



Simulation Studies for Comparing Genetics Models with Additive-Dominance-Maternal Effects and GE Interaction Effects

ABDERRAHMANE Achouch¹ ZHU Jun²

(1 Institute of Bioinformatics, Zhejiang University, Zhejiang Hangzhou 310029 China)

Abstract: Additive-dominance-maternal genetic models are employed for eight inbred lines by using mating design of diallel crosses with F_1 's and reciprocal F_1 's. Monte Carlo simulations are conducted for comparisons among models by including or ignoring maternal effects, maternal by environment effects and genotype by environment interaction effects. When σ_M^2 and σ_{ME}^2 exist in the full model, the significance of non-zero σ_A^2 and σ_M^2 can be detected around 97% and that for $\sigma_D^2, \sigma_{AE}^2, \sigma_{DE}^2, \sigma_{ME}^2$ and σ_e^2 can be more than 99%. When maternal and maternal by environment effects are ignored, overestimation of variance components is observed except for dominance variance. The overestimation is highest for σ_e^2 followed by σ_A^2 when particularly total genotype by environment effects is ignored in the additive-dominance model. When maternal effects and genotype by environment interaction effects are ignored, variances of predictor increase significantly.

Key words: Genetic variance component ; Predicted genetic effects ; Mixed model approach ; Monte Carlo simulation

CLC number: Q212.1; S330 **MR number:** 62J05; 62J10

Document code: A **Article ID:** 1001-9626(2002)02-0208-07

0 Introduction

Many kinds of genetic models were used to study phenotypic traits of the individuals or to analyze their differences, and their relationships (Eisen et al., 1966, Searle, 1971; Cockerham and Weir, 1977; Zhu, 1997). An understanding of the inheritance of these differences is of fundamental significance in the study of evolution and in the application of genetics to animals and plant breeding (Falconer 1996). Idea about polygene theory proposed by Johannsen (1909), Nilsson-Ehle (1909) and East (1916) indicates that a quantitative trait is controlled by many Mendelian factors or genes, groups of genes (polygene systems) each of which has small effects on the considered

Received date: 2000-04-19

Biography: Abderrahmane Achouch(1951-), male, genetics and plant breeding, associate professor.

trait and is easily influenced by the environments. For most of the quantitative traits, geneticists think that the phenotypic variation is due also, in large part, to the influence of the environments. Most of them are quantitative with continuous phenotypic variation, which can depend on the variations of main genetics effects and genotype by environment (*GE*) interaction effects. Many genetic models were developed and are still now analyzed by ANOVA methods. However, when the genetic model presents decimal values of coefficients and/or correlated random effects with unbalanced data, this statistical technique can not be applied. Improvement of quantitative genetics is due at the same time to the genetic studies and to the statistical methods merged progressively with the computational methods in data analysis (Weir, 1996; William, et al, 1992).

From 1967's statisticians developed some new methods capable to handle complicated models. These techniques based on the mixed linear models are Maximum likelihood (Hartley and Rao, 1967); Restricted Maximum likelihood (Patterson and Thompson, 1971) and Minimum Norm Quadratic Unbiased Estimation (MINQUE) (Rao, 1971a, b and 1972). Then many researches concerning the estimation of variance components were conducted (Henderson, 1979; Keele and Harvey, 1989; Searle et al, 1992; Zhu, 1992). Cockerham (1980) proposed a general genetic model for partitioning the total genotypic effect (*G*). Zhu (1994) extended Cockerham's general genetic model by including *GE* interaction. Monte Carlo simulations (Zhu, 1989; Zhu and Weir, 1994a, b, 1996) showed that MINQUE has more advantages than REML. It involves easier computation and needs no requirement for normality distribution. Monte Carlo simulation proved also that AUP would give unbiased mean and variance for predicted effects (Zhu and Weir, 1996).

The object of the present study is to conduct Monte Carlo simulation for comparing genetic full and reduced models with additive, dominance and maternal effects as well as genotype by environment interaction effects.

1 Models and Methodology

For the additive-dominance-maternal (*ADM*) models with *GE* interaction effects in multiple environments, the phenotypic mean of parent ($i = j$) and F_1 ($i \neq j$) in the k th block within the h th environment is

$$y_{hijk} = \mu + E_h + A_i + A_j + D_{ij} + M_i + AE_{hi} + AE_{hj} + DE_{hij} + ME_{hi} + B_{hk} + \varepsilon_{hijk} \quad (1)$$

where μ is the population mean, E_h is the environment effect; A_i or $A_j \sim (0, \sigma_A^2)$ is the additive effect; $D_{ij} \sim (0, \sigma_D^2)$ is the dominance effect; $M_i \sim (0, \sigma_M^2)$ is the maternal effect; AE_{hi} or $AE_{hj} \sim (0, \sigma_{AE}^2)$ is the additive by environment interaction effect; $DE_{hij} \sim (0, \sigma_{DE}^2)$ is the dominance by environment interaction effect; $ME_{hi} \sim (0, \sigma_{ME}^2)$ is the maternal by environment interaction effect; $B_k \sim (0, \sigma_B^2)$ is the block effect; $\varepsilon_{hijk} \sim (0, \sigma_\varepsilon^2)$ is the residual effect.

The *ADM* model in (1) can be expressed in a general matrix form of a mixed linear model,

$$y = Xb + U_{AE}e_A + U_D e_D + U_M e_M + U_{AE}e_{AE} + U_{DE}e_{DE} + U_{ME}e_{ME} + U_B e_B + e_\varepsilon = Xb + \sum_{u=1}^8 U_u e_u$$

where e_u are the random effects with a known coefficient matrix U_u .

When maternal effects and maternal by environment effects are ignored, the *ADM* model

can be reduced to the *AD* model in the following expression

$$y_{hijk} = \mu + A_i + A_j + D_{ij} + AE_{hi} + AE_{hj} + DE_{hij} + B_{hk} + \varepsilon_{hijk}. \quad (2)$$

The *AD* model in (2) can be presented in a general matrix form of a mixed linear model

$$y = Xb + U_A e_A + U_D e_D + U_{AE} e_{AE} + U_{DE} e_{DE} + U_B e_B + e_\varepsilon = Xb + \sum_{u=1}^6 U_u e_u.$$

ADM model and *AD* model are employed for genetic entries $i \times j$ for ($i = 1, 2, \dots, 8; j = 1, 2, \dots, 8$) by using mating design of diallel crosses with F_1 's and reciprocal F_1 's. Randomized complete block design of three blocks is used with a total of 192 genetics entries per block in multiple environments. The three environments are assumed fixed with values of 50, 100, 150, respectively. Genetic effects are assumed independent and random. The genetic entries are assigned at random within each block. Variance components are estimated by MINQUE (1) method and predicted for genetic effects by AUP method. Jackknifing over block method (Miller 1974; Efron, 1982) is conducted for estimating the standard error. Pseudo-random normal deviates with zero mean and unit variance (0,1) are generated by the method of Kinderman and Monahan (1977). Given the eight-parent number, F_1 's generation, true values of variance components and the three environment numbers, balanced data are used for *ADM* models and *AD* models. 500 Simulations are used for computing the sample mean of estimate, bias, M.S.E. and power value. Predicted mean, predicted variance and power are also computed for each specific effect of the *ADM* models and *AD* models. *MSE* is calculated by $Var(\hat{\theta}) + (bias)^2$, which is usually used as a main criterion for comparing the efficiency of estimation methods. Bias is calculated as $\hat{\theta} - \theta$. If $Bias/\theta \leq 5\%$, the estimate $\hat{\theta}$ is considered as unbiased (Graybill and Wortham, 1956). Sampling variance of estimates is calculated by $Var(\hat{\theta}) = \frac{1}{n-1} \sum (\hat{\theta} - \bar{\hat{\theta}})^2$.

2 Results and Analysis

Simulations are conducted for making comparison among *ADM* models according to the consideration of maternal effects and maternal by environment interaction effects (Table 1). When σ_M^2 and σ_{ME}^2 are significantly superior to zero in the full model, the power of detecting significance of non-zero σ_A^2 and σ_M^2 is around 97% and that for $\sigma_D^2, \sigma_{AE}^2, \sigma_{DE}^2, \sigma_{ME}^2$ and σ_ε^2 is more than 99%. Robustness of estimation is tested by simulation under the conditions of no specific variation. If $\sigma_M^2 = \sigma_{ME}^2 = 0$, all parameters are still unbiased and the variance components $\sigma_A^2, \sigma_D^2, \sigma_{AE}^2, \sigma_{DE}^2, \sigma_\varepsilon^2$ can be estimated with similar bias, *MSE* and power value to those when maternal effects and maternal by environment interaction effects are present. Significance of maternal and maternal by environment variance components can be falsely detected with probability of about 1%. When both maternal and maternal by environment interaction effects are ignored in the full models, bias and *MSE* increase highly and consequently the estimated variance components are overestimated except for σ_D^2 . The increases of biases of additive variance and additive by environment interaction variance are around 14.6% of the parameter values and that of dominance

by environment interaction variance component reaches 11.5%. Residual variance component is particularly overestimated with the highest value of bias around 244.2% of the parameter value.

Table 1 Bias, MSE and Power According to Maternal Effects and Maternal by Environment Interaction Effects

Parameter	True Value	Full model			With $M = ME = 0$			Ignored $M&ME$		
		Bias	MSE	Power	Bias	MSE	Power	Bias	MSE	Power
σ_A^2	80	2.17	3039.06	0.97	2.17	3037.99	0.97	11.71	3705.91	0.98
σ_D^2	50	0.25	237.62	1.00	0.25	237.61	1.00	0.19	237.56	1.00
σ_M^2	40 or 0	-1.58	604.21	0.97	0.01	0.04	0.01	-	-	-
σ_{AE}^2	40	1.01	282.71	1.00	1.00	282.60	1.00	5.87	371.57	1.00
σ_{DE}^2	25	-0.11	25.79	1.00	-0.12	25.79	1.00	2.88	35.62	1.00
σ_{ME}^2	20 or 0	-0.13	63.12	1.00	-0.03	0.13	0.01	-	-	-
σ_e^2	10	0.00	0.52	1.00	0.01	0.52	1.00	24.42	709.09	1.00

Simulation results for prediction of genetic effects concerning maternal model are listed in Table 2. Mean and variance of predicted effects are almost the same whether or not maternal effects and maternal by environment interaction effects exit. However when both effects are ignored, mean and variance of predicted effects are significantly different. Additive, dominance, additive by environment and dominance variance components increase significantly.

Table 2 Mean and Variance of Predicted Effects According to the Consideration of Maternal Effects and Maternal by Environment Effects

Predictor	Variance	Full model		With $M = ME = 0$		Ignored $M&ME$	
		Mean	Variance	Mean	Variance	Mean	Variance
\hat{e}_A	80	0.00164	82.41	0.00164	82.41	-0.00113	91.90
\hat{e}_D	50	0.00021	50.99	0.00021	50.99	-0.00015	55.83
\hat{e}_M	40 or 0	0.00078	38.79	0.00023	0.25	-	-
\hat{e}_{AE}	40	0.00055	41.50	0.00055	41.46	-0.00037	46.13
\hat{e}_{DE}	25	0.00007	26.98	6.88E - 05	26.97	-4.41E - 05	31.93
\hat{e}_{ME}	20 or 0	0.00026	20.87	5.07E - 05	0.40	-	-

Simulation results of bias, MSE and power value for AD models are shown in Table 3. In the full model with significant genotype by environment interaction effects, bias of the respective σ_{AE}^2 and σ_{DE}^2 can be around 4.7% and 0.2% of the parameter values. The significance of non-zero $\sigma_{AE}^2, \sigma_{DE}^2$ can be detected with probability of over 99%. For diallel crosses with F_1 's and reciprocal F_1 's from eight inbred lines within three blocks and three environments, robustness of estimation is tested under the conditions of no specific variation. When $\sigma_{AE}^2 = \sigma_{DE}^2 = 0, \sigma_A^2, \sigma_D^2$ and σ_e^2 can be estimated with similar bias, MSE and power value. However with no genotype by environment interaction effects, the bias, MSE and power value of both σ_{AE}^2 and σ_{DE}^2 become

smaller. Bias of σ_{AE}^2 decreases from 1.88 to 0.01 and that of σ_{DE}^2 from 0.06 to 0.02. When genotype by environment interaction effects are ignored, the main effects and residual effects tend to be overestimated by eight-parent diallel crosses with F_1 's and reciprocal F_1 's. Additive and dominance variance components are respectively overestimated with values of bias around 14% and 13% of the corresponding parameter values. Residual variance component is particularly overestimated with the highest value of bias around 745.5% of the parameter value.

Table 3 Bias, MSE and Power According to Genotype by Environment Interaction Effects

Parameter	True Value	Full model			With $GE = 0$			With GE		
		Bias	MSE	Power	Bias	MSE	Power	Bias	MSE	Power
σ_A^2	80	0.61	2765.91	0.97	-0.160	1988.74	1.00	11.08	2865.52	0.97
σ_D^2	50	0.87	253.91	1.00	0.23	183.01	1.00	6.80	295.73	1.00
σ_{AE}^2	40 or 0	1.88	264.32	1.00	-0.01	0.01	0.03	—	—	—
σ_{DE}^2	25 or 0	0.06	29.42	1.00	-0.02	0.21	0.01	—	—	—
σ_e^2	10	-0.02	0.41	1.00	-0.02	0.41	1.00	74.55	6047.10	1.00

Simulations are also conducted for the predicted $\hat{e}_{u(1)}$ (Table 4). Mean and variance of predicted effects are used to compare full model including genotype by environment interaction effects and full model ignoring interaction effects. AUP method gives extremely low biases for mean of predicted genetic effects with absolute value between $4.35 \cdot 10^{-6} \sim 29.10 \cdot 10^{-4}$. With ignored genotype by environment interaction effects, additive and dominance variances of predictor increase with respective values of bias around 17% and 19% of the corresponding parameter true values.

Table 4 Distance of Genetic Predictor According to the Consideration of Genotype by Environment Interaction

Predictor	Variance	Full model		With $GE = 0$		Ignored GE	
		Mean	Variance	Mean	Variance	Mean	Variance
\hat{e}_A	80	-0.00112	80.86	-0.00127	78.50	-0.00294	94.63
\hat{e}_D	50	-0.00015	51.61	-0.00016	50.99	-0.00034	59.40
\hat{e}_{AE}	40 or 0	-0.00037	42.14	-1.58E-05	0.09	—	—
\hat{e}_{DE}	25 or 0	-5.25E-05	27.14	-4.35E-06	0.44	—	—

3 Discussion

ADM model with maternal and maternal by environment interaction effects gives unbiased estimates with biases less than 5% of the parameter values. No matter whether or not the two effects exist, additive, dominance and residual variance components can be obtained with small biases and high efficiency. If maternal and maternal by environment interaction effects are negligible or null, the values of σ_M^2 and σ_{ME}^2 should be around zero or equal to zero. Under this condition, maternal effects and environment effects are not interdependent. Therefore the full

ADM model can be reduced to *AD* model.

According to the simulation results (Table 1), maternal and maternal by environment effects are ignored in the full model but effectively they exist. In this situation, the reduced model can not be an appropriate model because interaction effects are not included in it. Therefore, induced by some genes sensible to the environment, these effects are added to the original effects and in large part to the residual effects. Consequently the different variance components of the specific effects are overestimated. When σ_{AE}^2 and σ_{DE}^2 are ignored in the additive-dominance model, their effects are distributed into the main genetic effects and mostly into the residual effects. The effects of ignoring σ_M^2 and σ_{ME}^2 are added to additive, additive by environment, dominance by environment and particularly to residual effects. For diallel crosses with F_1 's and reciprocal F_1 's from the eight inbred lines, overestimation of dominance variance component cannot be detected in the maternal model. Maternal and maternal by environment effects are important in plant and animal breeding. More important are the ignored genetic effects; higher can be the overestimation of variance components. The overestimation is highest when particularly total genotype by environment effects is ignored in the additive-dominance model (Table 3). However the detection and evaluation of genotype by environment effects need more control on the experimental design, more entries and replications than those of main effects. When both additive-dominance and maternal model are used with eight-parent modified diallel crosses, the adjusted unbiased prediction with prior value 1 for all variances gives small biases of mean prediction. Mean of predictor is almost near the zero value. The ignoring genotype by environment interaction effects gives high variances of predictor. Before the start of an experiment, the breeder must necessarily have previous information about genotype by environment interaction effects. By default, the appropriate model must be the full *AD* genetic model or *ADM* model. The reduced model *AD* and *ADM* models in single environment can be used only when genotype by environment interaction effects is known as non-significant for the traits studied.

References

- [1] Cockerham C C, Weir B S. Quadratic analysis of reciprocal crosses[J]. *Biometrics*, 1977, **33**:187-203.
- [2] Cockerham C C. Random and fixed effects in plant genetics[J]. *Theor Appl Genet*, 1980, **56**:119-131.
- [3] Efron B. *The jackknife, the bootstrap and other resampling plans*[M]. Society for Industrial and Applied Mathematics, Philadelphia, Pa, 1982.
- [4] Eisen E J. Sex-lined and maternal effects in the diallel cross[J]. *Aust J Biol Sci*, 1966, **19**:1061-1071.
- [5] Falconer D S. *Introduction to Quantitative Genetics*[M]. Ed 4. New York: Longman Press, 1996.
- [6] Graybill F A, Wortham A W. A note on uniformly best unbiased estimators for variance components[J]. *J Am Stat Assoc*, 1956, **51**:266-268.
- [7] Hartley H D, Rao J N K. Maximum-likelihood estimation for the mixed analysis of variance model[J]. *Biometrika*, 1967.
- [8] Henderson C R. *Using estimates in predictions of breeding values under a selection model*[M]. In: Van Vleck L D, Searle S R (eds) *Variance components and animal breeding*. Cornell University, Ithaca, N Y, 1979. 217-227.
- [9] Keele J W, Harvey W R. Estimation of components of variance and covariance by symmetric differences squared and minimum norm quadratic unbiased estimation: a comparison[J]. *J Anim Sci*, 1989, **67**:348-356.

- [10] Kinderman A J, Monathan J F. Computer generation of random variables using the ratio of normal deviates[J]. *Assoc Comput Math Trans Math Soft*, 1977, 3:257-260.
- [11] Miller R G. The jackknife-a review[J]. *Biometrika*, 1974, 61:1-15.
- [12] Paterson A H, Thompson R. Recovery of inter-block information when block sizes are unequal[J]. *Biometrika*, 1971, 58:545-554.
- [13] Rao C R. Estimation of variance and covariance components MINQUE theory[J]. *Journal of Multivariate Analysis*, 1971a, 1:257-275.
- [14] Rao C R. Minimum variance equadratic unbiased estimation of variance components[J]. *Journal of Multivariate Analysis*, 1971b, 1:445-456.
- [15] Rao C R. Estimation of variance and covariance components in linear models[J]. *Journal of Multivariate Analysis*, 1972, 67:112-115.
- [16] Searle S R. *Linear models*[M]. Wiley:New York, 1971.
- [17] Searle S R, Casella G, McCulloch C E. *Variance Components*[M]. New York: Wiley, 1992.
- [18] William H P, Saul A, William T et al. *Numerical recipes in C: the art of scientific computing*[M]. Press Syndicate of the University of Cambridge, USA, 1992.
- [19] Weir B S. *Genetic data analysis II*[M]. Sinauer Associates, Inc. Publishers Sunderland, 1996.
- [20] Zhu J. Mixed model approaches for estimating genetic variances and covariances[J]. *Journal of Biomathematics*, 1992, 7(1):1-11.
- [21] Zhu J. General genetic models and new analysis methods for quantitative traits[J]. *Journal of Zhejiang Agricultural University*, 1994, 20(6):551-559.
- [22] Zhu J. *Analysis methods for genetic models (Chinese)*[M]. Beijing: Agricultural Publishing House of China, 1997, 295.
- [23] Zhu J, Weir B S. Analysis of cytoplasmic and maternal effects: I. genetic model for diploid plant seeds and animals[J]. *Theor Appl Genet*, 1994a, 89:153-159.
- [24] Zhu J, Weir B S. Analysis of cytoplasmic and maternal effects: II. genetic model for triploid endosperm[J]. *Theor Appl Genet*, 1994b, 89:160-166.
- [25] Zhu J, Weir B S. Diallel analysis for sex-linked and maternal effects[J]. *Theor Appl Genet*, 1996, 92(1):1-9.
- [26] Zhu J, Weir B S. Mixed model approaches for genetic analysis of quantitative traits[C]. In advanced Topics in Biomathematics: Proceedings of International Conference on Mathematical Biology, 321-330. Edited by Chen L S, Ruan S G, Zhu J. World Scientific Publishing Co., Singapore, 1998.

加性 - 显性 - 母体效应及 GE 互作效应遗传模型的模拟比较

阿布德拉芒 朱 军

(浙江大学 生物信息学研究所, 浙江 杭州 310029)

摘 要: 运用蒙特卡罗方法比较了 8 个亲本正反交的加性 - 显性 - 母体效应的全模型及缩减模型。当 σ_M^2 和 σ_{ME}^2 存在时, 检测各项遗传方差分量的功效高达 97% 以上。如果存在 σ_M^2 和 σ_{ME}^2 而缩减模型未包括这两项效应时, 除显性效应以外的各项方差分量都被高估。对于加性 - 显性模型, 如果忽略了基因型与环境互作, σ_i^2 和 σ_A^2 将被高估。当母体效应和基因型与环境互作被忽略时, 将显著地增加遗传效应预测值的方差。

关键词: 遗传方差分量; 遗传效应预测; 混合模型方法; 蒙特卡罗模拟