

The Alpha Lattice Design in Plant Breeding and Agronomy: Generation and Analysis

H.Barreto, G.Edmeades, S.Chapman and J.Crossa¹

Summary

This document presents a system that integrates available PC software for developing and analyzing alpha lattice designs. The introduction presents theoretical considerations on the use of incomplete block designs for evaluating genotypes and their potential use in agronomic trials. This type of software module is expected to offer an additional resource in improving the efficiency of agricultural research in national programs.

Introduction²

The traditional design for evaluating agricultural experiments is one in which the experimental units are grouped in complete blocks that each contain all treatments. This design assumes that variation among units within a block is less than that among units of different blocks, so that experimental precision is increased by error control (Steel and Torrie, 1980). Also, with this type of design each treatment occurs with the same frequency in each block; i.e. blocks and treatments are orthogonal. Orthogonality facilitates the mathematical calculation needed to produce an analysis of variance.

Table 1 shows a general classification of the experimental designs that are commonly used in agriculture. Table 2 summarizes the principal types of square and rectangular lattices. However, when the number of treatments is large the randomized complete block (RCB) and the Latin square are generally unsuitable because the size of the block increases and soil

¹ Maize Program Agronomist, Coordinator of the Maize program Physiology and Agronomy Unit, Maize program Physiologist and Biometrics Unit Coordinator, respectively. CIMMYT, International Maize and Wheat Improvement Center, Apdo. Postal 6-641, 06600, Mexico, D.F., Mexico.

² The introduction to incomplete block designs is based on class notes of Dr. Carlos González' (SAS Institute, Mexico) for CIMMYT wheat/maize training programs.

heterogeneity effects also increase. This may cause the variation within blocks to be larger than the variation between blocks, which tends to increase the experimental error and diminish the researcher's ability to observe statistical differences between any two treatments.

In an incomplete block design (e.g., lattice), the number of plots per block (k) is smaller than the total number of treatments (v). An incomplete block design is considered resolvable if the blocking units are arranged in complete replicates (a complete set of treatments is obtained by putting together the blocks of one replicate). In the field, an incomplete block design is indistinguishable from a randomized complete block design. However, practical considerations dictate that most designs used for agricultural trials be resolvable. Because they are resolvable, lattices can never be less efficient than a complete block design.

The quality of an incomplete block design is judged by the Harmonic-Mean Efficiency Factor (HMEF), which is the ratio of the average variance with which a complete block design would estimate treatment differences and the average variance with which the incomplete block design would estimate these differences, if the error mean square were the same in both cases. The value of HMEF is always bigger than 0 and smaller than 1. The difference measures the confounding between treatments caused by using incomplete blocks. When laying out the trial in the field, we should try to achieve maximum homogeneity possible within each incomplete block depending on soil conditions, fertility gradients, moisture, slope, etc. The rule of thumb for selecting the size (number of plots per block) of an incomplete block is to take a value that is the same or smaller than the square root of the number of treatments ($k = \sqrt{v}$).

To judge whether the incomplete block arrangement was more effective than the complete block design, we compare the corresponding experimental errors. The relative efficiency (RE) of a lattice design is defined as the ratio between the average variance of the differences between treatments (ignoring the use of incomplete blocks and assuming that the replicates were complete blocks), and the average variance of the differences between

treatments in the incomplete block design (including recovery of inter-block information). The RE ratio is analogous to the definition of the HMEF, but in the former, observed estimates of the experimental error are used. If incomplete blocking has been successful in removing variation (i.e., if the reduction in the error mean square compensates for the effect of confounding caused by incomplete blocks), then the relative efficiency will be greater than 1. As mentioned above, the fact that lattices can revert to complete blocks if the efficiency is less than 1 ensures that lattices can never be less efficient than a complete block design.

Table 3 lists some historical events that were important in the development of experimental designs, including alpha designs devised by Patterson and Williams (1976). These are type (0, 1) incomplete block designs. The term in parentheses indicates the type of restriction the combination of paired entries have with respect to incomplete blocks. In the case of an alpha lattice (0, 1), every pair of treatments never appear together (0) or are together in only one incomplete block (1). Many alpha-lattice designs exist for a given number of treatments (v), block size (k) and number of replicates (r); however, only some have an HMEF that is high enough and most are unacceptable. The best alpha-lattice design for a specific combination of replicates, blocks, and treatments is obtained using an algorithm (Patterson and Patterson, 1984) that is the basis of the ALPHAGEN program.

In CIMMYT's Maize Program, alpha-lattice designs have been used since 1986, particularly in the physiology-agronomy unit for evaluating trials with more than 200 genotypes, generally in one-row plots 2.5 m long. Relative efficiencies of 1.0-2.0 have been observed, but their use has been restricted to genotype evaluation trials. There is no information on the use of alpha-lattices in agronomic experiments at CIMMYT, although they could be useful when there is no factorial type relationship among treatments, the number of treatments is large, and soil heterogeneity is potentially high. In these cases, the use of incomplete blocks is suggested in order to control random variation better.

Despite the fact that software programs for designing and analyzing alpha-lattices are becoming available to national program researchers, little use has been made of them. Among the most important reasons are the absence of a software interface for entering and generating data files, the difficulty of handling information in a common format, and the lack of integration with other analytical tools (graphs, selection algorithms, regression, etc.).

The objective of this document is to introduce a software module (ALPHA) that facilitates and integrates the use of the ALPHAGEN and LATANOVA (formerly ALPHANAL) programs (Scottish Agricultural Statistics Service, Edinburgh) for the design and analysis of alpha-lattices (0, 1) and later processing them with Selection Assistant (Barreto *et al.*, 1991) as part of the genotype evaluation and selection process in a maize improvement program. ALPHA is currently available in both English and Spanish versions.

Software Description

At this time, CIMMYT is using ALPHAGEN for generating randomization and LATANOVA for analysis of variance of alpha-lattices. Both programs were developed originally by the Scottish Agricultural Statistics Service, Edinburgh, in FORTRAN computer language and were brought to CIMMYT in their original codes by C. González to be used on the VAX main frame computer. Later ALPHAGEN was modified by CIMMYT's Computer Services staff to produce output in the form of CIMMYT field books and was adapted for use on IBM compatible microcomputers. ALPHANAL was later modified by G.O. Edmeades to incorporate statistics from randomized complete blocks, and its name was changed to LATANOVA. Both programs were further modified especially to improve the output format, by S.C. Chapman and G.O. Edmeades.

Some general characteristics of both programs as currently implemented for use on IBM-compatible personal computers are presented in the following paragraphs.

ALPHAGEN generates randomization and field layouts for alpha-lattice (0, 1) designs. Its present limitations are: maximum number of

replications = 6; maximum number of entries = 300. The program requests input from the user and displays the randomization plan on screen or stores it in a disk file. Output includes block layout, replication, block, and entry codes, and the HMEF value. Processing this information further on a word processor has been a pre-requisite for using the analysis program.

LATANOVA performs the analysis of variance and generates the lattice adjusted means. The adjusted means can later be used in comparisons or in selecting treatments. This program has strict requirements in terms of input file characteristics (three different files), which, to a degree have made it hard to use. The number of variables that can be analyzed at a time is restricted to a maximum of ten.

CIMMYT's Maize Program developed data conversion software programs to facilitate the use of both programs. They were later integrated into a software system³ which assumes that all information is stored in MSTAT-C (1990) files. This step has made information easy to handle since MSTAT-C is currently being used in many national programs for entering data and analyzing trial results.

Structure of the Integrated System

To start the system, the programs are loaded from the distribution diskette onto a subdirectory on the hard disk. Then the word ALPHA is keyed in at the operating system prompt. Fig 1 illustrates the options menu available for designing and analyzing alpha-lattice (0, 1) experiments. The desired option is selected with the cursor or by pressing the letter describing that option. Options on the main menu include: 1) Produce a new randomization; 2) Analysis of variance and means adjusted for lattice; 3) Selection assistant; and 4) Editing and transforming data files.

Each option has a specific subprogram to enter the information needed to carry out the desired operation. Questions for the option of generating new randomization (1) and analysis

of variance (2) are shown in Figs 2 and 3, respectively. The description of the Selection Assistant option is not presented in this document (Barreto *et al.*, 1991). The option for editing and transforming data files (4) is illustrated in Figure 4. For using the latter, the program follows the software conventions described in the MST Data Assistant program (Barreto and Raun, 1988).

The following paragraphs present a detailed description of the questions included in options 1 and 2 for generating and analyzing alpha-lattices, respectively.

Option 1: Generating New Randomization

Number of replications in the design

Number of replications (r) per treatment.

Number of total entries

Total number of genotypes or treatments (including checks).

Number of entries per block

This defines the size of the block (k). The current software version only allows a block size that is an integer multiple of the total number of entries. For example, if the total number of entries is 33, possible block sizes can only be 11 or 3.

Randomize the first replication (Y,N)?

The user can request that the first replication not be randomized but ranked by entry order (which is generally not advisable). It should be noted that using this option does not alter the alpha-lattice structure, but there is a sequential rearrangement of the entry numbers on replicate 1, along with the corresponding rearrangement of entries in all remaining replications as to maintain the lattice structure.

Number of checks

If a certain number of cultivars are to be included as checks, ALPHAGEN attempts to place each check in a different block. In randomization, the checks use entry numbers from 1 to the specified number of checks.

³ The ALPHA system was developed in Turbo Pascal language (c)v 7.0 for MS-DOS by Borland International 1992.

Name of the MSTAT output file

This is the master file generated by option 1 which should be used for entering data of all variables to be analyzed and for using the analysis program. This design information in MSTAT format is stored in files having the extensions .TXT and .DAT. However, the program under option 1 generates intermediate files in ASCII format having the same name but different extensions and which contain all the ALPHAGEN output. These files are: Extension .FDS contains all output information including HMEF. Extension .DSN contains the field books in ASCII format (Rep, Block, Plot, and Entry), in plot order).

Option 2: Analysis of Variance of Alpha-Lattices

File containing data

This is the MSTAT file generated by the design option, to which have been added variables with data to be analyzed (e.g., yield, flowering date, ASI, etc.).

Variable with replication number

The number of the column in the MSTAT file that contains the replication code for the alpha-lattice design generated by ALPHAGEN.

Variable with block number

The number of the column in the MSTAT file that contains the block code for the alpha-lattice design generated by ALPHAGEN.

Variable with plot number

The number of the column in the MSTAT file that contains the plot code for the alpha-lattice design generated by ALPHAGEN.

Variable with entry number

The number of the column in the MSTAT file that contains the entry or treatment code for the alpha-lattice design generated by ALPHAGEN.

Variable(s) to be analyzed

Numbers of the columns in the MSTAT file that contain data to be processed in the analysis of variance option. The software can analyze a maximum of 10 variables at the same time. These are specified by their corresponding column number, separated by spaces (e.g., 3 4 5) or using a dash (e.g., 3-5).

First case to be included

The first case of the MSTAT file to be included in the analysis.

Last case to be included

The last case of the MSTAT file to be included in the analysis.

MSTAT file for adjusted means

Name of file supplied by user to be used for storing lattice adjusted means. The name should be different from that of the master file containing the original data.

Efficiency of the alpha-lattice design

This is the HMEF value generated by ALPHAGEN. This value is recorded in the file with an .FDS extension and in the MSTAT file as the file descriptor.

Title of work

This is an alternative descriptor defined by the user to identify the work in the analysis program. It appears as a title at the top of each analysis of variance.

Hardware requirements

To be used in an integrated manner, the software system requires an IBM compatible computer (8086, 286, 386, or 486, with or without a mathematical coprocessor), equipped with a minimum of 640 K RAM, a hard disk with at least 2 megabytes available space and MS-DOS operating system (ver. 3.3 or higher).

However, on a 8086 or 286 without a coprocessor, it takes considerably more time to run an analysis of variance.

The amount of available memory for the program after loading the operating system should be at least 590 K to run ALPHAGEN. The other options require less memory (512 K). The amount of memory available can be obtained by executing the CHKDSK or MEM command (only on DOS version 5.0 or higher).

Types of Output Files

Option 1: Generate new Randomization

As indicated, option 1 produces a series of files in ASCII and binary format that contain the randomization and additional information on the lattice design. These files have all the same name (defined by the user), but have different extensions. Depending on their extensions, files contain the following:

ALPHAGEN.IN files: This file contains information needed to run the ALPHAGEN.EXE program using entry redirection or piping (e.g., ALPHAGEN <ALPHAGEN.IN). This step is transparent to the user because ALPHAGEN is run directly from option 1. The file is in ASCII format and contains the answers to questions in ALPHAGEN.EXE, one answer per line, including random seeds used to generate randomization (these are automatically provided by the program in option 1). Since the name of this file is always the same (and is regenerated every time option 1 is used), the file on the disk is always the same as for the last randomization performed.

Files with an .FDS extension: Output information (ASCII format) that includes HMEF, the number of reps, blocks, and entries (including checks), and the field book with randomization ordered by plot number and with headings for every column (Tables 4 and 5).

Files with a .DSN extension: Contain the field book in ASCII format (rep, block, plot and entry, in that order), but with no headings. These files may be used to generate field books using a word processor.

Files with .TXT and .DAT extensions: MSTAT binary files containing field layouts ordered by plot number. These files are ready to be edited by the user at the end of each option 1 session. The user should enter the variables of interest using the data editor to add variables (columns) and do transformations between variables using the conventions described in the MST Data Assistant (Barreto and Raun, 1988).

Option 2: Analysis of Variance of Alpha-Lattices

Option 2 performs the analysis of variance of alpha-lattice designs by calling the LATANOVA program. As indicated, a prerequisite for using this option is that the data to be analyzed be in MSTAT format. During the analysis, option 2 does a number of file conversions to make it easy for the user to utilize the option. Among those conversions are the creation of new ASCII files with the field layout of the design, the data to be analyzed and storage of the lattice-adjusted means in a binary MSTAT file. As in the previous option, the output files have the same name as defined by the entry file but different extensions. Depending on their extensions, files contain the following:

LATANOVA.IN file: These files contain information needed to run the LATANOVA.EXE program using entry redirection or piping (e.g., LATANOVA <LATANOVA.IN). The file is in ASCII format and contains the answers to the questions required by LATANOVA.EXE, one answer per line. Since the name is the same (and is regenerated each time the option is used), the file on the disk is always the same as for the last analysis done using option 2.

Files with a .DSN extension: Contain the field book in ASCII format (rep, block, plot and entry, in that order), but without heads. This file can be used to generate field books using a word processor.

Files with a .LAD extension: Data to be analyzed (ASCII format) which include plot number and data variables separated by spaces with heads for each column.

Files with a .VAR extension: A FORTRAN reading format for each variable and names of variables to be analyzed.

Files with a .AOV extension: A LATANOVA.EXE output file in ASCII format that contains the analysis of variance tables for each variable analyzed and the means with and without lattice adjustment (Table 6).

Adjusted means files with .TXT and .DAT extensions: Files in binary MSTAT format that contain the lattice-adjusted means for the variables of analysis selected in option 2. The file always contains the first variable with the entry number (treatment) and the last variable called INDEX This makes it easy to use this file with the Selection Assistant (Barreto et al., 1991). This file can be modified using the data editor in option 4.

Examples of Output Files

Option 1: Generate new randomization

Table 4 show an example of randomization for a file having 24 entries in incomplete blocks of size 6 and two reps. The file presented is a .FDS extension file (complete design). Notes in the right margin indicate the important parts of this output file. Table 5 briefly describes each of these sections.

Option 2: Analysis of Variance of an Alpha-Lattice

Table 6 gives an example of an output file (extension .AOV) generated by LATANOVA. Shown in the right margin are the descriptions of each of the statistical parameters and elements for analysis.

References

Barreto, H.J. y W.R. Raun. 1988. *MST Assistant. Guide to software operation. Training Material. CIMMYT. Lisboa 27, Apdo. postal 6-641 06600 México D.F. México.*

Barreto, H.J., J.A. Bolaños y H.S. Córdova. 1991. *INDICE DE SELECCION: Guía para la Operación del Software. Manual de Capacitación Regional. Programa Regional de Centromérica y el Caribe. Guatemala. Apdo Postal 231-A, Guatemala*

González, C. 1989, *Curso de estadística aplicada. CIMMYT unidad de Capacitación. Notas de Clase. Mimeo.*

MSTAT-C. 1990. *A microcomputer program for the design, management, and analysis of agronomic research experiments. Crop and Soil Sciences, Michigan State University, East Lansing, MI.*

Paterson, L.J. and H.D.Patterson. 1984, *An algorithm for generating alpha-lattice designs. ARS Combinatoria. 16A:87-98.*

Patterson H.D. and E.R. Williams, 1976. *A new class of resolvable block designs. Biometrika 63:83-92.*

Steel R.G.D. y J.H. Torrie..1980. *Principles and procedures of statistics. A biometrical approach. Second edition. McGraw Hill. New York.*

Table 1. General Classification of experimental designs.

Type of Design	Block design	Row-column designs
Complete block	randomized blocks	latin square
Resolvable incomplete block designs	Square lattices * simple * triple	lattice square
	rectangular lattice * simple * triple	
	alpha designs	row-column alpha designs

Source: C.González class notes.

Table 2. Main types of square and rectangular lattices.

Name	No. genotypes	No. reps multiple of	No. blocks multiple of
Simple lattice	k^2	2	2k
Triple lattice	k^2	3	3k
Simple rectangular lattice	$k(k+1)$	2	$2(k+1)$
Triple rectangular lattice	$k(k+1)$	3	$3(k+1)$

N.B. number of plots per block is k in all cases

Source: C.González class notes.

Table 3. Some important events in the history of experimental design.

Year	Event
1910	The concept of replication is introduced in agricultural research
1920's	R.A.Fisher introduces the randomization concept and the complete block designs (variation control in one direction) and latin square (variation control in two direction perpendicular among each other).
1936	Frank Yates introduces square Lattices
1940	Frank Yates introduces Lattice square designs
1949	Harshbarger introduces Rectangular lattices designs
1976	Patterson and Williams introduce alpha designs
1978	Patterson, Williams and Hunter produce a catalog of alpha designs for experiments in the range of 20-100 treatments and up to 4 replications
1984	Paterson and Patterson developed an algorithm that automatically generates alpha lattice designs. In 1987 this algorithm is made available for use in microcomputers.

Source: C.González class notes.

Table 5. Explanation of an output file generated by ALPHAGEN.EXE for an experiment with 2 replicates, 24 entries and 4 incomplete blocks of 6 plots each.

This is a text file with four sections but only sections A and D are important.

SECTION A

Characteristics of selected design

DESIGN 1	TYPE OF DESIGN				
2 24 4 6	<u>.780439</u>				

respectively, NO. REPS/TRTS/BLKS/BLK SIZE and Efficiency Factor of Harmonic Mean

1

SECTION B

Field layout with plot 1 on the upper left corner in block 1.

8	7	20	12	16	21	BLOCK 1
17	9	15	10	1	4	BLOCK 2
19	11	5	23	14	2	BLOCK 3
6	24	22	18	13	3	BLOCK 4

SECTION C

Entries in each rep and block. The entries in first rep are always presented in sequence and the numbers have been correspondingly swaped in the other reps.

SECTION D

This is the most important section as it includes the randomization in a format that can be utilized for preparation of field books of imported into MST.

Table 6. Example of output file from LATANOVA (.AOV) with ANOVA tables and unadjusted and adjusted means.

#PLOTS= 450 #REPS= 2 # VARS= 225 # BLOCKS/REP= 15
 EFFY= .8889 CRIT= .00 MISS. VAL. CODE=-99.

CODE FOR DESIGN
HMEF (FIRST VALUE)

ANOVA of Trial 1603M

 ANALYSIS OF VARIANCE FOR VARIABLE AD

VARIABLE ANALYZED

LATTICE ADJUSTED ANALYSIS

ANOVA FOR LATTICE

SOURCE	DF	MS
REPS	1	.0200
BLK(ADJ)	28	3.9279
TRT	224	10.7597
TRT(ADJ)	224	10.3873
RESID	196	2.6914
TOTAL	449	6.7877

F (REPS)= .007 (1, 196 DF)
 F (TRTS)= 3.859 (224, 196 DF)

AV. SE 1.1922 LSD(0.05) 3.3247 F SIG 1.0 DF 196. CV 2.2

RCBD ANALYSIS

ANOVA FOR RCBD

SOURCE	D.F.	SUMS OF SQUARES	MEAN SQUARES	F VALUES
REPS	1	.25000000	.25000000	.088
ENTRIES	224	2410.25000000	10.76005000	3.783
ERROR	224	637.16460000	2.84448500	
TOTAL	449	3047.66500000		

REP MEANS

REP MEANS

1 74.102
 2 74.116

S.E.(DIFF) 1.68656(224)D.F.
 COEFF. OF VARIATION 2.28 OVERALL MEAN 74.1089

STANDARD ERROR OF DIFFERENCE
CV AND OVERALL MEAN

RELATIVE EFFICIENCY RCBD/LATT ADJ:
 CALCULATED AS (SED(RCBD)**2/SED(LATT.ADJ)**2)= 1.0007

RELATIVE EFFICIENCY

1

UNADJUSTED MEANS								
ENT	AD	ASI	GY	EPP	GPE	100KW	LOGG%	LROLSC
1	75.00	1.00	.81	.83	166.05	11.05	16.66	2.16
2	74.00	2.00	1.19	.65	287.72	12.85	18.06	3.00
3	72.50	-1.00	1.55	.94	212.12	14.55	11.11	2.50
.
225	68.50	-.50	1.71	1.00	264.05	12.30	11.11	3.50

UNADJUSTED MEANS RCBD

1

ADJUSTED MEANS								
ENT	AD	ASI	GY	EPP	GPE	100KW	LOGG%	LROLSC
1	74.88	1.20	.71	.82	165.69	10.55	23.31	2.42
2	74.00	1.99	1.19	.67	287.33	13.13	20.00	3.03
3	72.48	-1.09	1.56	.97	211.99	14.52	12.28	2.47
.
225	68.58	-.41	1.70	.99	264.33	12.18	6.45	3.43
MEAN	74.11	.40	1.03	.86	198.62	11.74	17.01	2.95
LSD	3.32	2.61	.58	.30	74.69	2.73	30.44	.96
CV	2.21	318.91	27.36	16.67	19.02	11.36	87.57	15.76

LATTICE ADJUSTED MEANS

DESCRIPTIVE STATISTICS MEANS

GIMMYT

Generation and Analysis of Alpha-Lattice Designs

```
N NEW RANDOMIZATION
A ANOVA ALPHA-LATTICE
S SELECTION ASSISTANT
V VIEW/EDIT DATA FILE
T TERMINATE
```

Figure 1. Main menu with options for generation and analysis of alpha-lattice designs.

```
Generate Randomization Alpha-Lattice (0,1)
Number of replications in design      | 2
Number of treatments (e.g. genotypes) | 4
Number of plots per block             | 2
Randomize first replication (Y,N)?    | Y
Number of check treatments            | 0
Output file name                     | TEMPO
```

Figure 2. Questions required for creating a new randomization for an alpha lattice design.

```
Analysis of Variance with LATANOVA
MST File name containing data
Variable containing replicate number  | 1
Variable containing block number     | 2
Variable containing plot number      | 3
Variable containing entry number     | 4
Variables to be used in analysis     | 1
First case to be included            | 1
Last case to be included             | 1
MSTAT filename for adjusted means
Harmonic mean efficiency factor      | 0.9999
Job title
```

Figure 3. Screen for using a MSTAT master file and executing LATANOVA.

```
Create / Read MST data files
File name |
```

Figure 4. Screen for accessing data editor