

**SOME REMARKS ON THE ANALYSIS OF DIALLEL CROSS EXPERIMENTS  
WITH RECOVERY OF INTER-BLOCK INFORMATION**

**B. Ceranka, H. Chudzik and S. Mejza**

Department of Mathematical and Statistical Methods  
Academy of Agriculture, Poznań, Poland

*Summary*

The paper gives a survey of analyses of diallel cross experiments carried out in block designs. It is assumed that the genotype effects are fixed while the block effects are random. The variance-covariance matrix of the random terms of the considered linear model is related to the structure of the randomization model. The statistical analysis is based on the so called intra- and inter-block analyses and on the combined analysis.

**1. INTRODUCTION**

The paper deals with the diallel cross experiments. This kind of experiments is commonly performed by geneticists who are interested in selecting lines and strains of plants or animals for further breeding. Observations are made on the offspring of the crosses of pairs of inbred lines. The main effect of a given line is called its general combining ability (g.c.a.). The interaction between two lines is called their specific combining ability (s.c.a.). The reader is referred to the paper by Griffing (1956) for more detailed discussion.

Here we consider a situation in which  $p$  inbred lines are chosen and certain crosses are made among them. Griffing (1956) discussed four types of diallel crossing systems. In types I and II the  $p$  parental lines are also included; in types III and IV they are omitted. In types I and III the reciprocal crosses are included, i.e. both, the cross in which the  $i$ -th line is the male line and the  $j$ -th line is the female line and the reciprocal cross with the  $j$ -th as the male line and the  $i$ -th as the female line. In types II and IV the reciprocal crosses are omitted.

Griffing (1956) gave the analysis of the above four types of diallel

crossing systems only for data obtained in a randomized complete block design. In the present paper a generalization of some Griffing's (1956) results connected with the diallel crossing systems concerns the experimental design only. In particular the analysis of interesting genetical characteristics based on experimental data obtained in a proper connected block design is given. It is assumed that genotype effects determined by the diallel crossing systems are fixed. In our approach to the analysis of block experiments the blocks are considered as being a random sample from a finite population of blocks, i.e. the block effects are treated as random. Their statistical properties result from the applied scheme of randomization, which is described in the next section.

The paper is divided into two parts. In the first part, the basis of the analysis of the randomization model for the block designs is given. The problems presented in this part are then adopted to the analysis of diallel cross experiments. It composes the second part of the paper.

It may also be interesting to mention some other papers concerning similar problems connected with the analysis of diallel crosses. In particular Ceranka and Mejza (1987, 1988a, 1988b, 1988c) give the analysis of genotypes for all types of Griffing's classifications for experiments laid out in balanced incomplete block designs. The analysis in these papers is based on another mixed model, in which the genotype effects are fixed and block effects are random, but such in which blocks represent a random sample from an infinite population of blocks. This situation is less realistic in diallel cross experiments than that considered in the present paper. For the classical fixed linear model the similar problems were considered by Ceranka and Kiełczewska (1984, 1985, 1986a, 1986b).

## 2. A RANDOMIZATION MODEL FOR BLOCK DESIGNS

Planning of experiments and inference problems are of fundamental importance for every experimenter using experiments in his corner research work. Usually inferences from experiments are based on the linear model of variables observed in the experiment. There are two approaches to the model building for experimental observations.

In the first approach the form of the model is assumed a priori, usually without taking the actual randomization into account. Such an approach can be questioned because the amount of data necessary to verify the assumptions is usually beyond the reach of the experimenter in collecting them.

In the second approach, the model is strictly connected with the given experiment, i.e. with the structure of its experimental material and with the method of assigning treatments to the units of the material, the so called scheme of randomization. In the paper we consider a two step randomization, i.e. randomization of blocks and randomization of plots

within blocks. Let a population of units (set of available potential units) be divided into  $b$  blocks of size  $k$  each. Let  $v$  denote the number of treatments.

Suppose that randomization is performed as described by Nelder (1954) by randomly permuting labels of blocks within a total area of them and by randomly permuting labels of units within the blocks. Furthermore, it is assumed that the randomizations of units within the blocks are among the blocks independent and that they are also independent of the randomization of blocks.

In our modeling the usual unit treatment additivity (as it is understood by Nelder, 1965, and by White, 1975) is assumed. Practically we assume that the variation of the responses among the available experimental units does not depend on the treatments applied.

Last assumption connected with the modeling concerns the so called technical errors. Namely, it should be noticed that when observing the responses of the units in reality, any observation may be not precise because of an error arising from variations in experimental technique and in various extraneous factors, also perhaps caused by measuring instruments or by the observer. We assume, as usual references, that the technical errors are uncorrelated, with the zero expectation and a constant variance, independent of the treatments in particular.

Finally, adopting the above additivity assumption that the treatment applied on an experimental unit can change only the constant term in the model, the model gets the form

$$y = \Delta' \tau + D' \beta + \eta + \epsilon, \quad (2.1)$$

where  $y$  is an  $nx1$  vector of observations (lexicographically ordered),  $\Delta'$  is an  $nxv$  design matrix for treatments,  $D'$  is an  $nxb$  design matrix for blocks, so that  $N = \Delta D'$  is the  $vxb$  incidence matrix of the block design,  $\tau$  is an  $v \times 1$  vector of treatment parameters,  $\beta$  is a  $b \times 1$  random vector of block effects,  $\eta$  is an  $nx1$  vector of unit errors and  $\epsilon$  is an  $nx1$  vector of technical errors.

The randomization leads to the following statistical properties of the random terms  $\beta$  and  $\eta$  in the model (2.1):

$$E(\beta) = 0, \quad E(\eta) = 0, \quad E(\beta\beta') = \sigma_{\beta}^2(I_b - (1/b)\mathbf{1}_b\mathbf{1}_b'), \quad E(\beta\eta') = 0,$$

$$E(\eta\eta') = \sigma_{\eta}^2(I_b \otimes (I_k - (1/k)\mathbf{1}_k\mathbf{1}_k')),$$

where  $\sigma_{\beta}^2$  denotes the variance of blocks,  $\sigma_{\eta}^2$  denotes the variance of units,  $I_a$  denotes an  $axa$  identity matrix,  $\mathbf{1}_a$  denotes an  $ax1$  vector of ones and  $\otimes$  denotes the Kronecker product of matrices.

It is assumed that  $E(\epsilon) = 0$ ,  $E(\epsilon\epsilon') = \sigma_{\epsilon}^2 I_n$ ,  $\sigma_{\epsilon}^2$  stands for the variance of technical errors,  $E(\beta\epsilon') = 0$ ,  $E(\eta\epsilon') = 0$ . It will be also assumed that  $\beta$ ,  $\eta$  and  $\epsilon$  are normally distributed.

The covariance matrix of the vector  $y$  is then equal to

$$\text{Cov}(y) = \sigma_{\beta}^2(D'D - (1/b)1_n 1_n') + \sigma_{\eta}^2(I_{b_n} \otimes (I_k - (1/k)1_k 1_k')) + \sigma_e^2 I_n.$$

The covariance structure of the model (2.1) is such as structure of typical randomization model. Hence analysis appropriate for this kind of models will be adopted (c.f. Nelder 1965, Houtman and Speed, 1983). In this analysis the set of pairwise orthogonal projectors summing up to the identity matrix plays an important role. In the considered case we have  $P_0 = (1/n)1_n 1_n'$ ,  $P_1 = (1/k)D'D - (1/n)1_n 1_n'$ ,  $P_2 = I_n - (1/k)D'D$ .

Let us note that  $P_i P_i = P_i$ ,  $P_i P_{i'} = 0$ ,  $i, i' = 0, 1, 2$ ,  $i \neq i'$ ,  $P_0 + P_1 + P_2 = I_n$ .

Then the covariance matrix of the vector  $y$  can be written as

$$\text{Cov}(y) = \gamma_0 P_0 + \gamma_1 P_1 + \gamma_2 P_2 \quad (2.2)$$

where  $\gamma_0 = \sigma_e^2$ ,  $\gamma_1 = k\sigma_{\beta}^2 + \sigma_{\eta}^2$ ,  $\gamma_2 = \sigma_{\eta}^2 + \sigma_e^2$ . It means that a proper block design has the orthogonal block structure (c.f. Nelder, 1965, Houtman and Speed, 1983).

The variance components (so-called "strata variances")  $\gamma_0$ ,  $\gamma_1$  and  $\gamma_2$  are usually unknown. In the situation such as the one considered in the paper, the so called within stratum analyses can be helpful. The within stratum  $f$  analysis is based on the model

$$y_f = P_f y, \quad E(y_f) = P_f \Delta' \tau, \quad \text{Cov}(y_f) = \gamma_f P_f, \quad f=0,1,2. \quad (2.3)$$

The covariance matrices in the linear models (2.3) are of ranks  $r(P_0)=1$ ,  $r(P_1)=b-1$ ,  $r(P_2)=n-b$ , and such that  $\text{Cov}(y_f) P_f \Delta' = \gamma_f P_f \Delta'$  for each  $f$ , which, as can be proved (see for example Bailey, 1981, Houtman and Speed, 1983, Rao and Mitra, 1971, Th. 8.2.1 with Corollary 3), are sufficient conditions for the BLUE's of estimable linear functions in the models (2.3) to be equal to the BLUE's obtainable by the simple least squares method based on the linear model

$$y_f^* = P_f y, \quad E(y_f^*) = P_f \Delta' \tau, \quad \text{Cov}(y_f^*) = \gamma_f I_{f'}, \quad f=1,2. \quad (2.4)$$

Note that the normal equations for estimating the fixed effects of model (2.4) have the form

$$\Delta P_f \Delta' \tau_f^0 = \Delta P_f y, \quad (2.5)$$

i.e.  $A_f \tau_f^0 = Q_f$ , where  $A_f = \Delta P_f \Delta'$ ,  $Q_f = \Delta P_f y$ ,  $f=0,1,2$ .

The estimability of a function  $c' \tau$  within the  $f$ -th stratum can be verified by the criterion  $c' A_f^{-} A_f c$  (cf. Rao and Mitra, 1971, Th. 7.2.1).

From the fact  $A_f 1 = 0$  hold for  $f=1,2$  we have, that if a linear function  $c' \tau$  is estimable in any of the two strata then it must be a contrast, i.e.  $c' 1 = 0$ . On the other hand, the only functions estimable in stratum 0 are those of the type  $c' \tau = (c' 1/n) r' \tau$ , for any  $c$  such that  $c' 1 = 0$ .

If a function  $c' \tau$  is estimable within the  $f$ -th stratum then its

BLUE within that stratum has the form  $(\hat{c}'\tau_f) = c'\tau_f^0 = c'A_f^{-1}AP_f'y$  with  $\text{var}((\hat{c}'\tau_f)) = \gamma_f c'A_f^{-1}c$  for  $f=0,1,2$ , where  $A_f^{-1}$  denotes generalized inverse of matrix  $A_f$ .

Let us consider any contrast of treatment parameters  $c'\tau$ . This contrast can be estimated only in one of the strata or in both strata. In the first case, BLUE of this contrast is BLUE in the general model (2.1).

There are some problems with estimation of contrast in the second case. It is known that the strata estimators are unbiased in the model (2.1). To improve the statistical properties (to minimize variance) we can use some methods of combining estimators (cf Khatri and Shah, 1974, 1975, Bhattacharya, 1980). The methods of combining estimators utilize usually the set of some contrasts called basic contrasts.

The analysis of model (2.1) is based on two strata. The first stratum, connected with the projector  $P_1$ , is called inter-block stratum. The second, connected with the projector  $P_2$ , is called intra-block stratum. The latter stratum is practically the most important in the planning and analysis of the block designs.

In the analysis of variance in both strata certain contrasts, called basic, play an important role. Let  $\epsilon_i$  and  $w_i$  be eigenvalues and eigenvectors, respectively, of the matrix  $A_2$  with respect to the matrix  $R$ , i.e.  $A_2 w_i = \epsilon_i R w_i$ ,  $i=1, \dots, v$ , where  $R$  is a diagonal matrix with the diagonal elements equal to the number of replications. The eigenvectors  $w_i$  can be chosen to be mutually orthogonal. Since  $A_2 1 = 0$ , the last (say) eigenvector  $w_v$  may be taken equal to  $(1/\sqrt{n})1$ . It may be shown that all the eigenvalues are in the range from 0 to 1. It is convenient to order the eigenvectors so that these which correspond to the eigenvalues equal to 1,  $g$  in number, are at the beginning and these with nonzero (less than 1) eigenvalues,  $v-g-1$  in number, come at the end (cf Pearce et al., 1974).

For any block designs, contrasts of treatment parameters  $s_i'\tau$ ,  $i=1, \dots, v-1$ , are said to be basic contrasts if  $s_i = R w_i$ , where  $w_i$  are pairwise orthonormal eigenvectors of matrix  $A_2$  with respect to the matrix  $R$ , i.e.  $w_i' R w_i = 1$  and  $w_i' R w_{i'} = 0$ ,  $i, i'=1, \dots, v-1$ ,  $i \neq i'$ .

It can be proved (see Pearce et al., 1974, Mejza, 1985) that

(i) in the first stratum (inter-block) only these basic contrasts  $s_i'\tau$  are estimable which are defined by the eigenvectors  $w_{g+1}, w_{g+2}, \dots, w_{v-1}$  of the matrix  $A_2$  with respect to the matrix  $R$ , that correspond to the nonzero eigenvalues of matrix  $A_2$ ;

(ii) the BLUEs of any of these contrasts are given by

$$\tilde{s}_i'\tau = (1/\epsilon_{oi}) w_i' Q_1, \quad i=g+1, \dots, v-1,$$

and the variances of  $\tilde{s}_i'\tau$  are

$$\text{var}(\tilde{s}_i'\tau) = \gamma_1 / \epsilon_{oi}, \quad i=g+1, \dots, v-1,$$

where  $\epsilon_{oi} = 1 - \epsilon_i$ ,

(iii) all estimators  $\tilde{s}_i^{\tau}$ ,  $i=g+1, \dots, v-1$ , are independently distributed according to  $\tilde{s}_i^{\tau} \sim N(s_i^{\tau}, \gamma_1/\epsilon_{oi})$ .

Additionally, it can be also proved that:

(i) in the second stratum (intra-block) only these basic contrasts  $\tilde{s}_i^{\tau}$  are estimable which are defined by the eigenvectors  $w_1, w_2, \dots, w_{v-1}$  of the matrix  $A_2$  with respect to the matrix  $R$ ,

(ii) the BLUEs of any of these contrasts are given by

$$\tilde{s}_i^{\tau} = (1/\epsilon_i)w_i'Q_2, \quad i=1, \dots, v-1,$$

and the variances of  $\tilde{s}_i^{\tau}$  are

$$\text{var}(\tilde{s}_i^{\tau}) = \gamma_2/\epsilon_i, \quad i=1, \dots, v-1,$$

(iii) all estimators  $\tilde{s}_i^{\tau}$ ,  $i=1, \dots, v-1$ , are independently distributed according to  $\tilde{s}_i^{\tau} \sim N(s_i^{\tau}, \gamma_2/\epsilon_i)$ .

Any contrast of treatment parameters  $c^{\tau}$  can be expressed as  $c^{\tau} = \sum_{i=1}^{v-1} d_i \tilde{s}_i^{\tau}$ . In this form two cases can be considered. The first one is such that some of  $d_i$  are different from zero, i.e.  $d_i \neq 0$  for  $i=1, \dots, g$ , and the rest of  $d_i$  are equal to zero,  $d_i=0$ ,  $i=g+1, \dots, v-1$ . The second situation is such that  $c^{\tau}$  is a linear combination of the basic contrasts  $\tilde{s}_i^{\tau}$ ,  $i=1, \dots, v-1$ , but at least one of  $d_i$ ,  $i=g+1, \dots, v-1$ , is a nonzero coefficient. The best situation is the first one. In this case the estimator of  $c^{\tau}$  is the BLUE in the model (2.1). Then the BLUE of  $c^{\tau}$  in the model (2.1) is equal to BLUE of  $c^{\tau}$  in the second stratum (intra-block analysis). In the second case the first and the second stratum estimators will be used to obtain a new, combined, estimator, which possesses some desirable statistical properties. As was mentioned previously the Khatri and Shah's (1974) method of combined estimators will be recommended.

According to this method the combined estimator of  $c^{\tau}$  has a form

$$c^{\tau} = \sum_{i=1}^g d_i \hat{s}_i^{\tau} + \sum_{i=g+1}^{v-1} d_i [(1-d_i)(\hat{s}_i^{\tau} - \tilde{s}_i^{\tau}) + \tilde{s}_i^{\tau}], \quad (2.6)$$

where

$$d_i = \epsilon_i^{-1} s_2^2 / \{ \epsilon_i^{-1} s_2^2 + \epsilon_{oi}^{-1} (n-b-v+3) [s_1^2 + \sum_{i=g+1}^{v-1} (\hat{s}_i^{\tau} - \tilde{s}_i^{\tau})^2 \epsilon_{oi}] (b-3)^{-1} \},$$

where  $s_1^2$ ,  $s_2^2$  are the stratum mean squares for error.

The combined estimator (2.6) has uniformly smaller variance than the estimator of  $c^{\tau}$  obtained in intra-block analysis only, provided  $b > 3$ .

The analysis of data obtained in the considered design usually contains the tests of general and particular hypotheses. The tests can be obtained from the within stratum analysis of variance as given in Table 1.

Tab.1. Analysis of variance for the stratum  $f$ .

Source of variation	d.f.	S.S.	E(S.S.)
Treatments (in $f$ )	$\nu_{Tf}$	$SST_f$	$\nu_{Tf}\gamma_f + \tau'A_f\bar{\tau}$
Error (in $f$ )	$\nu_{Ef}$	$SSE_f$	$\nu_{Ef}\gamma_f$
Total (in $f$ )	$\nu_f$	$SSY_f$	$\nu_f\gamma_f + \tau'A_f\bar{\tau}$

The symbols occurring in Table 1 denote  $SSY_f = y'P_f y$  - the total sum of squares,  $SST_f = Q'_f A_f Q_f$  - the treatment sum of squares,  $SSE_f = SSY_f - SST_f$  - the error sum of squares,  $\nu_{Tf} = r(A_f)$ ,  $\nu_f = r(P_f)$ ,  $\nu_{Ef} = \nu_f - \nu_{Tf}$ , for the  $f$ -th stratum,  $f=1, 2$ .

Under the assumption that the random terms of the linear model (2.1) are normally distributed it is easy to obtain an exact test (F-test) of the hypothesis  $H_{0f} : \tau'A_f\bar{\tau} = 0$ ,  $f=1, 2$ .

Let us consider  $h$  independent contrasts of treatment parameters,  $c'_i \tau$ , where  $c'_i 1 = 0$ , for  $i=1, \dots, h$ . If we are interested in testing the hypothesis  $H_{0f}^* : C'\tau = 0$ , where  $C = [c_1, \dots, c_h]$ , with  $r(C) = h$ , and all of the contrasts  $c'_i \tau$  are estimable in the  $f$ -th stratum, then the appropriate F-statistic is

$$F_f = \frac{Q'_f A_f C (C' A_f C)^{-1} C' A_f Q_f}{SSE_f / \nu_{Ef}} \quad (2.7)$$

The F-statistic (2.7) under  $H_{0f}^*$  has the F-distribution with  $h$  and  $\nu_{Ef}$  degrees of freedom.

For some of hypothesis,  $H_{0f}^* : C'\tau = 0$ , two independent tests are available, i.e. in the first and the second stratum. Then the combined test can be used to improve the statistical properties of the stratum tests. According to results obtained by Prasad and Subramanyam (1986) the procedure based on Fisher's method of combining tests can be recommended.

The statistic

$$z = -2 \log \alpha_1 \alpha_2 \quad (2.8)$$

is distributed as a  $\chi^2$  with 4 degrees of freedom, where  $\alpha_1$  and  $\alpha_2$  are the significance levels of  $F_f$  given by (2.7).

### 3. ESTIMATION IN DIALLEL EXPERIMENT

#### 3.1. Type I.

Let  $p$  denote the number of parental lines in all four types of diallel crossing.

Let us consider the type I of diallel crossing, i.e. the genotypes

obtained in a diallel crossing system including parents, one set of  $F_1$ 's and reciprocal  $F_1$ 's. The vector of genotype effects can be presented as  $\tau = [\tau_{11}, \dots, \tau_{1p}, \dots, \tau_{p1}, \dots, \tau_{pp}]'$ . The effect  $\tau_{ij}$  is assumed to be

$$\tau_{ij} = \mu + g_i + g_j + s_{ij} + w_{ij}, \quad i, j = 1, 2, \dots, p,$$

where  $\mu$  is the general parameter,  $g_i(g_j)$  is the general combining ability (g.c.a.) effect of the  $i$ -th ( $j$ -th) parent,  $s_{ij}$  is the specific combining ability (s.c.a.) effect of the cross between the  $i$ -th and  $j$ -th parents, such that  $s_{ij} = s_{ji}$ , and  $w_{ij}$  is the reciprocal effect involving the reciprocal crosses between the  $i$ -th and  $j$ -th parents such that  $w_{ij} = -w_{ji}$ .

In the paper the g.c.a. always will be denoted by  $g_i$ , s.c.a. - by  $s_{ij}$  and reciprocal effect - by  $w_{ij}$ .

As we have stated, the g.c.a., s.c.a. and reciprocal effects are certain contrasts which can be defined as follows:  $g_i = c'_i \tau$ ,  $i = 1, 2, \dots, p$ ,  $s_{ij} = c'_{ij} \tau$ ,  $i, j = 1, 2, \dots, p$ , and  $w_{ij} = d'_{ij} \tau$ ,  $i \neq j$ ,  $i, j = 1, 2, \dots, p$ . This convention is used in whole paper. Hence, it is sufficient to describe how to form the vectors defining the above contrasts i.e.,  $c_i$ ,  $c_{ij}$ ,  $d_{ij}$ . Since the elements of the vector  $\tau$  have a double index we introduce also the double indexes of the elements of vectors  $c_i$ ,  $c_{ij}$  and  $d_{ij}$ , respectively.

The  $(k, l)$ -th element of the vector  $c_i$ ,  $i, k, l = 1, 2, \dots, p$ , (defining the  $g_i$ ) is as follows:

$$c_{ikl} = \frac{1}{2p^2} \begin{cases} 2(p-1), & k=l=i, \\ p-2, & k=i \text{ or } l=i, \\ -2, & \text{otherwise.} \end{cases}$$

The  $(k, l)$ -th element of the vector  $c_{ii}$ ,  $i, k, l = 1, 2, \dots, p$ , (defining the  $s_{ii}$ ) is as follows:

$$c_{iikl} = \frac{1}{p^2} \begin{cases} (p-1)^2, & k=l=i, \\ -(p-1), & k=i \text{ or } l=i, \\ 1, & \text{otherwise.} \end{cases}$$

The  $(k, l)$ -th element of the vector  $c_{ij}$ ,  $i, j, k, l = 1, 2, \dots, p$ ,  $i \neq j$ , (defining the  $s_{ij}$ ) is as follows:

$$c_{ijkl} = \frac{1}{2p^2} \begin{cases} -2(p-1), & k=l=i \text{ or } k=l=j, \\ p^2 - 2p + 2, & k=i \text{ and } l=j \text{ or } k=j \text{ and } l=i, \\ -(p-2), & k=i \text{ or } l=i \text{ or } k=j \text{ or } l=j, \\ 1, & \text{otherwise.} \end{cases}$$

The  $(k, l)$ -th element of the vector  $d_{ij}$ ,  $i, j, k, l = 1, 2, \dots, p$ ,  $i \neq j$ , (defining the  $w_{ij}$ ) is as follows:

$$d_{ijkl} = \frac{1}{2} \begin{cases} 1, & k=i \text{ and } l=j, \\ -1, & k=j \text{ and } l=i, \\ 0, & \text{otherwise.} \end{cases}$$



### 3.2. Type II.

Let us consider the type II of diallel crossing, i.e. the genotypes obtained in a diallel crossing system including parents and one set of  $F_1$ 's. The number of genotypes (treatments) is  $v=p(p+1)/2$ .

The vector of genotype effects has the structure  $\tau = [\tau_{11}, \dots, \tau_{1p}, \tau_{22}, \dots, \tau_{2p}, \dots, \tau_{pp}]'$ , where the effect  $\tau_{ij}$  is assumed to be expressed as

$$\tau_{ij} = \mu + g_i + g_j + s_{ij}, \quad i, j = 1, 2, \dots, p, \quad i \leq j.$$

As in previous case we have that the  $(k, l)$ -th element of the vector  $c_i$ ,  $i, k, l = 1, 2, \dots, p$ , is as follows:

$$c_{ikl} = \frac{1}{p(p+2)} \begin{cases} 2(p-1), & k=l=i, \\ p-2, & k=i \text{ or } l=i, \\ -2, & \text{otherwise.} \end{cases}$$

The  $(k, l)$ -th element of the vector  $c_{ii}$ ,  $i, k, l = 1, 2, \dots, p$ , is as follows:

$$c_{iikl} = \frac{1}{(p+1)(p+2)} \begin{cases} p(p-1), & k=l=i, \\ -2p, & k=i \text{ or } l=i, \\ 2, & \text{otherwise.} \end{cases}$$

The  $(k, l)$ -th element of the vector  $c_{ij}$ ,  $i, j, k, l = 1, 2, \dots, p$ ,  $i < j$ , is as follows:

$$c_{ijk l} = \frac{1}{(p+1)(p+2)} \begin{cases} -2p, & k=l=i \text{ or } k=l=j, \\ p^2+p+2, & k=i \text{ and } l=j \text{ or } k=j \text{ and } l=i, \\ -(p-1), & k=i \text{ or } l=i \text{ or } k=j \text{ or } l=j, \\ 2, & \text{otherwise.} \end{cases}$$

### 3.3. Type III.

Let us consider the diallel crossing of type III, i.e. the genotypes obtained in a diallel crossing system including one set of  $F_1$ 's and reciprocal  $F_1$ 's. The number of genotypes (treatments) is  $v=p(p-1)$ , while the vector of genotype effects now is presented as  $\tau = [\tau_{12}, \dots, \tau_{1p}, \dots, \tau_{p1}, \dots, \tau_{p,p-1}]'$ . The effect  $\tau_{ij}$  is assumed to be expressed as

$$\tau_{ij} = \mu + g_i + g_j + s_{ij} + w_{ij}, \quad i, j = 1, 2, \dots, p, \quad i \neq j.$$

As in previous cases we have that the  $(k, l)$ -th element of the vector  $c_i$ ,  $i, k, l = 1, 2, \dots, p$ , has the form:

$$c_{ikl} = \frac{1}{2p(p-2)} \begin{cases} p-2, & k=i \text{ or } l=i, \\ -2, & \text{otherwise.} \end{cases}$$

The  $(k, l)$ -th element of the vector  $c_{ij}$ ,  $i, j, k, l = 1, 2, \dots, p$ ,  $i \neq j$ , is as follows:

$$c_{ij_{kl}} = \frac{1}{2(p-1)(p-2)} \begin{cases} (p-2)(p-3), & k=i \text{ and } l=j \text{ or } k=j \text{ and } l=i, \\ -(p-3), & k=i \text{ or } l=j \text{ or } k=j \text{ or } l=i, \\ 2, & \text{otherwise.} \end{cases}$$

The  $(k,l)$ -th element of the vector  $d_{ij}$ ,  $i,j,k,l = 1,2,\dots,p$ ,  $i \neq j$ , is as follows:

$$d_{ij_{kl}} = \frac{1}{2} \begin{cases} 1, & k=i \text{ and } l=j, \\ -1, & k=j \text{ and } l=i, \\ 0, & \text{otherwise.} \end{cases}$$

### 3.4. Type IV.

Let us consider the type IV of diallel crossing, i.e. the genotypes obtained in a diallel crossing system including one set of  $F_i$ 's. The number of genotypes (treatments) is in this case equal to  $v=p(p-1)/2$ . The vector of genotype effects has the structure  $\tau = [\tau_{12}, \dots, \tau_{1p}, \tau_{23}, \dots, \tau_{2p}, \dots, \tau_{p-1,p}]'$ . The effect  $\tau_{ij}$  is assumed to be expressed as usually i.e.,

$$\tau_{ij} = \mu + g_i + g_j + s_{ij}, \quad i,j = 1,2,\dots,p, \quad i < j.$$

As in previous cases the  $(k,l)$ -th element of the vector  $c_i$ ,  $i,k,l=1,2,\dots,p$ , is as follows

$$c_{i_{kl}} = \frac{1}{p(p-2)} = \begin{cases} p-2, & k=i \text{ or } l=i, \\ -2, & \text{otherwise.} \end{cases}$$

The  $(k,l)$ -th element of the vector  $c_{ij}$ ,  $i,j,k,l = 1,2,\dots,p$ ,  $i < j$ , is as follows

$$c_{ij_{kl}} = \frac{1}{(p-1)(p-2)} \begin{cases} (p-2)(p-3), & k=i \text{ and } l=j, \\ -(p-3), & k=i \text{ or } l=j, \\ 2, & \text{otherwise.} \end{cases}$$

Using the above vectors it is possible to obtain the intra- and the inter-block estimators. As the final estimators in all cases we suggest to use the combined ones.

## 4. TESTING HYPOTHESES

In the analysis of diallel crossing experiments the interesting hypotheses concerning the g.c.a., s.c.a. and reciprocal effects can be formulated as follows:

1.  $H_0 : g_i = 0$ ; for all  $i$ ,
2.  $H_0 : s_{ij} = 0$ ; for all  $i,j$ ,
3.  $H_0 : g_i = 0$ ; for fixed  $i$ ,
4.  $H_0 : g_i - g_j = 0$ ;  $i \neq j$ ,

5.  $H_0 : s_{ij} = 0$ ; for fixed  $i, j$ ,
  6.  $H_0 : s_{ij} - s_{kl} = 0$ ;
  7.  $H_0 : w_{ij} = 0$ ; for all  $i, j$ ,
  8.  $H_0 : w_{ij} = 0$ ; for fixed  $i, j$ ,
  9.  $H_0 : w_{ij} - w_{kl} = 0$ ,
- where  $i, j, k, l = 1, \dots, p$ .

The hypotheses 1 - 6 are testable in the all types of diallel crossing systems. The hypotheses 7, 8 and 9 are testable in the type I and III only.

In the present paper, the combined test (2.8) is recommended. With respect to the above, we need the significance levels of the appropriate F-tests used to test the same hypothesis.

For testing the hypothesis 1 we have to construct matrix  $C$  describing the independent contrasts. It can be noticed that each effect  $g_i$ ,  $i=1, \dots, p$ , is the contrast of the form  $c_i' \tau$ . Taking into account the relation  $\sum_{i=1}^p g_i = 0$ , for constructing matrix  $C$  we can take any set of  $p-1$  vectors  $c_i$  defining  $g_i$ , which is a set of independent contrasts. Then the matrix  $C$  from (2.7) has, for example, the form  $C=[c_1, c_2, \dots, c_{p-1}]$ . Thus, for the testing of hypothesis 1 we take the above matrix  $C$  and F-statistics given by (2.7) for  $f = 1, 2$ , assuming  $h=p-1$ . In the next step we calculate the significance levels of  $F_1$  and  $F_2$ , which are later used in a combined test (2.8).

For testing the hypothesis 2 we also have to construct matrix  $C$  describing independent contrasts concerning the s.c.a. effects. From the definition of s.c.a. it can be seen that each effect  $s_{ij}$  is the contrast of treatment parameters of the form  $c_{ij}' \tau$ . Thus, for the testing of hypothesis 2 we take the matrix  $C$  and F-statistics given by (2.7) for  $f = 1, 2$ , assuming  $h=p(p-1)/2$  for type I and II or  $h=p(p-3)/2$  for type III and IV. In the next step we calculate the significance levels of  $F_1$  and  $F_2$ , which are later used in a combined test (2.8).

For testing the hypothesis 7 we take F-statistics given by (2.7) for  $f=1, 2$ , the matrix  $C$  which is formed by  $p(p-1)/2$  independent contrasts describing reciprocal effects  $w_{ij}$ , and  $h=p(p-1)/2$ .

For testing the hypotheses 3 - 6, 8 and 9, we use the F-statistics given by (2.7), and the combined test (2.8) as usually, assuming  $h=1$ .

## 5. AN EXAMPLE

For illustrating the theory given in this paper, let us consider, as an example, an experiment carried out in a block design with  $v=10$  genotypes resulting from the diallel crossing among  $p=5$  inbred lines of sunflower. In this experiment the seed yields was observed. The genotypes were allocated in  $b=5$  blocks of size  $k=4$ . Each genotype was replicated

r=6 times in the experiment. The experimental results are presented in Table 2.

Tab.2. Experimental results

Number of block	Number of treatments			
	1	2	3	4
1	13.4	14.4	16.6	15.4
2	13.4	17.6	13.0	13.5
3	13.5	19.1	15.0	13.0
4	13.5	18.0	15.1	21.6
5	14.0	13.1	18.4	17.1
6	14.6	15.6	13.9	16.6
7	18.0	20.3	15.9	18.4
8	18.0	18.0	19.0	17.2
9	18.7	13.2	13.9	18.0
10	20.5	20.4	15.1	18.1
11	20.8	14.2	20.1	20.0
12	20.7	17.0	20.7	20.1
13	16.6	18.8	14.1	16.2
14	19.1	17.7	17.1	21.5
15	21.4	19.0	17.2	19.0

The first step in the analysis is to calculate the sums of squares, the mean squares and the F-ratio in the intra- and - inter-block analyses, which are summarized in Table 3 and Table 4, respectively.

Tab.3. Intra-block analysis of variance

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F
Genotypes	183.01	9	20.33	12.89
Error	56.78	36	1.58	-
Total	239.79	45	-	-

Tab. 4. Inter-block analysis of variance

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F
Genotypes	157.23	9	17.47	10.14
Error	8.61	5	1.72	-
Total	165.84	14	-	-

The next step is the estimation of effects of genotypes in intra- and inter-block analyses, which elements are arranged in the following triangular tables:

- in the intra-block analysis the estimated genotypes effects are

i	j	1	2	3	4
1		14.95	18.38	19.54	18.54
2			14.19	15.65	18.97
3				14.80	17.36
4					18.91

- in the inter-block analysis the estimated genotypes effects are

i	j	1	2	3	4
1		7.67	15.32	18.39	19.99
2			14.34	19.84	20.17
3				14.89	21.82
4					18.94

Now, using the formulae given above we can estimate the g.c.a. and s.c.a. effects. The estimates of g.c.a. effects are as follows:

- in the intra-block analysis they are

$$\hat{g}_1 = 0.112, \quad \hat{g}_2 = -0.717, \quad \hat{g}_3 = -0.562, \quad \hat{g}_4 = 1.167,$$

- in the inter-block analysis they are

$$\tilde{g}_1 = -2.775, \quad \tilde{g}_2 = -0.279, \quad \tilde{g}_3 = 0.692, \quad \tilde{g}_4 = 2.362,$$

- in combined analysis they are

$$\bar{g}_1 = 0.102, \quad \bar{g}_2 = -0.716, \quad \bar{g}_3 = -0.557, \quad \bar{g}_4 = 1.172.$$

The estimates of s.c.a. effects are arranged in triangular tables:

- in the intra-block analysis

i	j	1	2	3	4
1					
1		-2.413	1.846	2.855	0.126
2			-1.510	-0.206	1.380
3				-1.132	-0.386
4					-0.560

- in the inter-block analysis

i	j	1	2	3	4
1					
1		-3.920	1.234	3.338	3.268
2			-2.237	2.293	0.947
3				-3.628	1.626
4					-2.920

- in combined analysis

i	j	1	2	3	4
1					
1		-2.426	1.819	2.876	0.265
2			-1.516	-0.096	1.361
3				-1.153	-0.297
4					-0.580

As we can see from the results of this example, the combined estimates are nearly the estimates obtained from intra-block analysis. This follows from the fact, that main part of information about contrasts of treatment parameters is included in intra-block analysis. In our example more than 83% of information is contained in that analysis. Only less than 17% of information is recovered from inter-block analysis by combination of contrasts. The differences between intra-block estimates and combined estimates are more perceptible if more information is contained in the inter-block analysis. In planning an experiment it is desirable to have more information in intra-block analysis, but such a situation is not always possible.

The comparison of the performances of the individual lines is of considerable interest for a plant breeder. Inferences on them can be made by testing some hypotheses 1-6.

For testing of hypothesis that there are no differences among the g.c.a. effects (hypothesis 1), we calculate  $F_1 = 64.256$  ( $\alpha_1 = 2 \cdot 10^{-3}$ ) and  $F_2 = 13.985$  ( $\alpha_2 = 3 \cdot 10^{-5}$ ) and hence  $z = 42,201$  ( $P(\chi_4^2/H_0) = 2 \cdot 10^{-7}$ ).

For testing of hypothesis that there are no differences among the s.c.a. effects (hypothesis 2), we calculate  $F_1 = 28.711$  ( $\alpha_1 = 10^{-3}$ ) and  $F_2 = 12.347$  ( $\alpha_2 = 2 \cdot 10^{-6}$ ) and hence  $z = 45.020$  ( $P(\chi_4^2/H_0) = 4 \cdot 10^{-8}$ ).

For testing of hypothesis 3, for example  $H_0: g_1 = 0$ , we calculate  $F_1 = 143,025$  ( $\alpha_1 = 7 \cdot 10^{-4}$ ) and  $F_2 = 0.316$  ( $\alpha_2 = 0.577$ ) and hence  $z = 20.174$ , ( $P(\chi_4^2/H_0) = 5 \cdot 10^{-3}$ ).

For testing of hypothesis 4, for example  $H_0: g_1 - g_2 = 0$ , we calculate  $F_1 = 43.386$  ( $\alpha_1 = 10^{-3}$ ) and  $F_2 = 6,539$  ( $\alpha_2 = 0.015$ ) and hence  $z = 21.843$  ( $P(\chi_4^2/H_0) = 2 \cdot 10^{-3}$ ).

For testing of hypothesis 5, for example  $H_0: s_{11} = 0$ , we calculate  $F_1 = 89.188$  ( $\alpha_1 = 2 \cdot 10^{-3}$ ) and  $F_2 = 46.158$  ( $\alpha_2 = 6 \cdot 10^{-7}$ ) and hence  $z = 50.017$  ( $P(\chi_4^2/H_0) = 4 \cdot 10^{-9}$ ).

For testing of hypothesis 6, for example  $H_0: s_{12} - s_{13} = 0$ . we calculate  $F_1 = 6.167$  ( $\alpha_1 = 0.057$ ) and  $F_2 = 1.937$  ( $\alpha_2 = 0.173$ ) and hence  $z = 9.293$  ( $P(\chi_4^2/H_0) = 0.054$ ).

#### REFERENCES

- Bailey R.A. (1981). A unified approach to design of experiments. *J.R. Statist. Soc. A* **144**, 214-223.
- Ceranka B., Kiełczewska H. (1984). Analiza potomstwa z trójkątnej tablicy diallelicznej porównywanego w układach blokowych. *Listy Biometryczne* **21**, 57-67.
- Ceranka B., Kiełczewska H. (1985). Analiza potomstwa z tablicy diallelicznej porównywanego w układach blokowych. *Listy Biometryczne* **22**, 13-23.
- Ceranka B., Kiełczewska H. (1986a). Analysis of triangular diallel table for experiments in block designs. *Biometrical Journal* **28**, 149-157.
- Ceranka B., Kiełczewska H. (1986b). Analysis of diallel table for experiments in block designs. *Biometrical Journal* **28**, 529-538.
- Ceranka B., Mejza S. (1987). Analysis of a triangular diallel table experiment in a BIB design - a mixed model. *Rivista di Statistica Applicata* **20**, 329-339.
- Ceranka B., Mejza S. (1988a). Analysis of diallel table experiments carried out in BIB designs - mixed model. *Biometrical Journal* **30**, 3-16.
- Ceranka B., Mejza S. (1988b). Analysis of offsprings of a diallel experiment in a BIB design - a mixed model. *EDV in Medizin and Biologie* **19**, 70-74.
- Ceranka B., Mejza S. (1988c). Intra- and inter-block analysis of a triangular diallel table experiment in a BIB design. *Genetica Polonica*, **29**, 143-152.

- Fisher R.A. (1975). *Statistical methods for research workers*. Oliver and Boyd, London.
- Griffing B. (1956). Concept of general and specific combining ability in relation to diallel crossing systems. *Australian J. Biol. Sci.* **9**, 463-493.
- Houtman A.M., Speed T.P. (1983). Balance in designed experiments with orthogonal block structure. *Ann. Statist.* **11**, 1069-1085.
- Khatri C.G., Shah K.R. (1974). Estimation of location parameters from two linear models under normality. *Commun. Statist.* **3**, 647-663.
- Khatri C.G., Shah K.R. (1975). Exact variance of combined inter- and intra-block estimates in incomplete block designs. *J. Amer. Statist. Ass.*, **70**, 402-406.
- Mejza S. (1985). On testing hypotheses in a mixed linear model for incomplete block designs. *Scand. J. Statist.* **12**, 241-247.
- Nelder J.A. (1954). The interpretation of negative components of variance. *Biometrika* **41**, 544-548.
- Nelder J.A. (1965). The analysis of randomized experiments with orthogonal block structure. *Proc. R. Soc. A* **283**, 147-178.
- Pearce S.C., Caliński T., Marshall T.F. de C., (1974). The basic contrasts of an experimental design with special reference to the analysis of data. *Biometrika* **61**, 449-460.
- Prasad M.S., Subramanyam A. (1986). Asymptotic efficiency of sequential and nonsequential procedures based on Fisher's method of combining tests. *Commun. Statist. A. Theory Methods* **15**, 1665-1680.
- Rao C.R., Mitra S.K. (1971). *Generalized inverse of matrices and its applications*. Wiley, New York.
- White R.F. (1975). Randomization and the analysis of variance. *Biometrics* **31**, 555-571.

Received 1 May 1990; revised 19 November 1990

PEWNE UWAGI O ANALIZIE DOŚWIADCZEŃ Z KRZYŻOWANIEM DIALLELICZNYM  
Z ODZYSKIWIANIEM INFORMACJI MIĘDZYBLOKOWEJ

Streszczenie

W pracy przedstawiono przegląd analizy dotyczącej krzyżowania diallelicznego dla doświadczeń zakładanych w układach bloków. Przyjęto założenia, że efekty genotypów są stałe, a efekty bloków losowe. Macierz kowariancji w modelu liniowym dotycząca składników losowych jest uzależniona od modelu randomizacji. Analiza statystyczna oparta jest na wewnątrz i międzyblokowej analizie wariancji oraz analizie kombinowanej.

**Słowa kluczowe:** krzyżowanie dialleliczne, ogólna zdolność kombinacyjna, efekt krzyżowania odwrotnego, specyficzna zdolność kombinacyjna, odzyskiwanie informacji międzyblokowej