SHARE PROGRAM LIBRARY AGENCY

SHARE

PROGRAM NUMBER

137001

University of Miami

1365 MEMORIAL DRIVE - CORAL GABLES, FLORIDA
(305) - 284-6257

SHARE Program Library Agency
Triangle Universities Computation Center
P. O. Box 12076
Research Triangle Park, N. C. 27709

CONTRIBUTED PROGRAM LIBRARY SUBM
(for IBM S/360, 1130 and 1800)

This form should be completed and submitted with the program package to PID at the address shown above. Standards and instructions for submitting programs are in your *User Group Reference Manual* or the *Contributed Program Submittal Standards Manual* available from PID.

(1) Program Order Number *(to be filled in by PID)* . . . 360D-13.7.001

(2) System Type *(machine)* . . . S/360

(3) Search Key . . . /GENERAL/ LEAST /SQUARES/ DIALLEL ANALYSIS /OF/ VARIANCE

(4) Programming Language . . . FORTRAN IV

(5) Author's Name and Address . . .
Henry E. Schaffer and Richard A. Usanis
Department of Genetics, N. C. State University
Raleigh, North Carolina 27607

(6) Direct Inquiries to Name and Address *(if different than Author)*

DISCLAIMER

(7) Title of Program . . . DIALL - General Least Squares Diallel Analysis of Variance

(8) Submitter's User Group Affiliation Code and Installation Code . . . S NCS

(9) Submitter's Own Program Identification and Suffix *(optional)* . . .

(10) Primary Subject Code . . . 13.7

(11) Secondary Subject Codes . . . 17.5

(12) Operating or Monitor System Required . . . OS/360

(13) New or Revision Code *(if revision, show prior Program Order Number in item 1)* . . . N

(14) Year Completed . . . 68

(15) Date of Submittal . . . 062369

(16) Documentation *(number of original pages submitted)* . . . 43

(17) **Abstract** *(should contain sufficient information for a reader to determine the value of the program). Listed on the reverse side of this form are subjects which may serve as a guide for a descriptive abstract.*

CONTRIBUTED PROGRAM LIBRARY SUBMITTAL FORM

Subject Guide

a. Purpose
b. Programming Language used
c. Version and modification level or release number of IBM Programming System used, or program order number for non-IBM authored program used
d. Field of application
e. Type of routine (main program, subroutine, etc.)
f. Specific description of machine requirements
g. Engineering Changes (EC) level of equipment (if pertinent)

ABSTRACT

Two main programs, DIALL and DIALLC, which do a general least squares analysis for a general (unbalanced) diallel experiment, are described in detail. The programs, with their subroutines, will compute the analysis of variance and analysis of cross-products tables for any number of variables including the expectations of the mean squares, calculate the estimates of the variance components and estimate the correlations between the effects for different traits. Two examples are given of the use of these programs.

For a complete analysis of a 5 line diallel experiment with two replicates and one variable approximately 42K bytes of memory are needed. The memory requirements increase more rapidly than an increase in the number of lines.

The programs are written in Fortran IV, G level and tested using OS/360. Output records up to 132 characters are produced. No special equipment is required.

(Please attach additional pages if necessary) . . . . **Total pages attached** ______

Permission to Publish

"I hereby give anyone permission to reprint, reproduce, and distribute this program to anyone else."

(18) **Signature of Submitter and Date** [signature] 6-23-69 [signature]

(19) **Signature of Installation Addressee** [signature]

T4SF

Acknowledgments

This program was written in the course of research supported by Grant GM-11546 of the National Institutes of Health. Computing time was provided by Grant FR-00011 of the NIH.

We thank the Australian Journal of Biological Science for permission to use the data from Dr. B. Griffing's paper in our example 2.

Table of Contents

Deck Key Sheet

Deck # 1 - 1	Source deck, main program DIALL; sequence DIAL 10 through DIAL5270 in cc 73-80; 530 cards.
Deck # 1 - 2	Source deck, subroutine DATINT; sequence DATI 10 through DATI 800 in cc 73-80; 81 cards.
Deck # 1 - 3	Source deck, subroutine GOTOIT; sequence GOTO 10 through GOTO 109 in cc 73-80; 109 cards.
Deck # 1 - 4	Source deck, subroutine COVRED; sequence COVR 10 through COVR 530 in cc 73-80; 53 cards.
Deck # 1 - 5	Source deck, subroutine EXPMSQ; sequence EXPM 10 through EXPM 560 in cc 73-80; 56 cards.
Deck # 1 - 6	Source deck, subroutine ANOVSS; sequence ANOV 10 through ANOV 440 in cc 73-80; 44 cards.
Deck # 2	Example input; sequence DATA 1 through DATA 108 in cc 73-80; Example 1: DATA 1 Through DATA 68; Example 2: DATA 69 Through DATA 108; 108 cards.
Deck # 3 - 1	Source deck, main program DIALLC; sequence DILC 10 through DILC1400 in cc 73-80; 140 cards.
Deck # 3 - 2	Source deck, subroutine MATINV; sequence MATI 10 through MATI 660 in cc 73-80; 66 cards.
Deck # 4	Program statements; Changes to DIALL for direct use with DIALLC; sequence numbers in cc 73-80 correspond to the placement of the cards in DIALL; 21 cards.

Introduction

The diallel cross is an experimental design used by geneticists and plant and animal breeders. The computing formulas for the analysis of several balanced versions of the diallel cross are given by Griffing (1956). When unbalance occurs the exact analysis is sufficiently difficult that it is common to use an analysis of unweighted means, or, in the case of missing observations (cells with no observations), to substitute in "estimated" values and then use the formulas for a balanced design.

According to the theory of general least squares, an analysis of variance exists for any set of data, balanced or not, and can be calculated exactly. For the general unbalanced case, the computation formulas, including those for the expectations of the mean squares, are so extensive and involve such a large number of computations as to make them practical only when used on a high speed digital computer.

In this report we present these formulas written in the form of a FORTRAN program which can be used directly on most large modern computers (possibly with minor changes required). This report includes an explanation of the preparation of data for the program, and a discussion of the interpretation of the analyses of variance and cross-products output by the program. Two complete examples are given.

The model analyzed includes as sources of variation, locations, replications, general combining ability, selfs, specific combining ability, maternal, reciprocal and general combining ability by location interaction. Any of these source(s) may be omitted from the analysis except for the general and specific combining abilities.

Given the model for a diallel experiment,

$$Y_{hijkn} = \mu + \ell_h + b_{i(h)} + g_j + g_k + s_{jk} + m_j + r_{jk} + (\ell g)_{hj} + (\ell g)_{hk} + e_{hijkn}$$

where

Y_{hijkn}	:	response or character measured
μ	:	population mean
ℓ_h	:	effect of the h^{th} location, $h = 1, \ldots, \ell$
$b_{i(h)}$	:	added effect of the i^{th} replication in the h^{th} location, $i = 1, \ldots, b$
g_j (g_k)	:	added effect of the j^{th} (k^{th}) line, i.e. general combining ability effect for the j^{th} female (k^{th} male) line, $j = 1, \ldots, g_1$; $k = 1, \ldots, g_2$
s_{jk}	:	specific combining ability effect for the cross between the j^{th} and the k^{th} lines, $s_{jk} = s_{kj}$
m_j	:	maternal effect of the j^{th} female line
r_{jk}	:	reciprocal effect involving the reciprocal crosses of the j^{th} and k^{th} lines, $r_{jk} = -r_{kj}$
$(\ell g)_{hj}, (\ell g)_{hk}$	:	location by general combining ability interaction involving the h^{th} location and the j^{th} and k^{th} lines
e_{hijkn}	:	random element associated with the n^{th} individual observation, uncorrelated with mean 0, variance σ_e^2,

a generalized least square analysis (based on all the observations) will be performed, minimizing

$$\Sigma e^2_{hijkn} = \Sigma (Y_{hijkn} - \hat{Y}_{hijkn})^2 \quad .$$

Any configuration of data will be analyzed, regardless of missing cells, unbalance, etc. The resulting analysis of variance table gives the Source, d.f., SS and MS adjusted only for the previously occurring effects. If an analysis based on cell means is desired, the cell means must be used as data.

The program provides options for specifying which sources are to be included in the model, for controlling the amount of output, for including more than one response variable, for doing crossproducts analyses, and for reading data and control cards from different input units.

II. Input Preparation

Input Item No. 1: Control Card

Card Columns	Fortran Labels	Meaning
1-3	NL	Number of Lines. (A line may be used as a male, a female or both.) Must be given.
4-6	NY	Number of Y variables to be analyzed, Standard number of variables = 1
7-9	CPD	Standard - analyses of variance only, if cross-products desired = 1
10-12	SELF	Selfed effects (the s_{jj}) - Standard - treated as specific, if desired to separate them from specific = 1, if desired to omit data on selfs even if in input = -1
13-15	MATE	Standard - maternal effects omitted from the model, if maternal effects desired = 1
16-18	RECP	Standard - reciprocal effects omitted from the model, if reciprocal effects desired = 1
19-21	REP	Number of replications per location, Standard - one replication, see Note 1 below.
22-24	LOC	Number of locations, Standard - one location, see Note 1 below
25-27	GXL	Standard - General C.A. by Location effects omitted from the model, if desired = 1
28-30	INUNIT	Fortran unit reference number for input data, Standard - Unit No. 1 (see Note 2 below)
31-33	ALLOUT	Standard - X, X'X, X'Y matrices not printed, if these matrices are desired = -1. This option involves a lot of output and is not necessary except for troubleshooting.

Fields may be left blank or set to zero if options labeled standard are desired. If all options are left blank a model containing only general and specific combining ability effects will be used.

Notes: 1. If REP (LOC) is specified as 1, replication (location) identification will be read from the data cards and thus must be specified in the input format, even though it will not be used in the P vector or X'X matrix. Thus if there is only one replication (location) and no replication (location) identifications are to be read, the REP (LOC) field should be left blank or set to zero. The same rep names must be used within each location. If location 1 has replication names R1, R2, R3 then location 2 should have replication names R1, R2, etc. Different numbers of replications are allowed within locations and then REP = largest number of replications in any location. (If different replication names are used in different locations, REP must be set to the total number of different replication names, although storage is likely to be exceeded in DATINT or COVRED, giving an error message.)

2. Changing the input unit: Except for the data (which can be read from any unit, parameter INUNIT above), all input is read on unit number 1. Thus if control information is to be read from another input unit the JCL should be changed to redefine unit 1, or the READ statements, cards DIAL 830, DIAL2730, and DIAL3030, should be changed. Output records are up to 132 characters and all input is on unit reference number 3.

<u>Input Item No. 2</u>: Variable Names

This item gives the names of the response variables (characters to be analyzed), each may be up to 16 alphameric characters in length. The variable names should be in the same order that the variables are specified in the input format, with four names per card starting in columns 1, 21, 41, 61. These names are only used by the program to label the output.

<u>Input Item No. 3</u>: Data Format

A single card containing the FORTRAN format by which the data will be read. Column 1 contains the opening parenthesis of the format and the closing parentheses must be in or before column 80. The order of the READ statement is:

Location, if present (LOC > 0),*

Replication, if present (REP > 0),*

Female line, Male line } must have identical format specification*

Response variable(s) (the observations).

*maximum of eight columns of alphameric data.

In this format statement, 'A' format specification must be used for location, replication, and the two line identifiers. Up to A8 format specification can be used for each identifier. The response variable(s) should be read with real 'E', 'F' format specification. Note that the order of reading the data card can be transposed by use of the 'T' format specification. If a format containing more than 75 characters is needed, the dimension of the format vector, FMT (card DIAL 310) must be increased and the DO loop in the READ statement (card DIAL3030) must be adjusted.

<u>Input Item No. 4</u>: Data

Data cards, in any order.

<u>Input Item No. 5</u>: End Card

Blank card. At the end of the data, the program will complete the analyses when the end of file condition is raised, if a blank card is missing. For multiple runs, repeat items 1 through 5.

III. Output Interpretation

A. Minimum Output

1. The first page printed is the title page which gives the information read from the control card, the sources included in the analysis, and the names of the variables to be analyzed.
2. The second page lists the coefficients of the expected mean squares for the sources of variation included in the analysis except for the error term, the error term coefficient always being equal to unity.

   The coefficients are presented in the customary order for an analysis of variance. The order is the same as the sources of variation listed at the left-hand side of the page, starting with the last source and proceeding to the mean. For an analysis with sources, MEAN, REPS, GCA, SCA, the expected mean square for the mean line would be

$$c_{11}\ \sigma^2_{SCA} + c_{12}\ \sigma^2_{GCA} + c_{13}\ \sigma^2_{REPS} + c_{14}\ \mu^2 \quad .$$

The second column from the left in the table lists the degrees of freedom for the various sources of variation. Therefore the table for the above analysis would appear as:

SOURCE	DF	COEFFICIENTS			
MEAN	1	c_{11}	c_{12}	c_{13}	c_{14}
REPS	3	c_{21}	c_{22}	c_{23}	c_{24}
GCA	3	c_{31}	c_{32}	c_{33}	c_{34}
SCA	6	c_{41}	c_{42}	c_{43}	c_{44}

If more than six sources of variation are included, the coefficients for the components corresponding to the last six sources listed will be on the first line and the remaining coefficients on the second line for this source. The last coefficient for any source is always the coefficient for μ^2.

3. The analyses of variance and crossproducts tables are printed starting on the third page, with each analysis beginning on a new page. The sources of variation will include the mean, the error and the total. The degrees of freedom, sum of squares and mean square for each line will be given (the mean square for the total is omitted). Below the ANOVA table a summary including the total number of effects started with, the degrees of freedom, the number of dependencies found in the X'X matrix (the sum of the last two equaling the first), and the sum of squares accounted for by the least squares fitting of the effects. The error mean square will contain all interactions involving replications and locations, except for general combining ability by location interaction, and any within cross variation.

B. Additional Output

If the parameter for additional output, ALLOUT, is negative, the X, X'Y, and X'X matrices are printed, each starting on a new page, after the title page and before the coefficients of the expected mean squares.

1. For the X matrix, the female and male line numbers (in integer form numbered in the order encountered by the program) are printed followed by the location in X for each source of variation included, which is followed by the value put in the X matrix for each effect. The first column is always the mean column.
2. The X'Y vector for each variable is printed, ten elements per line, in the order of the individual effects in the X'X matrix, starting with the mean.
3. The X'X matrix is printed, 25 columns at a time, with a label for each row and column. Only the upper right triangular part, including the main diagonal, is meaningful, even though a square matrix is printed.

Because of the number of pages printed, and because it is usually of no benefit, it is strongly recommended that this option not be used except for troubleshooting difficult problems.

C. Error Messages

1. 'MORE EFFECTS ARE INITIALIZED THAN PROVIDED FOR IN THE DIMENSION STATEMENTS, SDFL = ___, AND SHOULD BE < = ___, JOB TERMINATED.' The first blank is filled with the number of effects initialized. In the program ISDFL = 250 and if more than 250 effects are initialized the above message is printed. This message indicates that either the parameters on the first input card are incorrect or more than 250 effects are in the model to be analyzed. If the former, correct the parameter card. If the latter situation, then the dimensions must be increased and ISDFL (card DIAL 360) changed (see Section V.B.).
2. 'THE NO. OF RESPONSE VARIABLES SPECIFIED, ___, IS GREATER THAN THE DIMENSION OF Y; WHICH IS, ___, JOB TERMINATED. The first blank is filled with the number of response variables specified on the control card and the second is filled with the value of INY (card DIAL 370). This message indicates either an error in the parameter card or that more variables are to be read than allowed by the dimensions (see Section V.B.).
3. 'TOO MANY ___ NAMES ENCOUNTERED, JOB TERMINATED.' This message is printed by SUBROUTINE DATINT; the blank will be filled with LOCATION, REPLICATION, or LINE and indicates that the number specified for this effect on the parameter card is incorrect or too large or that there is an error in the names on the data cards. The dimensions in DATINT allow a maximum of 30 different names of any one type. This may be changed, see Section V.B.

IV. Examples

A. Description

1. Cards DATA 10 through DATA 68. This problem shows a complete diallel analysis. Since the first column of data is read twice, the first and third variables are the same. This gives only two distinct analyses of variance and one distinct analysis of crossproducts for the six tables printed.

2. Cards DATA 690 through DATA1080. This problem is given by B. Griffing (1956). Since only the means over plots and reps are given as data, there are no degrees of freedom or sums of squares associated with error.

B. Input for examples

```
 4  3  1  1  1  1  2  2  1  1  1                                          DATA0010
    TRAIT 1            TRAIT 2                 TRAIT 3                    DATA0020
(2A1,A3,2X,A3,2F10.0, T11, F10.0)                                         DATA0030
11   1    1          1.        5.                                         DATA0040
11   1    2          2.        2.0                                        DATA0050
11   1    3          3.        10.                                        DATA0060
11   1    4          4.        8.                                         DATA0070
11   2    1          5.        6.                                         DATA0080
11   2    2          6.        9.                                         DATA0090
11   2    3          7.        7.                                         DATA0100
11   2    4          8.        8.                                         DATA0110
11   3    1          9.        9.                                         DATA0120
11   3    2         10.        10.                                        DATA0130
11   3    3          9.        09.                                        DATA0140
11   3    4          8.        8.                                         DATA0150
11   4    1          7.        07.                                        DATA0160
11   4    2          6.        06.                                        DATA0170
11   4    3          5.        09.                                        DATA0180
11   4    4          4.        04.                                        DATA0190
12   1    1          1.        01.                                        DATA0200
12   1    2          2.        02.                                        DATA0210
12   1    3          3.        03.                                        DATA0220
12   1    4          4.        04.                                        DATA0230
12   2    1          5.        09.                                        DATA0240
12   2    2          6.        06.                                        DATA0250
12   2    3          7.        07.                                        DATA0260
12   2    4          8.        09.                                        DATA0270
12   3    1          9.        09.                                        DATA0280
12   3    2         10.        09.                                        DATA0290
12   3    3          9.        09.                                        DATA0300
12   3    4          8.        07.                                        DATA0310
12   4    1          7.        06.                                        DATA0320
12   4    2          6.        06.                                        DATA0330
12   4    3          5.        05.                                        DATA0340
12   4    4          4.        04.                                        DATA0350
21   1    1          1.        03.                                        DATA0360
21   1    2          2.        03.                                        DATA0370
21   1    3          3.        05.                                        DATA0380
21   1    4          4.        06.                                        DATA0390
21   2    1          5.        07.                                        DATA0400
21   2    2          6.        08.                                        DATA0410
21   2    3          7.        09.                                        DATA0420
21   2    4          8.        10.                                        DATA0430
21   3    1          9.        01.                                        DATA0440
21   3    2         10.        02.                                        DATA0450
21   3    3          9.        03.                                        DATA0460
21   3    4          8.        04.                                        DATA0470
21   4    1          7.        05.                                        DATA0480
21   4    2          6.        06.                                        DATA0490
21   4    3          5.        07.                                        DATA0500
```



```
21   4    4          8.        08.                                        DATA0510
22   1    1          1.        09.                                        DATA0520
22   1    2          2.        08.                                        DATA0530
22   1    3          3.        07.                                        DATA0540
22   1    4          4.        06.                                        DATA0550
22   2    1          5.        05.                                        DATA0560
22   2    2          6.        04.                                        DATA0570
22   2    3          7.        01.                                        DATA0580
22   2    4          8.        10.                                        DATA0590
22   3    1          9.        02.                                        DATA0600
22   3    2         10.        01.                                        DATA0610
22   3    3          9.        07.                                        DATA0620
22   3    4          8.        09.                                        DATA0630
22   4    1          9.        01.                                        DATA0640
22   4    2          6.        05.                                        DATA0650
22   4    3          5.        06.                                        DATA0660
22   4    4          8.        04.                                        DATA0670
                                                                          DATA0680
  9  3  1                                                                 DATA0690
  TOTAL YIELD          COB WEIGHT          SHELLED CORN WT                DATA0700
(2A3,4X,3F10.0)                                                           DATA0710
  1  2      240.0      31.8      208.2                                    DATA0720
  1  3      260.0      34.7      225.3                                    DATA0730
  1  4      230.4      32.3      198.1                                    DATA0740
  1  5      257.0      45.0      212.0                                    DATA0750
  1  6      241.5      39.0      202.5                                    DATA0760
  1  7      266.9      35.1      231.8                                    DATA0770
  1  8      240.1      35.7      204.4                                    DATA0780
  1  9      300.4      40.1      260.3                                    DATA0790
  2  3      209.0      27.9      181.1                                    DATA0800
  2  4      217.3      30.8      186.5                                    DATA0810
  2  5      233.1      39.6      193.5                                    DATA0820
  2  6      229.5      33.1      196.4                                    DATA0830
  2  7      266.9      30.9      236.0                                    DATA0840
  2  8      216.3      30.9      185.4                                    DATA0850
  2  9      214.2      28.5      185.7                                    DATA0860
  3  4      183.7      25.2      158.5                                    DATA0870
  3  5      253.7      41.4      212.3                                    DATA0880
  3  6      250.1      35.5      214.6                                    DATA0890
  3  7      268.8      34.9      233.9                                    DATA0900
  3  8      222.3      32.1      190.2                                    DATA0910
  3  9      252.1      32.4      219.7                                    DATA0920
  4  5      233.8      42.6      191.2                                    DATA0930
  4  6      213.7      35.7      178.0                                    DATA0940
  4  7      255.7      35.6      220.1                                    DATA0950
  4  8      197.4      32.7      164.7                                    DATA0960
  4  9      281.0      41.3      239.7                                    DATA0970
  5  6      206.8      40.1      166.7                                    DATA0980
  5  7      272.2      43.6      228.6                                    DATA0990
  5  8      242.9      41.8      201.1                                    DATA1000
  5  9      260.8      44.2      216.6                                    DATA1010
  6  7      261.8      39.1      222.7                                    DATA1020
  6  8      270.3      43.5      226.8                                    DATA1030
  6  9      283.9      41.5      242.4                                    DATA1040
  7  8      273.2      38.3      234.9                                    DATA1050
  7  9      302.2      41.1      261.1                                    DATA1060
  8  9      259.8      35.2      224.6                                    DATA1070
                                                                          DATA1080
```

C. Output of examples

```
**** G E N E R A L   L E A S T   S Q U A R E S   D I A L L E L   A N A L Y S I S ****

                    H. E. SCHAFFER  AND  R. A. USANIS
                        NCSU, DEPT.OF GENETICS
                            NOVEMBER,1968

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

          NUMBER OF LOCATIONS                         =    2
          NUMBER OF REPLICATIONS                      =    2
          MAXIMUM NUMBER OF MALE - FEMALE LINES       =    4
          NUMBER OF RESPONSE VARIABLES                =    3
          ANALYSIS OF CROSSPRODUCTS DESIRED           -   YES
          EFFECTS ACCOUNTED FOR ARE :
                                           MEAN
                                           LOC
                                           REPS
                                           GCA
                                           SELF
                                           SCA
                                           MATR
                                           RECP
                                           GXL

          VARIABLES ANALYZED ARE :          TRAIT 1
                                            TRAIT 2
                                            TRAIT 3

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
```


E X P E C T A T I O N S   O F   M E A N   S Q U A R E S
A N D   M E A N   P R O D U C T S

F O R   C H A R A C T E R S

TRAIT 1
TRAIT 2
TRAIT 3

SOURCE	DF	COEFFICIENTS					
MEAN	1	32.000000000	0.0	0.0	6.0000000000	1.0000000000	64.000000000
		16.000000000	32.000000000	64.000000000			
LOC	1	32.000000000	0.0	0.0	0.0	0.0	0.0
		16.000000000	32.000000000	0.0			
REPS	2	0.0	0.0	0.0	0.0	0.0	0.0
		16.000000000	0.0	0.0			
GCA	3	16.000000000	0.0	0.0	4.0000000000	2.0000000000	32.000000000
		0.0	0.0	0.0			
SELF	4	0.0	0.0	0.0	3.5000000000	2.2500000000	0.0
		0.0	0.0	0.0			
SCA	2	0.0	0.0	0.0	8.0000000000	0.0	0.0
		0.0	0.0	0.0			
MATR	3	0.0	8.0000000000	32.000000000	0.0	0.0	0.0
		0.0	0.0	0.0			
RECP	3	0.0	8.0000000000	0.0	0.0	0.0	0.0
		0.0	0.0	0.0			
GXL	3	16.000000000	0.0	0.0	0.0	0.0	0.0
		0.0	0.0	0.0			

A N A L Y S I S   O F   V A R I A N C E

TRAIT 1

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES
MEAN	1	2413.26562500	2413.26562500
LOC	1	0.140625000000	0.140625000000
REPS	2	1.65625000000	0.828125000000
GCA	3	176.593750000	58.8645833333
SELF	4	19.6406250000	4.91015625000
SCA	2	45.5000000000	22.7500000000
MATR	3	145.625000000	48.5416666667
RECP	3	2.12500000000	0.708333333333
GXL	3	5.84375000000	1.94791666667
ERROR	42	34.6093750000	0.824032738095
TOTAL	64	2845.00000000	

S U M M A R Y

39 EFFECTS 22 DF
17 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS 2810.39062500

A N A L Y S I S   O F   C R O S S   P R O D U C T S

TRAIT 1 - BY - TRAIT 2

SOURCE	DF	SUM OF PRODUCTS	MEAN PRODUCTS
MEAN	1	2333.43750000	2333.43750000
LOC	1	-2.15625000000	-2.15625000000
REPS	2	4.15625000000	2.07812500000
GCA	3	51.2500000000	17.0833333333
SELF	4	5.31250000000	1.32812500000
SCA	2	-7.87500000000	-3.93750000000
MATR	3	13.5625000000	4.52083333333
RECP	3	2.56250000000	0.854166666667
GXL	3	8.06250000000	2.68750000000
ERROR	42	-2.31250000000	-0.550595238095D-01
TOTAL	64	2406.00000000	

S U M M A R Y

39 EFFECTS 22 DF
17 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS 2408.31250000

A N A L Y S I S   O F   C R O S S   P R O D U C T S

TRAIT 1 - BY - TRAIT 3

SOURCE	DF	SUM OF PRODUCTS	MEAN PRODUCTS
MEAN	1	2413.26562500	2413.26562500
LOC	1	0.140625000000	0.140625000000
REPS	2	1.65625000000	0.828125000000
GCA	3	176.593750000	58.8645833333
SELF	4	19.6406250000	4.91015625000
SCA	2	45.5000000000	22.7500000000
MATR	3	145.625000000	48.5416666667
RECP	3	2.12500000000	0.708333333333
GXL	3	5.84375000000	1.94791666667
ERROR	42	34.6093750000	0.824032738095
TOTAL	64	2845.00000000	

S U M M A R Y

39 EFFECTS 22 DF
17 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS 2810.39062500

A N A L Y S I S   O F   V A R I A N C E

TRAIT 2

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES
MEAN	1	2256.25000000	2256.25000000
LOC	1	33.0625000000	33.0625000000
REPS	2	15.3125000000	7.65625000000
GCA	3	23.3125000000	7.77083333333
SELF	4	15.6458333333	3.91145833333
SCA	2	9.04166666667	4.52083333333
MATR	3	38.0625000000	12.6875000000
RECP	3	10.1875000000	3.39583333333
GXL	3	79.9375000000	26.6458333333
ERROR	42	249.187500000	5.93303571429
TOTAL	64	2730.00000000	

S U M M A R Y

39 EFFECTS 22 DF
17 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS 2480.81250000

A N A L Y S I S   O F   C R O S S   P R O D U C T S

TRAIT 2   - BY -   TRAIT 3

SOURCE	DF	SUM OF PRODUCTS	MEAN PRODUCTS
MEAN	1	2333.43750000	2333.43750000
LOC	1	-2.15625000000	-2.15625000000
REPS	2	4.15625000000	2.07812500000
GCA	3	51.2500000000	17.0833333333
SELF	4	5.31250000000	1.32812500000
SCA	2	-7.87500000000	-3.93750000000
MATR	3	13.5625000000	4.52083333333
RECP	3	2.56250000000	0.854166666667
GXL	3	8.06250000000	2.68750000000
ERROR	42	-2.31250000000	-0.5505952380950-01
TOTAL	64	2406.00000000	

S U M M A R Y

39 EFFECTS   22 DF
17 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS   2408.31250000

A N A L Y S I S   O F   V A R I A N C E

TRAIT 3

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES
MEAN	1	2413.26562500	2413.26562500
LOC	1	0.140625000000	0.140625000000
REPS	2	1.65625000000	0.828125000000
GCA	3	176.593750000	58.8645833333
SELF	4	19.6406250000	4.91015625000
SCA	2	45.5000000000	22.7500000000
MATR	3	145.625000000	48.5416666667
RECP	3	2.12500000000	0.708333333333
GXL	3	5.84375000000	1.94791666667
ERROR	42	34.6093750000	0.824032738095
TOTAL	64	2845.00000000	

S U M M A R Y

39 EFFECTS   22 DF
17 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS   2810.39062500

**** G E N E R A L   L E A S T   S Q U A R E S   D I A L L E L   A N A L Y S I S ****

H. E. SCHAFFER AND R. A. USANIS
NCSU, DEPT.OF GENETICS
NOVEMBER,1968

* *

MAXIMUM NUMBER OF MALE - FEMALE LINES = 9
NUMBER OF RESPONSE VARIABLES = 3
ANALYSIS OF CROSSPRODUCTS DESIRED - YES
EFFECTS ACCOUNTED FOR ARE :
MEAN
GCA
SCA

VARIABLES ANALYZED ARE : TOTAL YIELD
COB WEIGHT
SHELLED CORN WT

* *

E X P E C T A T I O N S   O F   M E A N   S Q U A R E S
A N D   M E A N   P R O D U C T S

F O R   C H A R A C T E R S

TOTAL YIELD
COB WEIGHT
SHELLED CORN WT

SOURCE	DF	COEFFICIENTS		
MEAN	1	1.0000000000	16.000000000	36.000000000
GCA	8	1.0000000000	7.0000000000	0.0
SCA	27	1.0000000000	0.0	0.0

A N A L Y S I S   O F   V A R I A N C E

TOTAL YIELD

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES
MEAN	1	2184877.34065	2184877.34065
GCA	8	18605.9200847	2325.74001059
SCA	27	9164.84224087	339.438601514
ERROR	0	0.0	0.0
TOTAL	36	2212647.68359	

S U M M A R Y

55 EFFECTS 36 DF
19 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS 2212648.10298

A N A L Y S I S   O F   C R O S S   P R O D U C T S

TOTAL YIELD - BY - COB WEIGHT

SOURCE	DF	SUM OF PRODUCTS	MEAN PRODUCTS
MEAN	1	323513.998415	323513.998415
GCA	8	1872.27852225	234.034815281
SCA	27	1160.08205789	42.9660021439
ERROR	0	0.0	0.0
TOTAL	36	326546.285156	

S U M M A R Y

55 EFFECTS 36 DF
19 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS 326546.358995

A N A L Y S I S   O F   C R O S S   P R O D U C T S

TOTAL YIELD - BY - SHELLED CORN WT

SOURCE	DF	SUM OF PRODUCTS	MEAN PRODUCTS
MEAN	1	1861363.56026	1861363.56026
GCA	8	16733.6550860	2091.70688574
SCA	27	8004.76007302	296.472595297
ERROR	0	0.0	0.0
TOTAL	36	1886101.79687	

S U M M A R Y

55 EFFECTS 36 DF
19 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS 1886101.97542

A N A L Y S I S   O F   V A R I A N C E

COB WEIGHT

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES
MEAN	1	47902.6008569	47902.6008569
GCA	8	727.848006592	90.9810008240
SCA	27	188.674219361	6.98793405040
ERROR	0	0.0	0.0
TOTAL	36	48819.1193848	

S U M M A R Y

55 EFFECTS 36 DF
19 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS 48819.1230829

A N A L Y S I S   O F   C R O S S   P R O D U C T S

COB WEIGHT   - BY -   SHELLED CORN WT

SOURCE	DF	SUM OF PRODUCTS	MEAN PRODUCTS
MEAN	1	275611.429841	275611.429841
GCA	8	1144.43278731	143.054098413
SCA	27	971.407588502	35.9780588334
ERROR	0	0.0	0.0
TOTAL	36	277727.215088	

S U M M A R Y

55 EFFECTS   36 DF
19 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS   277727.270217

A N A L Y S I S   O F   V A R I A N C E

SHELLED CORN WT

SOURCE	DF	SUM OF SQUARES	MEAN SQUARES
MEAN	1	1585752.31617	1585752.31617
GCA	8	15589.2335505	1948.65419382
SCA	27	7033.35262472	260.494541656
ERROR	0	0.0	0.0
TOTAL	36	1608374.78125	

S U M M A R Y

55 EFFECTS   36 DF
19 DEPENDENCIES
TOTAL SS ACCOUNTED FOR BY EFFECTS   1608374.90234

* * * * T H A T ' S   I T   F O L K S * * * *

V. Procedure

A. General

The main program, DIALL, develops the X'X matrix (matrix of coefficients of the normal equations) in upper triangular form, the X'Y vectors (right-hand sides of the normal equations), TSS (total sums of squares and crossproducts of the observed variables) and a vector (P) of effect names from the control card and data read. The control card is read first and the P vector, which contains a name for each single effect in the X'X matrix, is filled in. Then the response variable names and the format for the data are read. The data are read one observation at a time with the identifications read into REAL*8 variables (allowing up to 8 alphameric characters per identifier) and the response variables read into the REAL*4 vector Y. SUBROUTINE DATINT, ENTRY DATCNV is then called to convert the alphameric identification names into three consecutive sequences of integers, each beginning with 1, for the location, replication and line numbers. Next the X'X, X'Y, and TSS matrices are developed. All matrices are stored in vector form.

In addition, the main program writes the title page and, if desired, the X matrix. When the matrices are complete SUBROUTINE GOTOIT is called, which then calls the remaining subroutines. GOTOIT writes the headings for the coefficients of the expected mean squares, the analyses of variance and, if specified on the control card, prints the unreduced X'X and X'Y matrices. It also calculates the locations for the responses variables in the X'Y vector, completes the error line in the ANOVA table and prints a summary table.

The first call in GOTOIT is to SUBROUTINE COVRED which reduces the X'X and X'Y matrices in place, leaving in X'X the geometric means of the A-B rows of the forward Doolittle (reduced X'X). For details see Schaffer (in preparation). Next a call to SUBROUTINE EXPMSQ is made which, using the reduced X'X matrix, computes and prints the coefficients of the expected mean squares assuming a random model (omitting the error component which has coefficient

= 1). Finally a call is made to SUBROUTINE ANOVSS for each analysis of variance and crossproducts to be computed. This subroutine uses the reduced X'Y matrix and calculates and prints the source, d.f., sums of squares, and mean square for each line of the analysis.

Even if more than one response variable is being analyzed in one run (a control card with NY > 1), only one X'X matrix is developed. Different missing observations in the different response variables will cause the analysis of variance (and crossproducts) to be incorrect except for those variables that have all observations on the data cards that are read. Missing observations are treated correctly only when the observations are missing from the input data stream, not when they are read as a zero or blank.

B. Changing the Capacity of the Program

The dimensions given are sufficient for a model with 250 effects. Omission of any or all of the LOC, REP, SELF, MATERNAL, RECIPROCAL or GENERAL by LOCATION INTERACTION sources allows for a larger number of male - female lines to be analyzed for general and specific combining ability.

Sources	Number of Effects
Mean	1
Location	LOC
Replications	REP*LOC
General C. A.	NL
Selfs	NL
Specific C. A.	(NL*(NL-1))/2
Maternal	NL
Reciprocal	(NL*(NL-1))/2
G.C.A. by Location	(LOC*NL)

If less than 250 effects are in the model the memory required by the program can be decreased by decreasing the dimensions of XPY, XPX, and P (cards DIAL 290, DIAL 330). In addition, the repetition factor for filling P with blanks should be decreased to 1 less than the dimension of P and the error check ISDFL should be decreased to these new dimensions. (See Section VIII. A. for vector dimensions.)

The number of response variables is 10 with the present dimensions. To decrease the memory requirements for response variables, vectors TSS, XPY, LABP and Y should be decreased (cards DIAL 290, DIAL 310). Also the response variable error check constant, INY, should be decreased to be equal to the dimension of Y. To increase the capacity of the program these changes must be made in the opposite direction.

Approximately 2300_{10} bytes of memory can be saved by removing SUBROUTINE DATINT. This subroutine is only used to convert the identifications for the observations from alphameric to integers. Its removal means that the identifiers must be integers, since they are used as subscripts in the program. When the subroutine is not used, the locations, replications, and lines must each be numbered from 1 to the highest number of units within each classification. Also the data format, input item number 3, should then have an 'I' format specification for these identifiers. Removal of the subroutine is accomplished by making the following changes in the program:

1. Removing the call statements to it, cards DIAL 940 and DIAL3390.
2. Removing the G in front of the identification variables in the 4 read statements, cards DIAL3250, DIAL3270, DIAL3290, and DIAL3310.
3. Changing the variable BL in card DIAL3380 to 0 (zero). This variable is used to check if a blank card has been read to end the set of data.

In addition the REAL*8 statement can be changed, removing the comma from the first card (DIAL 290) and deleting the second card (DIAL 300).

To increase the capacity of DATINT, the 30 on cards DATI 220, DATI 230 should be increased.

With some compilers it will be necessary to change the dimension statements in the subroutines when the dimension statements in the main program (DIALL) are changed.

Because the X'X matrix can become large, all calculations are done in double precision to reduce errors due to "rounding". If the program is converted to single precision all REAL*8 statements should be changed to REAL*4 and the calculations of EPS (card DIAL5170) must be changed. Changing to single precision is <u>not recommended</u> because of the increased error. If this change must be made the input data used should be coded to a mean of approximately zero.

The program requires (including OS/360 routines) approximately:

$$33{,}000 + 4\cdot(\text{SDFL}\cdot(\text{SDFL} + 1)) + 8\cdot\text{NY}\cdot\text{SDFL} + 4(\text{NY}\cdot(\text{NY} + 1)) \text{ bytes}$$

where $\text{SDFL} = 1 + 2\text{NL} + \text{NL}^2 + \text{LOC}(1 + \text{REP} + \text{NL})$, and

a double precision word is eight bytes long.

The present dimensions in the program are sufficient for SDFL = 250, and NY = 10. Approximately 304K bytes of memory are required for program and storage. Execution time for a 9 line diallel with 4 reps and 1400 observations was approximately 45 seconds on an IBM 360/75.

For a complete analysis of a 5 line diallel experiment with two replicates and one variable approximately 42K bytes of memory are needed. The memory requirements increase more rapidly than an increase in the number of lines.

The program is written in FORTRAN IV, G level, and tested using OS/360. Output records up to 132 characters are produced. No special equipment is required.

VI. Estimation of Variance Components, Standard Deviations, and Correlations

A. Introduction

This main program, DIALLC, computes the variance components, their standard deviations, and the correlations between the effects for different traits when crossproduct analyses were computed. The input for this program is computed by the program DIALL.

B. Input

1. General information

The first read statement (card DILC 180) of this program reads the number of variables analyzed, a code telling if crossproduct analyses were computed, and the number of lines in the analysis excluding the error line. Next the names, degrees of freedom and coefficients of the expected mean squares for each source in the analysis are read (card DILC 220). Then, for each analysis of variance or crossproducts that was computed, the name of the variable(s) used (card DILC 350), the computed mean squares and the error degrees of freedom are read (cards DILC 370, DILC 390). When all the input for a set of data has been read the variance components, standard deviation, and correlations are computed. If more than one set of data was analyzed, the above sequence is repeated. All input is read from FORTRAN Unit Number 1.

2. Changing DIALL

Although DIALL produces all the above information, not all of it is stored or saved because of the additional memory that it would require. However, if the cards listed below are added to DIALL the input needed for this program will be produced in proper form (on cards or disk) for input for DIALLC. On these cards the X in column 73 indicates that it is a card added to DIALL. The next three columns are the first three letters of the name of the section in DIALL (main program or subroutines) where they are to be inserted. Columns 77-79 designate the cards which they follow in their respective sections. The last column gives the order of the addition cards. Thus card XDIA2675 is inserted directly after DIAL2670 and is followed by DIAL2680. The addition of these cards, therefore, does not disturb the sequence numbering in columns 77-80.

Cards to be inserted in DIALL if input to DIALLC is wanted.

```
C     CARDS WITH X IN COLUMN 73 ARE ADDED TO THE MAIN PGM DIALL AND      XDIA0104
C          ITS SUBROUTINES BY THE CODES IN COLUMNS 74 - 80.             XDIA0105
C                                                                        XDIA0106
      NAL = 1                                                            XDIA2645
      NAL = NAL + 1                                                      XDIA2675
      WRITE (4,44) NY,CPD,NAL                                            XDIA2684
   44 FORMAT (3I4)                                                       XDIA2685
      WRITE (4,41) (LABP(J,I),J=1,2),(LABP(J,K),J=1,2)                   XGOT0815
   41 FORMAT (2A8,' - BY - ',2A8)                                        XGOT0816
      WRITE (4,42) (LABP(J,I),J=1,2)                                     XGOT0875
   42 FORMAT (12X,2A8)                                                   XGOT0876
      WRITE (4,43) EDF,ERMS                                              XGOT1015
   43 FORMAT (I4,A8)                                                     XGOT1016
      WRITE (4,45) CON,DF,(EXPMS(J),J=1,K)                               XEXP0525
   45 FORMAT (A4,I4,9A8)                                                 XEXP0526
      REAL*8 DXMS(9)                                                     XANO0055
      IC = 0                                                             XANO0205
      IC = IC + 1                                                        XANO0405
      DXMS(IC) = XMS                                                     XANO0406
      WRITE (4,40)  II,IJ,(DXMS(J),J=1,IC)                               XANO0415
   40 FORMAT (2I4,9A8)                                                   XANO0416
```


C. Output

For each set of data the first page of output will be the inverse of the expected mean squares coefficient matrix. Following this, the variance (and covariance) components for each variable will be printed. Each variable is begun on a new page. The names of the components, their estimated values and their estimated standard deviations are printed. Then, if crossproduct analyses were computed, the correlations between variables will be printed. If either variance component entering this correlation is less than or equal to zero eight asterisks are printed instead of the correlation. A message is also printed at the bottom of the table. The correlation table lists the variables across the page and the sources down the page. Correlations for up to six variables are printed across the page and up to 18 variables per page. All output is to FORTRAN Unit Number 3.

The output of this program, for the examples used in the program DIALL, is given below.

D. Procedure Used

The estimates of the variance (crossproduct) components are calculated by inverting the matrix of coefficients of the expected mean squares and multiplying the computed mean squares (products) for each variable by this inverse. The counter method of inversion by row operations is used. The standard deviation of the component is computed as

$$SDVC = \sqrt{\sum_i \frac{2\, a_i^2\, (MS_i)^2}{DF_i + 2}}$$

where the a_i are the coefficients of the linear combination of the mean squares (i.e. the rows of the inverse) used to estimate the component (see Anderson and Bancroft. 1952). The correlations are computed by taking the crossproduct component and dividing by the square root of the product of the respective variance components.

E. Output of examples

INVERSE OF THE EXPECTED MEAN SQUARES COEFFICIENTS MATRIX

	MEAN	LOC	REPS	GCA	SELF	SCA	MATR	RECP	GXL	ERR
MEAN	0.1562D-01	-0.1562D-01	0.0	-0.3125D-01	0.2083D-01	-0.5208D-02	0.0	0.0	0.3125D-01	-0.1562D-01
LOC	0.0	0.3125D-01	-0.3125D-01	0.0	0.0	0.0	0.0	0.0	-0.6250D-01	0.6250D-01
REPS	0.0	0.0	0.6250D-01	0.0	0.0	0.0	0.0	0.0	0.0	-0.6250D-01
GCA	0.0	0.0	0.0	0.3125D-01	-0.2778D-01	-0.3472D-02	0.0	0.0	-0.3125D-01	0.3125D-01
SELF	0.0	0.0	0.0	0.0	0.4444	-0.1944	0.0	0.0	0.0	-0.2500
SCA	0.0	0.0	0.0	0.0	0.0	0.1250	0.0	0.0	0.0	-0.1250
MATR	0.0	0.0	0.0	0.0	0.0	0.0	0.3125D-01	-0.3125D-01	0.0	[illegible]
RECP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1250	0.0	-0.1250
GXL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6250D-01	-0.6250D-01
ERR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000

VARIANCE COMPONENTS FOR

TRAIT 1

NAME	VALUE	STANDARD DEVIATION
LOC	-0.09172712	0.07998136
REPS	0.00025577	0.03820997
GCA	1.58901057	1.16806003
SELF	-2.44732763	3.37247125
SCA	2.74074591	2.01095482
MATR	1.49479167	0.95949106
RECP	-0.01446243	0.06015078
GXL	0.07024275	0.07777714
ERR	0.82403274	0.17568437

VARIANCE COMPONENTS FOR

TRAIT 1 - BY - TRAIT 2

NAME	VALUE	STANDARD DEVIATION
LOC	-0.30373419	0.12814677
REPS	0.13332403	0.09184395
GCA	0.42492870	0.34259121
SELF	1.36966766	0.63972029
SCA	-0.48530506	0.34803221
MATR	0.11458333	0.09093167
RECP	0.11365327	0.06754374
GXL	0.17140997	0.10623530
ERR	-0.05505952	0.01173873

VARIANCE COMPONENTS FOR

TRAIT 1 - BY - TRAIT 3

NAME	VALUE	STANDARD DEVIATION
LOC	-0.09172712	0.07998136
REPS	0.00025577	0.03820997
GCA	1.58901057	1.16806003
SELF	-2.44732763	3.37247125
SCA	2.74074591	2.01095482
MATR	1.49479167	0.95949106
RECP	-0.01446243	0.06015078
GXL	0.07024275	0.07777714
ERR	0.82403274	0.17568437

VARIANCE COMPONENTS FOR

TRAIT 2

NAME	VALUE	STANDARD DEVIATION
LOC	-0.50060454	1.36232172
REPS	0.10770089	0.34747484
GCA	-0.52878534	0.55367224
SELF	-0.62388393	1.22218769
SCA	-0.17652530	0.42973481
MATR	0.29036458	0.25958526
RECP	-0.31715030	0.31156648
GXL	1.29454985	1.05623191
ERR	5.93303571	1.26492746

VARIANCE COMPONENTS FOR

TRAIT 2 - BY - TRAIT 3

NAME	VALUE	STANDARD DEVIATION
LOC	-0.30373419	0.12814677
REPS	0.13332403	0.09184395
GCA	0.42492870	0.34259121
SELF	1.36966766	0.63972029
SCA	-0.48530506	0.34803221
MATR	0.11458333	0.09093167
RECP	0.11365327	0.06754374
GXL	0.17140997	0.10623530
ERR	-0.05505952	0.01173873

VARIANCE COMPONENTS FOR

TRAIT 3

NAME	VALUE	STANDARD DEVIATION
LOC	-0.09172712	0.07998136
REPS	0.00025577	0.03820997
GCA	1.58901057	1.16806003
SELF	-2.44732763	3.37247125
SCA	2.74074591	2.01095482
MATR	1.49479167	0.95949106
RECP	-0.01446243	0.06015078
GXL	0.07024275	0.07777714
ERR	0.82403274	0.17568437

CORRELATIONS COMPUTED FROM VARIANCE (COVARIANCE) COMPONENTS

TRAITS:	TRAIT 1 - BY - TRAIT 2	TRAIT 1 - BY - TRAIT 3	TRAIT 2 - BY - TRAIT 3
COEFFICIENT			
LOC	********	********	********
REPS	25.4025	1.0000	25.4025
GCA	********	1.0000	********
SELF	********	********	********
SCA	********	1.0000	********
MATR	0.1739	1.0000	0.1739
RECP	********	********	********
GXL	0.5684	1.0000	0.5684
ERR	-0.0249	1.0000	-0.0249

******** MEANS THAT THE VARIANCE COMPONENT(S) FOR THAT CORRELATION IS LESS THAN OR EQUAL TO ZERO.

INVERSE OF THE EXPECTED MEAN SQUARES COEFFICIENTS MATRIX

	MEAN	GCA	SCA	ERR
MEAN	0.2778D-01	-0.6349D-01	0.3571D-01	-0.2837D-15
GCA	0.0	0.1429	-0.1429	0.1332D-14
SCA	0.0	0.0	1.000	-1.000
ERR	0.0	0.0	0.0	1.000

VARIANCE COMPONENTS FOR

TOTAL YIELD

NAME	VALUE	STANDARD DEVIATION
GCA	283.75734415	149.13077587
SCA	339.43860151	89.14094388
ERR	0.0	0.0

VARIANCE COMPONENTS FOR

TOTAL YIELD - BY - COB WEIGHT

NAME	VALUE	STANDARD DEVIATION
GCA	27.29554473	15.03857260
SCA	42.96600214	11.28342495
ERR	0.0	0.0

VARIANCE COMPONENTS FOR

TOTAL YIELD - BY - SHELLED CORN WT

NAME	VALUE	STANDARD DEVIATION
GCA	256.46204149	134.09632026
SCA	296.47259530	77.85751786
ERR	0.0	0.0

VARIANCE COMPONENTS FOR

COB WEIGHT

NAME	VALUE	STANDARD DEVIATION
GCA	11.99900954	5.81847194
SCA	6.98793405	1.83512139
ERR	0.0	0.0

VARIANCE COMPONENTS FOR

COB WEIGHT - BY - SHELLED CORN WT

NAME	VALUE	STANDARD DEVIATION
GCA	15.29657708	9.23852339
SCA	35.97805883	9.44830113
ERR	0.0	0.0

VARIANCE COMPONENTS FOR

SHELLED CORN WT

NAME	VALUE	STANDARD DEVIATION
GCA	241.16566459	124.87793669
SCA	260.49454166	68.40921809
ERR	0.0	0.0

CORRELATIONS COMPUTED FROM VARIANCE (COVARIANCE) COMPONENTS

TRAITS:	TOTAL YIELD - BY - COB WEIGHT	TOTAL YIELD - BY - SHELLED CORN WT	COB WEIGHT - BY - SHELLED CORN WT
COEFFICIENT			
GCA	0.4678	0.9804	0.2844
SCA	0.8822	0.9970	0.8433
ERR	********	********	********

******** MEANS THAT THE VARIANCE COMPONENT(S) FOR THAT CORRELATION IS LESS THAN OR EQUAL TO ZERO.

* * * AND ALL IS DONE * * *

VII. Annotated List of References

1. Anderson, R. L. and T. A. Bancroft. 1952. Statistical Theory in Research. McGraw-Hill, New York, 399 pp.

2. Cockerham, C. C. 1963. Estimation of Genetic Variances. Symposium on Statistical Genetic and Plant Breeding. NAS-NRC Pub. 982:53-94.
   Relation of variance components to genetic variances. Read carefully - a non-standard use is made of the notation for the expectations of the mean squares.

3. Federer, W. T. 1967. Diallel cross designs and their relation to fractional replication. Der Zuchter 37:174-178.
   Discusses different portions of the complete diallel cross as fractional replications.

4. Griffing, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. Aust. J. Biol. Sci. 9:463-493.
   Standard reference to the different forms of balanced diallel experiments.

5. Kempthorne, O. 1957 An Introduction to Genetic Statistics. Wiley, New York, 545 pp.
   General treatment of the genetic aspects of diallel crosses.

6. Littlewood, R. K., S. G. Carmer, and C.N. Hittle. 1964. A computer program for estimating combining abilities in relation to diallel crossing systems. Crop Sci. 4:662-663.
   The program discussed is only for balanced diallel experiments, for which it would require much less computer memory and time than would DIALL.

7. Schaffer, H. E. (in preparation).
   A presentation of the algorithms used to perform the analysis in DIALL, and other general least squares analyses.

8. Searle, S. R. 1966. Matrix Algebra for the Biological Sciences. Wiley, New York, 296 pp.
   General treatment of matrices and their use in least squares analysis.

VIII. Appendix

A. Variables Used in the Program

Variables which are in the calls of the subroutines are listed in the calling program.

Fortran Labels	Type	Dimension	Meaning
			- DIALL: MAIN PROGRAM -
ALLOUT	I*4		Parameter to control printing
BL	R*8		Variable used to check for the end of a set of data, i.e., GFLN and GMLN are blank
CPD	I*4		Parameter to control whether or not cross-product analyses are computed
EPS	R*8		Epsilon such that smaller numbers are set equal to zero
FLN	I*4		Female line number converted to integer form
FMT	R*4	20	Contains format, read from input, for reading data cards
GFLN	R*8		Female line identifier as read from data in 'A' format specification, see FLN
GLN	R*8		Location line identifier as read from data in 'A' format specification, see LN
GMLN	R*8		Male line identifier as read from data in 'A' format specification, see MLN
GRN	R*8		Replication identifier as read from data in 'A' format specification, see RN
GXL	I*4		Parameter to indicate if GCA by Location Interaction is to be accounted for
I	I*4		Do loop parameter
ICY	I*4		Subscript variable used in getting total SS and SP
IFLN	I*4		Variable used to switch male and female line nos. when FLN > MLN
IGCA	I*4		First column before GCA effects in P, XPX, XPY vectors
IGXL	I*4		First column before GCA by Location effects in P, XPX, XPY vectors
IJ	I*4		Subscript variable giving location in XPY, XPX
ILX	I*4		Subscript variable giving location of RECP effect in XPX, XPY
IMLN	I*4		Variable used to switch male and female line numbers when FLN > MLN
INUNIT	I*4		Parameter for Fortran Unit Reference No. for data input
INY	I*4		Variable containing dimension of Y, checked against NY
IRECP	I*4		First column before Reciprocal effects in P, XPX, XPY
IREP	I*4		First column before Replication effects in P, XPX, XPY

Fortran Labels	Type	Dimension	Meaning
IRF	I*4		Switch to determine proper read statement depending on whether or not LOC and/or REP are > 1
ISCA	I*4		First column before SCA effects in P, XPX, XPY
ISDFL	I*4		Variable containing maximum number of effects that can be handled. Checked against the computed SDFL
ISELF	I*4		First column before Selfed effects in P, XPX, XPY
IY	I*4		As ICY
J	I*4		As I, except in developing vector where it is a counter for the no. of SDFL used
JGCA	I*4		Last column of GCA effects in P, XPX, XPY
JGXL	I*4		Last column of GCA by LOC effects in P, XPX, XPY
JLOC	I*4		Last column of Location effects in P, XPX, XPY
JMATR	I*4		Last column of Maternal effects in P, XPX, XPY
JRECP	I*4		Last column of Reciprocal effects in P, XPX, XPY
JREP	I*4		Last column of Replication effects in P, XPX, XPY
JSCA	I*4		Last column of SCA effects in P, XPX, XPY
JSELF	I*4		Last column of Selfed effects in P, XPX, XPY
K	I*4		Do loop parameter
KGGXL	I*4		Element number of GCA by Location effect in crossed progeny in the LX vector
KGCA	I*4		Element number for GCA effect for progeny in LX vector
KIGXL	I*4		Location of GCA by Location effect for a progeny in XPX, exclusive of line number
KMATR	I*4		Element number for maternal effect for progeny in LX vector
KRECP	I*4		Element number for Reciprocal effect for progeny in LX vector
KREP	I*4		Element number for Replication effect for progeny in LX vector
KSCA	I*4		Element number for SCA effect for progeny in LX vector
KSELF	I*4		Element number for Selfed effect for progeny in LX vector
KSGXL	I*4		Element number for GCA by Location effect for selfed progeny in LX vector
LABP	R*8	2,10	Labels for the Y (response or character) variables, dimensions are (2, NY)
LN	I*4		Location number on data card last read, converted to integer form
LOC	I*4		Parameter specifying the no. of locations
LX	I*4		Vector containing locations of progeny in X matrix

Fortran Labels	Type	Dimension	Meaning
			- SUBROUTINE DATINT -
			Entry DATCNV
CC	I*4		Contains the identifier being worked with; LOC, REP, FLN, MLN
I	I*4		Do loop parameter, for rows of LT
J	I*4		Do loop parameter, for columns of LT
K	I*4		Contains the no. of rows in LT to be checked
L	I*4	3	Vector containing the maximum no. of loc, reps and lines that are to be encountered, no. of rows in LT to be checked, put into K for checking
LT	R*8	30, 3	Contains the names of the locations, replications and lines as read from data
NN	I*4	4	Vector containing the integer values of the identification names
R	R*8	4	Vector containing the identifications for data just passed from DIALL as GFLIN, etc.
			- SUBROUTINE COTOIT -
EDF	I*4		Error degrees of freedom
ERMS	R*8		Error mean square
ESS	R*8		Error sums of squares
II	I*4		First element in XPY to be used in analysis
IJ	I*4		First element in XPY to be used in analysis (II = IJ for analysis of variance)
J	I*4		Do loop parameter
K	I*4		Last element of XPX to be written on next line
M	I*4		First element of XPX to be written on next line
NDEP	I*4		Number of dependencies found in XPX
RDF	I*4		Regression degrees of freedom
RSS	R*8		Regression sums of squares and products
			- SUBROUTINE COVRED -
II	I*4		Diagonal element in XPX for line being reduced
IJ	I*4		Element in XPX for line being reduced
IM	I*4		Element in XPX being used to reduce element ML
J	I*4		Do loop parameter
K	I*4		Diagonal element counter
L	I*4		Do loop parameter
M	I*4		Do loop parameter
MJ	I*4		Element in XPY being reduced

Fortran Labels	Type	Dimension	Meaning
T	R*8		Square root of XPX (II)
XPX			As in DIALL, except in reduced form when returned to GOTOIT
XPY			As in DIALL, except in reduced form when returned to GOTOIT
			- SUBROUTINE EXPMSQ -
CON	R*4		Element of P vector for effect being worked on
CONA	R*4		Next element in P vector after CON
DF	I*4		Degrees of freedom in the data
EXPMS	R*8		Coefficients of expected means squares for CON effect
I	I*4		Do loop parameter and line counter
J	I*4		Do loop parameter and column counter
K	I*4		Number of single degree lines in the CON effect
L	I*4		Column of XPX being summed
			- SUBROUTINE ANOVSS -
CON			As in EXPMSQ
DF			As in EXPMSQ
I			As in EXPMSQ
SS	R*8		Accumulation of sums of squares or products for each effect
XMS	R*8		SS/DF
			- DIALLC: MAIN PROGRAM -
CON	R*4	10	Vector of names of sources in the analysis
COR	R*8	10,7	Matrix of correlations of variance components, computed for seven characters at a time
CPD	I*4		Parameter controlling whether or not correlations are computed
DF	I*4	10	Vector of degrees of freedom for each line in the analysis
E	R*4		Variable containing name of error line-ERR
EXPMS	R*8	10,20	Matrix of coefficients of the expected mean squares; contains inverse after return from MATINV, with twice necessary no. of columns for inversion purposes
I	I*4		Do loop parameter
IAN	I*4		Variable containing variable number when error degrees of freedom do not check
IDF	I*4		Error degrees of freedom for variable in IAN
II	I*4		Number of first trait in correlation denominator

Fortran Labels	Type	Dimension	Meaning
MATR	I*4		Parameter specifying whether or not Maternal effects are desired
MLN	I*4		Male line number converted to integer form
NL	I*4		Parameter specifying maximum number of male-female lines
NKY	I*4		Counter for SS or SP in TSS vector
NO	R*4		Variable containing 'NO' for heading output
NY	I*4		Parameter specifying number of response variables
P	R*4	250	Vector containing name for each single degree of freedom line in the XPX, XPY vectors, with dimension (SDFL)
PN	R*4	8	Vector containing names used to develop P vector
RECP	I*4		Parameter specifying whether or not Recipro-effects are desired
REP	I*4		Parameter specifying number of replications
RN	I*4		Replication number on last read data card, converted to integer form
SDFL	I*4		Total number of single degree of freedom lines in ANOVA
SELF	I*4		Parameter specifying whether or not Selfs are included
STAR	R*4		Variable containing asterisks, for output only
TDF	I*4		Total degrees of freedom
TSS	R*8	55	Vector containing total sums of squares and products, with dimension ((NY*(NY+1))/2)
VCX	R*4	11	Vector containing value for XPX for crossed progeny
VINX	R*4		Variable used in negating value for XPX when FLN > MLN
VSX	R*4	6	Vector containing value for XPX for selfed progeny
WCPD	R*4		Output variable containing NO or YES for crossproduct analyses
XPX	R*8	31375	Upper right triangular part of X'X matrix in vector form, with dimension ((SDFL*(SDFL-1))/2)
XPY	R*8	2500	Matrix of X'Y values in vector form, with dimension (SDFL*NY) = (ISDFL*INY)
Y	R*4	10	Vector containing last read response variables, with dimension (NY)
YES	R*4		Variable containing YES, for heading output

Fortran Labels	Type	Dimension	Meaning
IJ	I*4		Number of trait for correlation numerator
IO	I*4		Switch to determine when new page is to be set when printing correlations
II	I*4		II plus 1 trait for correlation denominator
J	I*4		Do loop parameter
JJ	I*J		Number of second trait in correlation denominator
K	I*4		Do loop parameter
KL	I*4		Switch to determine when new set of data is to be read
LABP	R*8	10,55	Matrix of variable names
M	I*4		Number of correlation being worked on
N	I*4		Do loop parameter
NAL	I*4		Number of lines in analysis excluding the error line
NAL1	I*4		Number of lines in analysis
NNY	I*4		Number variance and crossproduct analyses
NOCOR	I*4		Switch to write message when variance components are < = 0 in correlation denominator
NY	I*4		Number of variables analyzed
NY1	I*4		NY -1, used in correlations
SDVC	R*8		Variable containing standard deviation of the variance component
SVC	R*8		Variable containing variance of variance component
VC	R*8	10,55	Matrix of variance components
VI	R*8		Individual product of EXPMS by XMS for computing variance components
XMS	R*8	10,55	Matrix of estimated mean squares

- SUBROUTINE MATINV -

Fortran Labels	Type	Dimension	Meaning
A	R*8	10,20	Matrix to be inverted
AIL	R*8		Minor diagonal element of row being reduced
AJL	R*8		Element used in reducing the Lth column
I,J,K,L	I*4		Do loop parameter and row and column counters
N	I*4		Number of rows in the matrix to be inverted
NN	I*4		2 times N
N1	I*4		N + 1