Journal of Agricultural Biotechnology and Sustainable Development Vol. 2(4), pp. 56-60, April 2010
Available online http://www.academicjournals.org/JABSD

Full Length Research Paper

Variance balanced designs for complete diallel cross

M. K. Sharma* and Sileshi Fanta

Addis Ababa University, Addis Ababa, Ethiopia.

Accepted 11 January, 2010

Optimal block designs for complete diallel cross system IV was investigated by several authors in presence and absence of specific combining ability parameters in the model. But there are drawbacks in their designs. Firstly, their designs consume more experimental units when these are used for experiments to estimate general and specific combining ability effects and secondly, these designs are not available for every value of p of parental lines. In the present paper we are proposing a simple method of construction of balanced and orthogonal block designs for every value for p of parental lines for complete diallel cross plans connected for cross effects through cycle permutation of initial block with minimum possible number of experimental units and at the same time these designs allow same precision for estimation of contrasts of general and specific combining ability of participating lines. The analysis is presented for data obtainable from proposed designs. The statistical analysis includes the analysis of variance and the estimation of general and specific combining abilities.

Key words: Incomplete block design, complete diallel cross, general combining ability, specific combining ability, mating-environment design, variance balanced.

INTRODUCTION

Various form of diallel crosses as mating designs are used in plant and animal breeding to study the genetic properties and potential of inbred lines or individuals. Let p denote the number of lines and let a cross between lines i and j be denoted by (ixj) = (jxi) i ≠ j = 1, 2,…, p and let v be number of crosses. Among the four types of diallel discussed by Griffing (1956), method 4 is the most commonly used diallel in plant breeding. This type of diallel crossing includes the genotypes of one set of F_1's, but neither the parents nor reciprocals i.e. v = p (p-1)/2. We shall refer to it as a complete diallel cross (CDC). The common practice with CDC is to evaluate the crosses in a completely randomized design or randomized complete block design as the environment designs. Due to limitation of homogeneous experimental units in a block to accommodate all the chosen crosses, the estimates of genetic parameters would not be precise enough if a complete block design was adopted for the large number of crosses. It is for this reason that the use of incomplete block designs as environment designs has been advocated by Bratten (1965), Aggarwal (1974), Ceranka and Meza (1988), Agarwal and Das (1990), Divecha and Ghosh (1994), Singh and Hinkelmann (1988) and Sharma (1996).

The problem of generating optimal mating designs for experiments with diallel crosses has been investigated by Gupta and Kageyama (1994), Dey and Midha (1996), Mukerjee (1997), Das et al. (1998), Parsad et al. (1999), Sharma (2004) and Sharma and Fanta (2009) under the assumption that the model does not include parameters representing specific combining abilities. Chai and Mukerjee obtained optimal designs for diallel crosses with specific combining abilities using triangular PBIB designs. These designs require experimental units in range of 30 - 405 to conduct experiment for complete diallel cross system IV for 5 - 10 parental lines and more over, their designs are not available for every value of p parental lines. In diallel cross experiments, it is very difficult to get large homogeneous experimental. So the experimenter can not use many of their designs for the diallel cross experiments. In this situation, the experimenter can use proposed variance balanced block designs for their experiments

In this paper we give a simple method of construction of variance balanced incomplete block designs for diallel crosses which are connected for cross effects for every value of p parental lines to fill this gap. These designs are found to be MS-optimal according to the criterion by Shah (1960) and Eccleston and Hedayat (1974). We also present their analysis. The statistical analysis includes

*Corresponding author. E-mail: mk_subash@yahoo.co.in.

the analysis of variance and the estimation of general and specific combining abilities.

METHODS OF CONSTRUCTION

The method of construction of incomplete block designs connected for cross effect for all values of p of parental lines is simply stated as: for p lines under evaluation numbered randomly from 1 to p, the first line is crossed with all (p-1) lines. These (p-1) crosses are the initial elements for the first block and other (p-1) blocks can be further obtained cyclically by developing (p-1) crosses of the first block , where symbols in each block are reduced mod (p).Thus we obtain incomplete block design d for CDC system IV with parameters v = p (p-1)/2, b = p , r = 2, k = p-1. Henceforth, d (v, b, k) will denote the class of all block design with v treatments, b blocks and block size k. This design is connected for cross effects.

Example 1

For p = 5, the first block will contain following crosses [0 × 1, 0 × 2, 0 × 3, 0 × 4] to be developed mod 5 to obtain following design with parameters v = 10, b = 5, r = 2 and k = 4.

B1	B2	B3	B4	B5
0×1	1×2	2×3	3×4	4×0
0×2	1×3	2×4	3×0	4×1
0×3	1×4	2×0	3×1	4×2
0×4	1×0	2×1	3×2	4×3

We note the following remarkable features of these two series of designs.

1. If we consider (p-1) row of the above mating-environment design as blocks, we get incomplete block design for CDC experiment with parameters v = p (p-1)/2, b = p-1, k = p, and r = 2. The efficiency of this design is equal to one in comparison to randomized block design.
2. These are the series of V-B designs with constant replication number. This constant, again, is the minimum possible value for a connected equireplicate mating design for CDC experiment. Thus, these designs are extremely cost-effective as they allow the maximum number cross to be tested on a given set up.
3. The members of the d series are quite plentiful, as for every positive integer, there is a design.
4. Due to feature 2, these designs can be used for screening of potential genotypes in preliminary trials.

ANALYSES

For the analysis of data obtained from design d, we will follow Singh and Hinkelmann (1998) two stage procedures for estimating gca and sca effects. The first stage is to consider estimating cross effects, say, $\tau = (\tau_{01},...,\tau_{p(p-1)/2})'$ by the following model.

$$y = \mu 1 + X\tau + D\beta + e \quad (1)$$

Where y be n × 1 vector of observations, 1 is the n × 1 vector of ones, X is the n × v design matrix for treatments and D is an n × b design matrix for blocks, that is, the $(h,u)^{th}$ ($(h,l)^{th}$) element of X (respectively, of D) is 1 if the h^{th} observation pertains to the u^{th} cross (to l^{th} block), and is zero otherwise (h = 1, …, n; u = 1, …, v ; and 1, … , b), μ is a general mean, τ is a v × 1 vector of treatment parameters, β is a b × 1 vector of block parameters and e is an n × 1 vector of residuals. It is assumed that vector β is fixed and e is normally distributed with E (e) = 0, V (e) = σ^2 I and Cov (β, e) = 0, where I is the identity matrix of conformable order.

Following Tocher (1952), Raghava et al. (1971), the least square method for the analysis of a proposed design leads to the following reduced normal equations for the crosses for model (1).

$$C_d\tau = Q_d \quad (2)$$

Where $C_d = r^{\delta} - N k^{-\delta} N'$ and $Q_{d1} = (Q_{1d1},...., Q_{vd1}) = T - N k^{-\delta} B$.

In the above expressions above, r^{δ} and k^{δ} are diagonal matrices of order v × v and b × b with elements 2 and p-1 in the diagonal, respectively of designs d. N = X′D is the v × b incidence matrix of the designs d, T = X′ Y and B = D′Y are vector of cross totals and block totals of order v × 1 and b × 1 respectively for design d. Hence a solution to (2) is given by

$$\hat{\tau} = C_d^{-} Q_d \quad (3)$$

Where C_d^{-} is the generalized inverse of C_d with property $C C^{-} C = C$. The sum of squares due to crosses is $Q_d' C_d^{-} Q_d$ with d.f. = rank (C_{d1}). Expectation and variance Q_d is given as

$$E(Q_d) = C_d \tau \text{ and } V(Q_d) = \sigma^2 C_d \quad (4)$$

Now we will utilize the above equations to estimate the genetic parameters in proposed design.

Note: According to Chai and Mukerjee (1999), if the C_d matrix of the design d, for some scalars ω_1, ω_2, and ω_3 can be written as

$$C_d = \omega_1 I_v + \omega_2 B + \omega_3 J_{vv} \quad (5)$$

Where for positive integer I_u is the u × u identity matrix, B is a v × v matrix with rows and columns indexed by the pairs (i, j) , for 0 ≤ i ≤ j ≤ p-1 , such that the { (i_1,j_1), (i_2,j_2)}th entry of b equals 1 if the sets (i_1,j_1) and (i_2,j_2) have exactly one element in common and 0 otherwise, J_{vv} is a square matrix of 1's. Then the design d will be balanced and orthogonal. Here ω_1 = r (p-2)/ (p-1), ω_2 = -1/(p-1) and ω3 = 0 and the C_d of design d can be written as (5). Hence design d is balanced and orthogonal.

The second stage is to utilize the fact that the cross effects can be expressed in terms of gca and sca effects. So we can write;

$$\tau_{ij} = g_i + g_j + s_{ij} \quad (6)$$

Where g_i (g_j) is the gca for the i^{th} (j^{th}) parent, s_{ij} ($s_{ij} = s_{ji}$) is the sca for the cross between the i^{th} and the j^{th} parent($i< j = 0,1, \ldots, p-1$). In matrix notation equation (6) can be written as

$$\tau = Z g + s \quad (7)$$

Where $Z = (z_{ui})$ ($u = 1, 2, \ldots, n : i = 0, 1, \ldots, p-1$) is the cross and gca relation matrix. $z_{ui} = 1$ if the u^{th} cross has a parent i, otherwise 0. In design d the cross and gca relation matrix and incidence matrix are same for respective design that is Z = N.

Following the approach used in Kempthorne and Curnow (1961), equation (2) can be written as $C_d \tau = C_d Z g + C_d s$ that is;

$$E(Qd) = C_d Z g + C_d s \quad (8)$$

Since the matrix C is singular, we use the unified of least square due to Rao (1973). So we get

$$\hat{g} = (Z' C_d C_d^{-} C_d Z)^{-} Z' Q_d = (Z' C_d Z)^{-} Z' Q_d \quad (9)$$

Here the matrix $(Z' C_d Z) = \frac{p(p-2)}{(p-1)}[I_p - \frac{1}{p} 1p\, 1'p]$, I_p is the p × p identity matrix, 1p is the p × 1 column vector with all elements unity.

$$\text{So } \hat{g} = \frac{(p-1)}{p(p-2)} Z' C_d \tau \quad (10)$$

Hence $\hat{g} = H_1 \tau$, where $H_1 = Z' C_d$

$$\text{Now Var}(\hat{g}) = H_1 C^{-} H_1 \sigma^2 = \sigma^2 \frac{p(p-2)}{(p-1)} I_p \quad (11)$$

Since the Var (ĝ) is a constant times the identity matrix, therefore the design d is balanced for the general combining abilities. Hence the proposed design d is variance balanced. We thus have the following result.

Theorem 1

Let p be a positive integer greater > 3. Then there exists a variance balanced incomplete block designs for CDC experiment IV with v = p (p-1)/2 crosses as treatments such that each cross appears two times in a design.

Now substituting the estimate of g in equation (7), we obtain the estimate of s for design d, respectively as

$$\hat{s} = (C_d - \frac{(p-1)}{p(p-2)} ZZ')Q_d$$

$$= (C_d^{-} - \frac{(p-1)}{p(p-2)} ZZ')C_d\tau$$

$$= H_2 \tau \quad (12)$$

Where $H_2 = (C_d^{-} - \frac{(p-1)}{p(p-2)} ZZ)C_d$ and is symmetric idempotent.

$$\text{Var}(\hat{s}) = H_2' C^{-} H_2 \sigma^2 \quad (13)$$

Since $H_1 1_v = 0$, $H_2 1_v = 0$, $H_1 H_2' = 0$. rank $(H_1) = p-1$ and rank $(H_2) = v-p$.

It follows that g and s represented by treatment contrasts that carry p-1 and v-p degrees of freedom respectively and that contrasts representing g are orthogonal to those representing s. It means the proposed design d allows for gca and sca effects to be estimated independently.

The sum of squares due to gca and sca for d2 are given by

$$SS(gca) = Q_d' Z (Z' C_d Z)^{-} Z' Q_d \quad (14)$$

$$SS(sca) = Q_d' (C_d^{-} - \frac{(p-1)}{p(p-2)} Z Z') Q_d \quad (15)$$

The ANOVA is then given in Table 1.

Optimality

According to James and Wilkinson (1971), the canonical efficiency factors of the designs d is p/(2(p-1). Since all the efficiency factors of design d are equal. Hence design d is MS- optimal and as well as efficiency balanced.

An example for illustrating the theory given in this paper, let us consider the data from an experiment on blooming days of sunflowers, reported by Ceranka and Mejza (1988). These authors used a balanced incomplete block design with v = 10, b = 15, r = 6, and k = 4 resulting from diallel crossing among p = 5 lines of sunflower and found out the estimates of gca and sca using intra and inter-block analysis of variance. We took the data of two replications randomly for each genotype for design d. The following exhibit the experimental design with data, vectors T, Q and matrices r^{δ}, k^{δ}, Z, C, and $(Z' C_d Z)^{-}$ (Table. 2).

T′ = (20.0, 30.0, 22.0, 24.0, 27.0, 24.1, 28.8, 23.5, 24.1, 27.9)

Q′ = (-4.625, 4.85, -1.9, -1.175, 1.225, -.425, 3.00, -1.55,-2.225, 2.825)

r^{δ} = (2, 2, 2, 2, 2, 2, 2, 2)

k^{δ} = (4, 4, 4, 4, 4)

Table 1. Intra-block analysis of proposed design d_2 as mating and environment design.

Source of variation	Degrees of freedom	Sum of squares
Blocks	p-1	$B' B/p - G^2/2p^2$
Crosses (adjusted for blocks)	rank $(C_d) = p(p-1)/2 - 1$	$Q_d' C_d^- Q_d$
gca	rank $(H_1) = p-1$	$Q_d' Z (Z' C_d Z)^- Z' Q_d$
sca	rank $(H_2) = p(p-3)/2$	$Q_d' H_2 Q_d$
Residual	$(p^2 -3p +4)/2$	by subtraction, R^2_o
Total	n-1	$y' y - G^2/p(p-1)$

Table 2. Experimental design with data in parentheses.

B1	B2	B3	B4	B5
1×2 (10.0)	2×3 (13.00	3×4 (11.5)	4×5 (14.0)	5×1 (12.0)
1×3 (15.0)	2×4 (13.5)	3×5 (12.1)	4×1 (11.0)	5×2 (14.8)
1×4 (11.0)	2×5 (14.0)	3×1 (15.0)	4×2 (10.6)	5×3(12.0)
1×5 (12.0)	2×1 (10.0)	3×2 (14.0)	4×3 (12.0)	5×4(13.9)
Total = 48.0	50.5	52.6	47.6	52.7

Table 3. Intra and inter block analysis of the data on blooming days.

Source	D.F	SS	MSS	F
Blocks	4	5.9170		
Cross (adjusted)	9	40.067	4.452	5.7
gca	5	5.620	1.405	1.79
sca	4	34.447	6.889	8.82
Intra- block error	6	4.688	.781	
Total	19	50.672		

Table 4. Intra-block estimates of gca effects and their estimated standard error on blooming days.

Parent	Estimates of (gca)	± SE
1	-0.76	0.095
2	0.22	0.095
3	0.61	0.095
4	-0.28	0.095
5	0.65	0.095

Table 5. Intra-block estimates of sca effects and their standard error on blooming days.

sca	Estimates of (sca)	± SE	sca	Estimates of (sca)	± SE
s_{12}	-1.70	0.244	s_{24}	0.10	0.244
s_{13}	2.52	0.244	s_{25}	1.23	0.244
s_{14}	-0.30	0.244	s_{34}	-0.98	0.244
s_{15}	-0.52	0.244	s_{35}	-1.90	0.244
s_{23}	0.37	0.244	s_{45}	1.18	0.244

$$N = Z = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}, (Z' C_{d2} Z)$$

$$= 15/4(I_5 - 1/5\ 1_5\ 1'_5),\ (Z' C_{d2} Z)^{-} = 4/15\ I_5$$

$$C = 1/4 \begin{bmatrix} 6 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 \\ -1 & 6 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & 0 \\ -1 & -1 & 6 & -1 & 0 & -1 & 0 & -1 & 0 & -1 \\ -1 & -1 & -1 & 6 & 0 & 0 & -1 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 & 6 & -1 & -1 & -1 & -1 & 0 \\ -1 & 0 & -1 & 0 & -1 & 6 & -1 & -1 & 0 & -1 \\ -1 & 0 & 0 & -1 & -1 & -1 & 6 & 0 & -1 & -1 \\ 0 & -1 & -1 & 0 & -1 & -1 & 0 & 6 & -1 & -1 \\ 0 & -1 & 0 & -1 & -1 & 0 & -1 & -1 & 6 & -1 \\ 0 & 0 & -1 & -1 & 0 & -1 & -1 & -1 & -1 & 6 \end{bmatrix}$$

The intra-inter block analysis is given in Table 3 and the intra and inter block estimates for gca and sca are given in Table 4 and 5.

REFERENCES

Aggarwal KR (1974). Analysis of Li(s) and triangular designs. J. Indian Soc. Agric. Stat. 26: 3-13.

Agarwal SC, Das MN (1990). Use of n-ary block designs in diallel crosses evaluation. J. Appl. Stat. 17(1).

Bratten OB (1965). The union of partial diallel mating designs and incomplete block environmental designs. Dissertation. North Carolina State University.

Ceranka B, Mejza S (1988). Analysis of diallel table for experiments carried out in BIB design mixed model, Biom. J. 30: 3-16.

Chai FS, Mukerjee R (1999). Optimal design for diallel crosses with specific combining abilities. Biometrika, 86(2): 453-458.

Das A, Dey A, Dean A (1998). Optimal design for diallel cross experiments. Statist. Prob. Letters 36: 427-436.

Dey A, Midha CK (1996). Optimal designs for diallel crosses. Biometrika 83(2): 484-489.

Divecha J, Ghosh DK (1994). Incomplete block designs for complete diallel crosses and their analysis. J. Appl. Stat. 21(5).

Eccleston JA, Hedayat A (1974). On the theory of connected designs: Characterization and optimality. Ann. Stat. 2: 1238-1255.

Gupta S, Kageyama S (1994). Optimal complete diallel crosses. Biometrika, 81: 420-424.

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

James AT, Wilkinson GN (1971). Factorization of the residual operator and canonical decomposition of nonorthogonal factors in the analysis of variance. Biometrika 58: 279-294.

Kempthorne O, Curnow RN (1961). The partial diallel cross. Biometrics 17: 229-250. {See also correction, Biometrics 18(1962): 128.

Mukerjee R (1997). Optimal partial diallel crosses. Biometrika 84(4): 937-946.

Parsad R, Gupta VK, Srivastava R (1999). Optimal designs for diallel crosses. J. Soc. Stat. Comp. Appl. 1: 35-52.

Raghavarao D (1971). Constuction and combinatorial problems in design of experiments. J. Wiley and Sons, New York.

Rao CR (1973). Linear statistical inference and its applications, 2nd edition J. Wiley and Sons, New York.

Shah KR (1960). Optimality criteria for incomplete block designs. Ann. Math. Statist. 22: 235-247.

Sharma MK (1996). Blocking of complete diallel crossing plans using balanced lattice designs. Sankhya B. 58(3): 427-430.

Sharma MK (2004). Optimal complete diallel cross. Advances in mating designs (Chapter 15), Dhanpat Rai &Co. (P) Ltd., N. Delhi.

Sharma M.K., Fanta S (2009). Incomplete block designs for CDC method I and III. Metron. 67(2): 209-226.

Singh M, Hinkelmann K (1998). Analysis of partial diallel crosses in incomplete blocks. Biom. J. 40: 165-181.

Tocher KD (1952). The design and analysis of block experiments (with discussion). J. Royal Stat. Soc. B. 14: 45-100.