects or no dominance effects.

Estimation of covariances between two traits were tested for unbiasedness and efficiency by MINQUE(0/1) with the jackknife procedure (Table 1). The estimate of average covariance between direct and maternal dominance effects, $\hat{\sigma}_{D/D_m}$, is slightly biased only for the unbalanced mating design. Other covariance components are well estimated for both balanced and unbalanced modified diallel crosses. The power values are over 50% for σ_{A/A_m} and $\sigma_{e/e}$, but less than 50% for other covariance components. More genetic materials are needed for constructing modified diallel crosses in order to detect the significance of covariance components of two traits.

In this study, the balanced and unbalanced mating designs had the same experimental sizes. There are no considerable differences of bias, MSE, and power value for variance and covariance components between these two mating designs. It is indicated that MINQUE(0/1) has almost an equal efficiency in estimating variance and covariance components for balanced and unbalanced modified diallel crosses with the same experimental sizes.

The linear unbiased prediction, $\hat{\mathbf{e}}_{u(0/1)}$ with prior values of 1 for all variance components and 0 for all covariance components, were compared to the BLUP $\hat{\mathbf{e}}_{u(\theta)}$ using parameter values. Simulations were conducted for estimating bias of predicted random effects and distance between predictor vector $\hat{\mathbf{e}}_u$ and sampling vector $\tilde{\mathbf{e}}_u$. Table 2 presents the simulation results with $A_i \sim N(0, 25)$, $D_{ij} \sim N(0, 16)$, $C_i \sim N(0, 10)$, $A_{m_i} \sim N(0, 36)$, $D_{m_i} \sim N(0, 25)$, $B_l \sim N(0, 20)$, $\varepsilon_{ij} \sim N(0, 20)$, $\mathrm{Cov}(A_i, A_{m_i}) = 15$, and $\mathrm{Cov}(D_i, D_{m_i}) = 10$. Both prediction methods give extremely low bias for mean of predicted genetic effects. Hence, MINQUE(0/1) gives unbiased prediction for random genetic effects just as the BLUP does. The BLUP $\hat{\mathbf{e}}_{u(\theta)}$ should give the smallest distance for the predicted genetic effects among all linear unbiased predictions (Henderson 1979). The distance of $\hat{\mathbf{e}}_{u(0/1)}$ to $\tilde{\mathbf{e}}_u$ approaches that of the BLUP $\hat{\mathbf{e}}_{u(\theta)}$ for all the genetic effects in balanced mating design. Except for dominance effect D, the distance of $\hat{\mathbf{e}}_{u(0/1)}$ is similar to that of the BLUP $\hat{\mathbf{e}}_{u(\theta)}$ for all of the other predicted genetic effects in unbalanced mating design. It is indicated that prediction choosing one for the prior variances and zero for the prior covariances is unbiased and quite efficient.

Example: seed-oil content in cotton

Data of a four-parent modified diallel cross with $F_{1_{ij}}$, $F_{1_{ij}} \times P_i$, $F_{1_{ij}} \times P_j$, and $F_{2_{ij}}$ from Dani and Kohel (1989) are used as an example for estimating variance and covariance components and for predicting the genetic merits of seed-oil percentage and seed-oil index (milligrams of oil per seed) in cotton (*Gossypium hirsutum* L.). Since only four parents would not normally be regarded as a random set of parents to justify the use of a random genetic model, the analysis serves mainly to demonstrate the application of the genetic model.

The cell mean of each genetic entry served as a resampling unit in the jackknife procedure. With jackknife estimates and their standard errors, a one-tail t-test was conducted for testing variance components while a two-tail t-test was used for testing covariances or genetic effects. The degrees of freedom for a t-test were 47. Estimates of variance and covariance components and their standard errors by MINQUE(0/1) method for cotton seed-oil content are summarized in Table 3. Highly significant variances were detected for maternal additive effects and residual errors of seed-oil percentage and seed-oil index. Since additive effects of maternal genes were the major contributions of genetic variation, commercial cotton varieties with high seed-oil content could be developed by selection based on maternal plants. Negative $\hat{\sigma}_C^2$ indicated no variation of cytoplasmic effects for these two traits. For seed-oil percentage, estimates of direct variance components and covariances between direct and maternal effects were small and not significant. Therefore, seed-oil percentage might be controlled mainly by maternal gene effects. For seed-oil index, significance of variance was observed for direct dominance effects. Although σ_A^2 was not significant, $\sigma_{A.A_m}$ was significantly negative. It was suggested that there might be some additive effects of seed genes and

Table 2 Prediction of genetic effects by LUP and BLUP methods for modified diallel crosses (Absolute bias for mean prediction of genetic effects is $10^{-5} \sim 10^{-6}$ for $\hat{\mathbf{e}}_{u(0/1)}$, and $10^{-3} \sim 10^{-5}$ for $\hat{\mathbf{e}}_{u(\theta)}$)

Parameter	Balanced design				Unbalanced design			
	Distance of $\hat{\mathbf{e}}_{u(0/1)}$		Distance of $\hat{\mathbf{e}}_{u(\theta)}$		Distance of $\hat{\mathbf{e}}_{u(0/1)}$		Distance of $\hat{\mathbf{e}}_{u(\theta)}$	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
A	5.51	5.64	5.48	5.66	6.3	5.99	6.30	5.95
D	11.20	7.15	10.78	6.73	12.27**	5.72	11.79	4.75
C	3.55	2.68	3.53	2.68	4.46	2.39	4.40	2.41
A_m	6.44	8.88	6.39	9.08	7.25	8.43	7.21	8.62
D_m	11.56	8.38	11.41	8.71	12.99	10.24	12.88	10.11

** Significantly different from the BLUP at the 0.01 significance level

that they could behave differently than those of maternal plant genes. Seed-oil index was controlled by both direct and maternal gene effects. About 64% of the total genetic variance was contributed by maternal genetic variances, among which 76% was maternal additive variance. Covariance components between two traits were significantly positive for maternal additive effects and residual errors. Other covariance components were negligible.

Predicted genetic effects and their standard errors are listed in Table 4 only for those with significant variances. Since maternal additive effects were the most important contribution of variance components for seed-oil content, the genetic merits of parental lines could be evaluated mainly based on the predicted maternal additive effects $\hat{A}_{m_i}$. Both parental line 1 (SA 1169) and line 4 (SA 59) were high seed-oil percentage lines. Line 1 ($\hat{A}_{m_1} = 1.049$) might be more superior to line 4($\hat{A}_{m_4} = 0.341$) for increasing seed-oil percentage in the breeding program. Though both parental line 1 and line 4 had high values of seed-oil index, their maternal additive effects ($\hat{A}_{m_1} = 2.258$, and $\hat{A}_{m_4} = -0.868$) were different. Heterosis for seed-oil content can be evaluated by the average of homozygous dominance effects. Since $(\Sigma_i \hat{D}_{ii})/4$ was -0.337 (± 0.101) for seed-oil index, inbreeding depression was expected for homozygous genotype of seed.

Discussion

Variances and covariances for the genetic model proposed in this paper can also be estimated by other mixed model approaches (Searle et al. 1992). As compared to methods of maximum likelihood (ML) (Hartley and Rao 1967) and restricted maximum likelihood (REML) (Patterson and Thompson 1971), MINQUE has the advantages of simple computation and no requirement for normality distribution. Estimates obtained by MINQUE methods are invariant and unbiased (Rao 1971). Though the choosing of different prior values will give unbiased estimates, they may not be equally efficient. MINQUE(1) suggested by Giesbrecht (1985) is a MINQUE method setting all the prior values as 1. Monte Carlo simulations have been conducted with MINQUE (1) for the genetic model (results are not presented in this paper). MINQUE(0/1) is more efficient than MINQUE(1), especially when the true covariances between direct and maternal genetic effects are very small. Another advantage of MINQUE (0/1) over MINQUE (1) is that MINQUE(0/1) has less computation for matrix $\mathbf{V}^{-1}_{(0/1)}$.

By MINQUE (0/1) with the jackknife procedure, unbiased estimates of direct, cytoplasmic and maternal variance, and covariance components could be obtained for the genetic model. It has been proved that the genetic model is robust under several situations of no specific effects. By the jackknife procedures, estimated sampling variances for estimates of variance and covariances components give powerful t-tests for detecting significance of variation. If the null hypothesis of zero variance or covariance is true, the power value (or P-value) is around the α- value (0.05) in most cases. The t-test for variance and covariance components appears to be appropriate.

MINQUE(0/1) with the jackknife procedure can also be conducted for unbiased estimation of genetic

Table 3 Estimates of variance and covariance components from cell mean data for seed-oil content in cotton

Variance	Seed-oil (%) Estimate ± SE	Seed-oil Index Estimate ± SE	Covariance	Oil (%) versus Index Estimate ± SE
σ^2_A	0.016 ± 0.119	1.727 ± 1.366	$\sigma_{A/A}$	0.098 ± 0.329
σ^2_D	0.092 ± 0.296	3.419* ± 1.958	$\sigma_{D/D}$	0.797 ± 0.699
σ^2_C	−0.046 ± 0.085	−0.020 ± 0.791	$\sigma_{C/C}$	−0.087 ± 0.185
$\sigma^2_{A_m}$	1.110** ± 0.381	4.710** ± 1.875	σ_{A_m/A_m}	2.013** ± 0.743
$\sigma^2_{D_m}$	0.338 ± 0.295	2.943 ± 2.001	σ_{D_m/D_m}	0.944 ± 0.706
$\sigma_{A.A_m}$	−0.029 ± 0.221	−3.599* ± 1.528	σ_{A/A_m}	−0.142 ± 0.471
$\sigma_{D.D_m}$	−0.056 ± 0.214	0.768 ± 1.159	σ_{D/D_m}	−0.282 ± 0.450
σ^2_e	0.567** ± 0.167	3.032** ± 1.014	$\sigma_{e/e}$	0.931* ± 0.376

* $P < 0.05$, ** $P < 0.01$

Table 4 Predicted genetic effects and standard errors from cell mean data for seed-oil content traits in cotton

	Parent			
	$i = 1$	$i = 2$	$i = 3$	$i = 4$
Seed–oil percentage				
$\hat{A}_{m_i}$	1.049** ± 0.197	−0.067 ± 0.171	−1.324** ± 0.186	0.341* ± 0.163
Seed–oil index (mg)				
$\hat{D}_{i1}$	−1.941** ± 0.449			
$\hat{D}_{i2}$	−1.415* ± 0.547	1.237* ± 0.521		
$\hat{D}_{i3}$	0.483 ± 0.536	−0.252 ± 1.033	−0.200 ± 0.683	
$\hat{D}_{i4}$	2.879** ± 0.603	−1.053* ± 0.546	−1.531* ± 0.882	−1.269* ± 0.491
$\hat{A}_{m_i}$	2.258** ± 0.374	−1.152** ± 0.416	−2.544** ± 0.572	−0.868* ± 0.469

* $P < 0.05$, ** $P < 0.01$

covariance components for multiple traits with the same design matrices. By this method covariance components between two traits are as easily estimable as in the one-trait case. In real situations for two variables $\mathbf{y}_a$ and $\mathbf{y}_b$ with equal design matrices, it is possible that $\mathrm{Cov}(A_{i_a}, A_{m_{i_b}}) \neq \mathrm{Cov}((A_{i_b}, A_{m_{i_a}})$, or $\mathrm{Cov}(D_{ij_a}, D_{m_{ij_b}}) \neq \mathrm{Cov}(D_{ij_b}, D_{m_{ij_a}})$. In this study, only average covariances of direct and maternal effects between two traits (σ_{A/A_m} and σ_{D/D_m}) are estimable.

In plant and animal breeding, genetic merits of breeding materials are sometimes more interesting to the breeder than the variance and covariance components. Random genetic effects in the genetic model are not estimable separately by constructing side conditions, but they are predictable by the BLUP procedure (Henderson 1963). Instead of using estimates of variances and covariances for predicting genetic effects, $\hat{\mathbf{e}}_{u(0/1)}$ with 0 for covariances and 1 for variances is easier to compute. It turns out that $\hat{\mathbf{e}}_{u(0/1)}$ is an unbiased and efficient predictor of $\mathbf{e}_u$.

We assume that all the genetic effects in the model are random effects. Inbred parents are considered as a random sample from a reference population. For the general analysis of diallel cross involving F_1s, a sample size of eight to ten parents is reasonable. Since the modified diallel cross including F_1s, BCs, and/or F_2s, the sample size is about three or four times larger than the general diallel analysis with the same number of parents. Plant breeders usually cannot afford very large sample sizes in conducting genetic research since seeds of F_1s and backcrosses should be produced by manual hybridization. Simulation studies suggest that quite good estimation could be obtained if five parents are selected at random for mating. When more parents are involved in mating, power of tests for detecting significance could be increased. Modified partial diallel crosses can be constructed for increasing sampling size but not experimental size.

Acknowledgments This research was supported in part by NIH Grant GM32518.

References

Beavis WD, Pollak E, Frey KJ (1987) A theoretical model for quantitatively inherited traits influenced by nuclear-cytoplasmic interactions. Theor Appl Genet 74:571–578

Cockerham CC (1980) Random and fixed effects in plant genetics. Theor Appl Genet 56:119–131

Cockerham CC, Weir BS (1977) Quadratic analyses of reciprocal crosses. Biometrics 33:187–203

Dani RG, Kohel RJ (1989) Maternal effects and generation mean analysis of seed-oil content in cotton (*Gossypium hirsutum* L.) Theor Appl Genet 77:569–575

Efron B (1982) The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia, Pa.

Eisen EJ, Hörstgen-Schwark G, Saxton AM, Bandy TR (1983) Genetic interpretation and analysis of diallel crosses with animals. Theor Appl Genet 65:17–23

Foolad MR, Jones RA (1992) Models to estimate maternally controlled genetic variation in quantitative seed characters. Theor Appl Genet 83:360–366

Gardner CO, Eberhart SA (1966) Analysis and interpretation of the variety cross diallel and related populations. Biometrics 22:439–452

Giesbrecht FG (1985) MIXMOD: A SAS procedure for analysing mixed models. Mimeo series No. 1659. Institute of Statistics, North Carolina State University, Raleigh, N.C.

Gowe RS, Fairfull RW (1982) Heterosis in egg-type chickens. In: 2nd World Congr Genet Appl Livestock Production 6:228–242, Madrid

Griffing B (1956) Concept of general and specific combining ability in relation to diallel crossing systems. Aust J Biol Sci 9:463–493

Hartley HD, Rao JNK (1967) Maximum-likelihood estimation for the mixed analysis of variance model. Biometrika 54:93–108

Henderson CR (1948) Estimation of general, specific and maternal combining abilities in crosses among inbred lines of swine. PhD thesis, Iowa State University, Ames, Iowa

Henderson CR (1963) Selection index and expected genetic advance. In: Hanson WD, Robinson HF (eds) Statistical genetics and plant breeding. Natl Acad Sci, NRC 982, pp 141–163

Henderson CR (1979) Using estimates in predictions of breeding values under a selection model. In: Van Vleck LD, Searle SR (eds) Variance components and animal breeding. Cornell University, Ithaca, N.Y., pp 217–227

Keele JW, Harvey WR (1989) Estimation of components of variance and covariance by symmetric differences squared and minimum norm quadratic unbiased estimation: a comparison. J Anim Sci 67:348–356

Kinderman AJ, Monahan JF (1977) Computer generation of random variables using the ratio of normal deviates. Assoc Comput Mach Trans Math Soft 3:257–260

Mather K, Jinks JL (1982) Biometrical genetics, 3rd edn. Chapman and Hall, London

Miller RG (1974) The jackknife – a review. Biometrika 61:1–15

Mosjidis JA, Yermanos DM (1984) Maternal effects and cytoplasmic inheritance of oleic and linoleic acid contents in sesame. Euphytica 33:427–432

Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. Biometrika 58:545–554

Quaas RL, Pollak EJ (1980) Mixed model methodology for farm and ranch beef cattle testing programs. J Anim Sci 51:1277

Rao CR (1970) Estimation of heteroscedastic variances in linear models. J Am Stat Assoc 65:161–172

Rao CR (1971) Estimation of variance and covariance components MINQUE theory. J Multivar Anal 1:257–275

Rao CR (1972) Estimation of variance and covariance components in linear models. J Am Stat Assoc 67:112–115

Rao CR, Kleffe J (1980) Estimation of variance components. In: Krishraiah PR (ed) Handbook of statistics, vol 1. North-Holland Publ, New York, pp 1–40

Searle SR, Casella G, McCulloch CE (1992) Variance components. Wiley, New York

Swallow WH, Monahan JF (1984) Monte Carlo comparison of ANOVA, MINQUE, REML and ML estimators of variance components. Technometrics 6:47–57

Tan NY, Shiue WW (1982) Comparative studies on the estimation of genetic variances by several variance component estimation methods. Biometrical J 24:543–557

Topham PB (1966) Diallel analysis involving maternal and maternal interaction effects. Heredity 21:665–674

Vencovsky R (1970) Alguns aspectors teóricos e aplicodos relativos a cruzamentos dialelicos de variedades. Tese de Livre-Docencia, ESALQ-USP, Piracicaba, Brazil

Willham RL (1963) The covariance between relatives for characters composed of components contributed by related individuals. Biometrics 19:18–27