

AN APPLICATION OF ANALYSIS OF COVARIANCE TECHNIQUE IN MATING DESIGN

BY

R C. HASIJA AND K.C. GEORGE*

Hissar Agricultural University, Hissar

(Received : June, 1979)

SUMMARY

In the present investigation, the analysis covariance technique was applied to the model I of method IV of Griffing (1956), by taking a concomitant variable, in order to increase the efficiency of the estimate of the g.c.a. and s.c.a. effects. The analysis of covariance of the above model was made. The adjusted estimates of the g.c.a. and s.c.a. effects were derived. The variances, and estimates of the variances of the various comparisons of the g.c.a. and s.c.a. effects were derived. Average variances of the different g.c.a. and s.c.a. comparisons were also derived. Finally an illustrative example was also worked out in order to show the usefulness of this technique.

1. INTRODUCTION

Generally the analysis of the mating designs, especially that of the diallel mating designs, is being carried out without exploiting the relationship of the main characteristic under study with that of the other characters which are independent to the varieties under trial and at the same time contributing considerably to the main character. It is of interest to note that these characters are influencing the main character to a considerable extent. Hence, it is essential to eliminate these types of influences before analysing the main character under study to make a proper test of significance as well as to get a reasonable estimate of the general and specific combining ability.

Mating designs were first suggested by Schmidt (1919). It largely determines the kinds of genetic information. From a practical point of view, the mating designs provide a very simple and convenient method of generating a number of crosses in one or two generations and it is mainly because of this mating designs are now the important tools to both plant and animal breeders.

*Present address : Kerala Agriculture University, Trichur.

The diallel cross which is composed of all single crosses among a group of inbred lines, is now a common plan of investigation in both plant and animal breeding. Modern use of diallel crossing started with the development of the concepts of general and specific combining ability by Sprague and Tatum (1942). The diallel cross is used to estimate the genetic components of variation among the yields of the crosses and to estimate the actual yielding capacities of crosses.

The first attempt to analysis of diallel cross was made by Yates (1947). Later the analysis of diallel crosses was presented by Jinks and Hayman (1953). Hayman (1954a, 1954b) considered on arbitrary degree of dominance for the cause where the parent in the diallel table were completely homozygous. Only two alleles per locus were considered and assumption of no epistasis was imposed. Dickinson and Jinks (1956) extended the methods of Hayman and Jinks to of include the inbreeding the parental material in the diallel table.

Kempthorne (1956) removed the assumption of no epistasis and of only two alleles per locus as given by Hayman (1954b) and presented the analysis of diallel crosses for arbitrary epistacy and thus making the approach more realistic and valid.

Griffing (1956a, 1956b) considered diallel crosses of homozygous individuals obtained from a random mating population and presented the genetic parameters in terms of additive, dominance and epistatic effects for an arbitrary number of loci with arbitrary alleles. The analysis given by Griffing (1956) in diallel crosses is based on the 'mean' of a selected number of plants.

Covariance technique has been recognized as an important technique to reduce the experimental error by taking advantage of the association which is believed to exist between the experimental variate 'y' and a concomitant variate 'x'. But this concept has not been utilized in mating designs so far.

Present investigation is an attempt to analyse the diallel cross data based on Griffing (1956) method IV of model I, by adjusting with a concomitant variable in order to increase the efficiency of *g.c.a.* and *s.c.a.* effects. Unlike regular Griffings model, here we are considering the data on individual observation wise.

2. THEORETICAL PROCEDURES

In the present investigation analysis of covariance model corresponding to the method IV of model I, Griffing (1956) was taken as

$$Y_{ijk} = \mu + g_i + g_j + s_{ij} + b(x_{ijk} - \bar{x}) + e_{ijk}$$

$i, j = 1, 2, \dots, p$
 $k = 1, 2, \dots, n$

where,

Y_{ijk} = is the response of (i, j) th cross in k th replication

μ = overall mean or general mean

g_i = g.c.a. effect for i th parent

g_j = g.c.a. effect for j th parent

s_{ij} = s.c.a. effect for the cross between i th and j th parents such that $s_{ij} = s_{ji}$

x_{ijk} = is the auxiliary variable or concomitant variable which has a linear regression with regression coefficient ' b ' on Y_{ijk} in the error line

e_{ijk} = error associated with (ijk) th individual is

$$NI\left(0, \sigma_e^2\right).$$

The above model was taken because of the following reasons :

(1) It represents the most commonly used diallel crossing system.

(2) When the necessary assumptions concerning the sampling nature of the set of inbreds can be validly made, it is possible to give an exact genetic interpretation to the general and specific combining ability variations.

(3) Reciprocal genotypic effects are generally non-existent in plant data and therefore, it is not necessary to introduce the additional complication of the reciprocal F^2 s.

The parameters of the above model have been estimated by the method of fitting constants. The estimates of different parameters are given below :

$$(i) \hat{\mu} = \frac{2y_{...}}{np(p-1)}$$

$$(ii) \hat{g}_i = \frac{y_{i..}}{n(p-2)} - \frac{2y_{...}}{np(p-2)} - \hat{b} \left[\frac{x_{i..}}{n(p-2)} - \frac{2x_{...}}{np(p-2)} \right]$$

$$(iii) \hat{g}_j = \frac{y_{.j.}}{n(p-2)} - \frac{2y_{...}}{np(p-2)} - \hat{b} \left[\frac{x_{.j.}}{n(p-2)} - \frac{2x_{...}}{np(p-2)} \right]$$

$$(iv) \hat{s}_{ij} = \frac{y_{ij.}}{n} - \frac{y_{i..} + y_{.j.}}{n(p-2)} + \frac{2y_{...}}{n(p-1)(p-2)} - \hat{b} \left[\frac{x_{ij.}}{n} - \frac{x_{i..} + x_{.j.}}{n(p-2)} + \frac{2x_{...}}{n(p-1)(p-2)} \right]$$

where
$$\hat{b} = \frac{\sum_i \sum_j \sum_k y_{ijk} x_{ijk} - \frac{1}{n} \sum_i \sum_j y_{ij} x_{ij}}{\sum_i \sum_j \sum_k x_{ijk}^2 - \sum_i \sum_j \frac{x_{ij}^2}{n}}$$

Sum of squares and mean squares have been calculated and are given in the analysis of variance table below :

TABLE I
Analysis of variance after adjusting with the concomitant variable

S.V.	d.f.	S.S.	M.S.
Due to g	(p-1)	S'_g	M'_g
Due to s	$\frac{p(p-3)}{2}$	S'_s	M'_s
Error	$\frac{p(p-1)(n-1)}{2} - 1$	S'_e	M'_e
Total	$\frac{np(p-1)}{2} - 2$	S'_t	

where

$$S'_g = \sum_i \frac{y_{i..}^2}{n(p-2)} - \frac{4y^2 \dots}{np(p-2)} - \hat{b} \left[- \sum_i \frac{x_{i..} y_{i..}}{n(p-2)} - \frac{4x \dots y \dots}{np(p-2)} \right]$$

$$S'_s = \frac{1}{n} \sum_i \sum_j y_{ij}^2 - \frac{1}{n(p-2)} \sum_i y_i^2 \dots + \frac{2y^2 \dots}{n(p-1)(p-2)} - \hat{b} \left[\frac{1}{n} \sum_i \sum_j y_{ij} x_{ij} - \frac{1}{n(p-2)} \sum_i x_i \dots y_i \dots + \frac{2x \dots y \dots}{n(p-1)(p-2)} \right]$$

$$S'_e = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{1}{n} \sum_i \sum_j y_{ij}^2 \dots - \hat{b} \left[\sum_i \sum_j \sum_k x_{ijk} y_{ijk} - \frac{1}{n} \sum_i \sum_j x_{ij} y_{ij} \dots \right]$$

$$S'_t = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{2y^2 \dots}{np(p-1)} - \hat{b} \left[\sum_i \sum_j \sum_k y_{ijk} x_{ijk} - \frac{2x \dots y \dots}{np(p-1)} \right]$$

and

$$M'_g = S'_g / (p-1)$$

$$M'_s = S'_s / \frac{p(p-3)}{2}$$

$$M'_e = S'_e / \left[\frac{p(p-1)(n-1)}{2} - 1 \right]$$

The variances of the estimates of general combining ability and specific combining ability can be derived and are as given below :

$$(i) V(\hat{g}_i) = \frac{(p-1)}{np(p-2)} \sigma_e'^2 + \frac{1}{E_{xx}} \left\{ \frac{x_{i...}}{n(p-2)} - \frac{2x_{..}}{np(p-2)} \right\}^2 \sigma_e'^2$$

$$(ii) V(\hat{s}_{ij}) = \frac{(p-3)}{n(p-1)} \sigma_e'^2 + \frac{1}{E_{xx}} \left\{ \frac{x_{ij.}}{n} - \frac{x_{i..} + x_{j..}}{n(p-2)} + \frac{2x_{...}}{n(p-1)(p-2)} \right\}^2 \sigma_e'^2$$

$i \neq j$

The variances of the difference comparisons of the estimates of general and specific combining ability effects are derived and are as given below :

$$(i) V(\hat{g}_i - \hat{g}_j) = \frac{2}{n(p-2)} \sigma_e'^2 + \frac{1}{E_{xx}} \left\{ \frac{(x_{i..} - x_{j..})}{n(p-2)} \right\}^2 \sigma_e'^2$$

$i \neq j$

$$(ii) V(\hat{s}_{ij} - \hat{s}_{kl}) = \frac{2(p-3)}{n(p-2)} \sigma_e'^2 + \frac{1}{E_{xx}} \left\{ \frac{x_{ij.} - x_{kl.}}{n} - \frac{x_{ij.} - x_{kl.}}{n(p-2)} \right\}^2 \sigma_e'^2$$

$i \neq j, k; j \neq k$

$$(iii) V(\hat{s}_{ij} - \hat{s}_{kl}) = \frac{2(p-4)}{n(p-2)} \sigma_e'^2 + \frac{1}{E_{xx}} \left\{ \frac{x_{ij.} - x_{kl.}}{n} - \frac{x_{i..} + x_{j..} - x_{k..} - x_{l..}}{n(p-2)} \right\}^2 \sigma_e'^2$$

$i \neq j, k, l; j \neq k, l; k \neq l$

It is clear from the above that the variances of the estimates of different effects depend upon the concomitant variable and it will not be a fixed value as in the case of analysis without adjustment with the concomitant variable. Hence for comparison purposes,

average variance of the estimate of the various comparison of the effects calculated as follows :

$$\overline{V}(\hat{g}_i - \hat{g}_j) = \left[\frac{2}{n(p-2)} + \frac{2}{E_{xx}p(p-1)} \sum_{i \neq j}^p \sum_j^p \left\{ \frac{x_{i..} - x_{j..}}{n(p-2)} \right\}^2 \right] \sigma_e'^2$$

$$V(\hat{s}_{ij} - \hat{s}_{ik}) = \left[\frac{2(p-3)}{n(p-2)} + \frac{6}{p(p-1)(p-2)E_{xx}} \sum_{i=j}^p \sum_{j=k}^p \sum_k^p \left\{ \frac{x_{jj..} - x_{ik..}}{n} - \frac{x_{ij..} - x_{ik..}}{n(p-2)} \right\}^2 \right] \sigma_e'^2 \quad i \neq j \neq k$$

$$\overline{V}(\hat{s}_{ij} - \hat{s}_{kl}) = \left[\frac{2(p-4)}{n(p-2)} + \frac{8}{p(p-1)(p-2)(p-3)E_{xx}} \sum_{i=j}^p \sum_{j=k}^p \sum_{k=l}^p \sum_l^p \left\{ \frac{x_{ij..} - x_{kl..}}{n} - \frac{x_{i..} + x_{j..} - x_{k..} - x_{l..}}{n(p-2)} \right\}^2 \right] \sigma_e'^2$$

$i=j, k=l, 1; j=k, 1; k=l$

Where $\sigma_e'^2 = \frac{2}{p(p-1)(n-1)} \left[\sum_{i,j,k}^p \sum_j^p \sum_k^p y_{ijk}^2 - \sum_{i,j}^p \frac{y_{ij}^2}{n} - b \left\{ \sum_{i,j,k}^p x_{ijk} y_{ijk} - \sum_{i,j}^p \sum_{i < j} \frac{x_{ij} \cdot y_{ij}}{n} \right\} \right] \quad (a)$

$$E_{xx} = \sum_{i,j,k}^p \sum_j^p \sum_k^p x_{ijk} - \sum_{i,j}^p \sum_{i < j} \frac{x_{ij}^2}{n}$$

$$b = \frac{\sum_{i,j,k}^p \sum_j^p \sum_k^p y_{ijk} x_{ijk} - \frac{1}{n} \sum_{i,j}^p \sum_{i < j} y_{ij} \cdot x_{ij}}{\sum_{i,j,k}^p \sum_j^p \sum_k^p x_{ijk}^2 - \sum_{i,j}^p \sum_{i < j} \frac{x_{ij}^2}{n}}$$

2.1 Comparison of efficiencies :

(i) The efficiency of analysis of covariance over the analysis of variance is given by

$$E = \frac{\sigma_e^2}{\sigma_e'^2}$$

Where $\sigma_e'^2 = \sum_{i,j,k}^p \sum_j^p \sum_k^p y_{ijk}^2 - \sum_{i,j}^p \sum_{i < j} \frac{y_{ij}^2}{n}$

and σ_e^2 is given by equation (a) above

(ii) Efficiency of analysis of covariance over the analysis of variance in comparing the g.c.a effects is given by

$$E_1 = \frac{\frac{2}{n(p-2)}\sigma_e^2}{\left[\frac{2}{n(p-2)} + \frac{2}{E_{xx}p(p-1)} \sum_{i \neq j} \sum_i \left\{ \frac{x_{i..} - x_{i.}}{n(p-2)} \right\}^2 \right] \sigma_e^2} \quad i \neq j$$

(iii) Efficiency of analysis of covariance over the analysis of variance in comparing the s.c.a effects (when one parent is common) is given below :

$$E_2 = \frac{\frac{2(p-3)}{n(p-2)}\sigma_e^2 \quad i \neq j, k; j \neq k}{\left[\frac{2(p-2)}{n(p-2)} + \frac{6}{p(p-1)(p-2)} \sum_i \sum_j \sum_k \left\{ \frac{x_{ij.} - x_{ik.}}{n} - \frac{x_{ij.} - x_{ik.}}{n(p-2)} \right\}^2 \right] \sigma_e^2}$$

(iv) Efficiency of analysis of covariance over the analysis of variance in comparing the s.c.a effects (when no parent is common) is given by :

$$E_3 = \frac{\frac{2(p-4)}{n(p-2)}\sigma_e^2}{\left[\frac{2(p-4)}{n(p-2)} + \frac{8}{p(p-1)(p-2)} \sum_i \sum_j \sum_k \sum_l \left\{ \frac{x_{ij.} - x_{kl.}}{n} - \frac{x_{i..} + x_{j..} - x_{k..} - x_{l..}}{n(p-2)} \right\}^2 \right] \sigma_e^2}$$

$i \neq j, k, l; j \neq k, l; k \neq l$

3. ILLUSTRATIVE EXAMPLE

To explain the theoretical procedure for the analysis of covariance of diallel cross data by taking one concomitant variable derived above, a data on diallel cross experiment has been obtained from the Department of Forage, conducted on eight varieties of oats (HFO-114, Weston-11, HFO-54, HFO-55, HFO-43, HFO-65, HFO-163 and HFO-184) in the year 1975 at HAU Farm Hissar. The data consists of observations on green fodder yield (γ), the main character and the number of tillers per plant \times as an auxiliary character.

The data has been analysed in the Randomized Block Design and it has been tested and it has been found that the varieties do not differ significantly with regard to the auxiliary character (number of tillers per plant). Further, it has been tested for the regression and

it was found to be highly significant. It means that the number of tillers per plant influence the green fodder yield. Hence, in order to estimate correctly the difference among the different varieties, the effect due to this concomitant variable needs to be removed. Hence, the analysis of variance of the diallel cross for yield with and without adjusting the effect of this concomitant variable (number of tillers per plant) were worked out as shown below :

TABLE 2
Analysis of variance table (without adjusting the effect of the concomitant variable)

Source	d.f.	SS(y)	M.S.(y)	F _{cal}
Due to <i>g</i>	7	120218.77	17174.11	2.99*
Due to <i>s</i>	20	157129.10	7856.46	1.37 (N.S.)
Error	28	161022.92	5750.82	

* Significant at 5% level.

TABLE 3
Analysis of variance table (after adjusting with the effect of the concomitant variable)

S.V.	d.f.	S.S.	M.S.	F _{cal}
Due to <i>g</i>	7	20321.49	930.0	1.03 N.S.
Due to <i>s</i>	20	85739.71	4286.99	1.52 N.S.
Error	27	75903.41	2811.24	

From table 2 it can be concluded that the differences due to *g.c.a.* effect was significant but it comes out to be non-significant in table 3, implying that the significance of the *g.c.a.* effect in main character in table 2 was due to the influence of the concomitant variable. Moreover, it can also be observed that the error mean square is considerably reduced due to the application of covariance technique.

The different results of the analysis after adjustment by the concomitant variable were as given in tables 4, 5, 6, 7.

TABLE 4

Mean of the main character (green fodder yield)

<i>Parents</i>	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	<i>Total</i>
P_1		315.00	159.90	136.25	105.00	211.25	437.50	307.05	1671.95
P_2			185.35	107.50	155.00	226.60	225.10	242.50	1457.05
P_3				160.75	208.30	239.35	261.25	158.30	1373.20
P_4					125.00	221.20	227.50	212.05	1190.25
P_5						291.60	220.00	228.30	1330.20
P_6							163.75	172.00	1525.75
P_7								243.30	1707.10
P_8									1563.50
Total									5946.65

TABLE 6

Adjusted general combining ability (diagonal elements) and specific combining ability (off diagonal elements) effects.

<i>Parents</i>	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
P_1	-1.10	113.09	193.54	150.20	171.90	201.01	278.83	218.18
P_2		-4.51	216.55	177.91	195.31	214.76	192.32	214.91
P_3			-14.57	178.53	213.34	198.99	164.08	139.81
P_4				-17.83	156.64	178.31	233.93	231.53
P_5					1.47	262.20	174.31	153.26
P_6						-7.92	192.90	178.87
P_7							39.05	190.47
P_8								12.65

Average variances of the different comparisons of the g.c.a. effects and s.c.a. effects were also worked out by using the formulae given already above as follows:

$$(i) \frac{\hat{\Delta}}{\hat{V}}(adj \hat{g}_i - adj \hat{g}_j) = 493.0 \quad i \neq j$$

$$(ii) \frac{\hat{\Delta}}{\hat{V}}(adj \hat{s}_{ij} - adj \hat{s}_{ik}) = 2520.43 \quad i \neq j, k; j \neq k$$

i.e. when one of the parent is common

$$(iii) \frac{\hat{\Delta}}{\hat{V}}(adj \hat{s}_{ij} - adj \hat{s}_{kl}) = 2021.25$$

$i \neq j, k, l; j \neq k, l; k \neq l$
i.e. when no parent is common

The efficiency of analysis of covariance over the analysis of variance of the different comparisons of g.c.a. effects and s.c.a. effects, calculated by using the formulae given above, are given below :

$$\hat{\sigma}_e^2 = 5750.82$$

$$\hat{\sigma}_s^2 = 2811.24$$

$$(i) \hat{E} = \frac{5750.82}{2811.24} = 2.05$$

$$\text{Estimated gain in efficiency} = \hat{E} - 1 = 2.05 - 1.00 = 1.05$$

$$(ii) \hat{V}(g_i - g_j) = 958.47$$

$$\frac{\hat{\Delta}}{\hat{V}}(adj \hat{g}_i - adj \hat{g}_j) = 493.00$$

$$\hat{E}_1 = \frac{958.47}{493.00} = 1.94$$

$$\text{Estimated gain in efficiency while comparing the g.c.a. effects} \\ = \hat{E}_1 - 1 = 1.94 - 1.00 = 0.94.$$

$$(iii) \hat{V}(s_{ij} - s_{ik}) = 4792.35$$

$$\frac{\hat{\Delta}}{\hat{V}}(adj \hat{s}_{ij} - adj \hat{s}_{ik}) = 2520.43$$

$$\hat{E}_2 = \frac{4792.35}{2520.43} = 1.90$$

Estimated gain in efficiency while comparing the g.c.a. effects
 $= \hat{E}_2 - 1 = 1.90 - 1.00 = 0.90$

$$(iv) \quad \hat{V}(s_{ij} - s_{kl}) = 3833.38$$

$$\hat{V}(adj\ s_{ij} - adj\ s_{kl}) = 2021.25$$

$$\hat{E}_3 = \frac{3833.38}{2021.25} = 1.90$$

Estimated gain in efficiency while comparing the s.c.a. effects
 $= \hat{E}_3 - 1 = 1.90 - 1.00 = 0.90$

Hence we can reasonably conclude that the analysis of covariance technique increases the efficiency in all comparisons over the corresponding analysis of variance technique.

ACKNOWLEDGEMENT

The authors are very much grateful to the Professor and Head, Department of Math. & Stat. for providing us the necessary facilities and to the Prof. and Head, Department of Plant Breeding for the data for the illustration.

REFERENCES

- Dickinson, A G. and Jinks, J.L. (1956) : A generalized analysis of diallel crosses. *Genetics* 41, (65-77).
- Griffing, B. (1956a) : Concept of general and specific combining ability in relation to diallel crossing systems. *Austr. J. Biol. Sci.* 9. (463-493).
- Griffing, B. (1956b) : A generalized treatment of the use of diallel crosses in quantitative inheritance. *Heredity* 10 pp. 31-50.
- Hayman, B.I. (1954a) : The analysis of variance of diallel crosses. *Biometrics* 10, pp. 235-244.
- Hayman, B.I. (1954b) : The theory and analysis of diallel crosses. *Genetics* 39, pp. 789-809.
- Jinks, J.L. and Hayman, B.I. (1953) : The analysis of diallel crosses. *Maize Genetic News Letter* 27, pp. 48-54.
- Kempthorne, O. (1956) : The theory of diallel cross. *Genetics* 41, pp. 451-459.
- Schmidt, J. (1919) : La valeur de l'individua titre de generateur appreciee suivant la methode du croisement diallele, *Compt : Rend. Lab. Carlsberg* 14. No. 633.
- Sprague, G.F. and Loyd, A. Tatum (1942) : General vs specific combining ability in single crosses of corn. 34, pp. 923-932.
- Yates, F. (1947) : The analysis of the data from all possible reciprocal crosses between a set of parental lines. *Heredity* 1, pp. 387-407.