

7.1 Introduction

So far the attention has been focused on designs in which the interest of the experimenter is to make all the possible pairwise treatment comparisons. However, this is not the case always. There exist situations where the treatments studied in an experiment comprise of two disjoint groups and each group of treatments has different importance to the researcher. The comparisons of treatments within groups may not be of interest or may be of less importance to the researcher while the comparisons of treatments between groups may be of interest or of more importance to the researcher. In other words, the interest of the experimenter is only in a subset of all the possible pairwise treatment comparisons. We illustrate this fact through examples studied in the literature.

Experimental Situation 1: (Federer, 1956). In a sugarcane breeding trial conducted at Hawaii, four sugarcane varieties *viz.*, A , B , C , or D , and eleven-seedling tests $e, f, g, h, i, j, k, l, m, n$, or o , were tried to evaluate the seedlings. Replicated plots on the individual seedlings were not possible because of the scarcity of seed cane and the large plots required for experimentation. One of the objectives of the trial was to make comparisons among the members of the two groups of sugarcane variety and seed cane. For obtaining the experimental error, sugarcane varieties were replicated. The trial was conducted in four blocks ($b = 4$) with seven plots in three blocks ($k_1 = k_2 = k_3 = 7$) and six plots in the fourth block ($k_4 = 6$). The layout of the design, without randomization, is the following:

B1	B2	B3	B4
A	A	A	A
B	B	B	B
C	C	C	C
D	D	D	D
e	h	k	n
f	i	l	o
g	j	m	

Experimental Situation 2: (Pearce, 1960). In a strawberry weed killer trial, it was intended to find out whether the application of any of the weed killers, A , B , C , or D , all of which were apparently suitable for controlling weeds in strawberry fields, would harm the growth of fruiting

strawberry plants. The trial was run in a design comprising of four blocks ($b = 4$) of seven plots each ($k = 7$) and there were four treatments comprising of four kinds of herbicides ($w = 4$) besides an untreated control (O). The layout, without randomization, is given as:

B1	B2	B3	B4
O	O	O	O
O	O	O	O
A	A	A	A
A	B	B	B
B	C	B	C
C	D	C	C
D	D	D	D

Interestingly, initially the design planned was randomized complete block (RCB) design with five test treatments (weed killers, A , B , C , D , or E) modified by adding two control replications in each block. But at last moment it was observed that supply of herbicide E still had not arrived and a decision had to be made quickly. It was thought of merging treatment E with any one of the treatments A , B , C , or D , in each block thus doubling the number of plots assigned to one of other substances. But sufficient supplies were not available for any of them (*i.e.* merging of treatments was not possible). Hence in desperation it was thought that in each of the four blocks treatment E will be exchanged with distinct treatments like A , B , C , and D in block I, II, III and IV, respectively. As a result of this a new class of designs, described above, was discovered.

Experimental Situation 3: In a trial, seven treatments (test treatments) were tried along with three controls. The purpose of the experiment was to make comparisons between the treatments in the two groups. The trial was laid out in ten blocks ($b = 10$), the first seven blocks having six plots each ($k_1 = k_2 = \dots = k_7 = 6$) and the last three blocks having ten plots each ($k_8 = k_9 = k_{10} = 10$). The design adopted is the following:

B1	1	2	4	A	B	C				
B2	2	3	5	A	B	C				
B3	3	4	6	A	B	C				
B4	4	5	7	A	B	C				
B5	5	6	1	A	B	C				
B6	6	7	2	A	B	C				
B7	7	1	3	A	B	C				
B8	1	2	3	4	5	6	7	A	B	C
B9	1	2	3	4	5	6	7	A	B	C
B10	1	2	3	4	5	6	7	A	B	C

Experimental Situation 4: (Ture, 1994). A certain type of synthetic fiber is used in the production of various consumer goods. The research team of the manufacturer of this fiber has developed three new types of synthetic fibers that can be used for the same purpose. Each of these alternative fibers is more cost-efficient than the present one and can replace it if any one of them is proved to be stronger. An experiment was conducted to compare the breaking strengths of all these synthetic fibers. Suppose that 4 testing machines and 5 operators are available for the experiment. Because variability between the machines and the operators is suspected, the experiment must be designed to control such variability. Assuming that each operator can work on each testing machine once only, the following design may be used which is efficient for making tests *vs* controls comparisons.

Machine ↓	Operator →				
	A	B	C	D	E
I	0	3	0	1	2
II	1	2	0	0	3
III	3	1	2	0	0
IV	2	0	1	3	0

For the experimental settings considered here, all the possible pair-wise comparisons among treatments are not of equal interest to the researcher. In fact, the researcher is interested only in a subset of comparisons comprising of tests *vs* controls comparisons or pairwise comparisons among treatments belonging to the two groups. The comparisons among tests and among controls may be of little or no consequence to the researcher. For this experimental setting the variance-balanced designs for making all possible pair-wise treatment comparisons may not be useful.

Suppose that there are 6 treatments tried in an experiment *viz.*, A, B, C, D and $0, 1$. If one is interested in making all the possible pairwise treatment comparisons, then there would be the following 15 comparisons: $(A, B), (A, C), (A, D), (A, 0), (A, 1), (B, C), (B, D), (B, 0), (B, 1), (C, D), (C, 0), (C, 1), (D, 0), (D, 1), (0, 1)$. In these comparisons one may notice that every treatment appears 5 times in the 15 pairwise comparisons. So using any design described so far with equal or as far as possible equal replications will be an obvious choice. But suppose that the treatments A, B, C, D are tests (new treatments) and treatments $0, 1$ are the controls (standard treatments or existing practices). The experimenter is not interested in making pairwise comparisons among treatments within groups. The interest is only in pairwise comparisons between the two groups. The following 8 pairwise comparisons of treatments are of interest now to the experimenter: $(A, 0), (A, 1), (B, 0), (B, 1), (C, 0), (C, 1), (D, 0), (D, 1)$. Now it may be seen that treatments A, B, C, D appear twice in the comparisons but treatments $0, 1$ appear 4 times each. Thus intuitively it is apparent that one needs to have designs with unequal replications of treatments with treatments $0, 1$ replicated more times than the treatments A, B, C, D . Obviously then the designs with equal replications of treatments may not be good for these experimental situations.

The experimental situations described above can be classified in two broad categories *viz.* Category A: where it is not possible to replicate the test treatments (Experimental Situation 1) and Category B: where it is possible to have replication of test treatments as well (Experimental Situations 2,3, and 4). The designing and analysis of these experimental situations is described in the sequel.

7.2 Category A experiments with single replication of tests (Augmented designs)

Category-A designs are essentially augmented designs. In genetic resources environment, which is a field in the forefront of biological research, an essential activity is to test or evaluate the new germplasm / provenance / superior selections (test treatments), etc. with the existing provenance or released varieties (control treatments). A problem in these evaluation studies is that the quantity of the genetic material collected from the exploration trips is very limited or cannot be made available since a part of this is to be deposited in Gene Bank. The available quantity of seed is often not sufficient for replicated trials. Moreover, the number of new germplasm or provenance to be tested is very high (usually about 1000-2000 and sometimes up to 3000 accessions), and it is very difficult to maintain the within block homogeneity. These experimental situations may also occur in the fields of entomology, pathology, chemistry, physiology, microbiology, agronomy and perhaps others for screening experiments on new material and preliminary testing of experiments on promising material. In some other cases (*e.g.* physics), a single observation on new material may be desirable because of relatively low variability in the experimental material. These types of situations came to be known to Federer around 1955 in screening new strains of sugarcane and soil fumigants used in pineapples. Augmented (Hoonuiaku) designs were introduced by Federer (1956) to fill a need arising in screening new strains of sugarcane at Experimental Station of Hawaii Sugarcane Planters Association on the basis of agronomic characters other than yield.

Thus, we have seen that we have to design an experiment in which the experimental material for new (test) treatments is just enough for a single replication. However, the connectedness property of the design is ensured by augmenting any standard connected design in control treatments with new (test) treatments and replications of the control provide the estimate of error. More precisely, ***an augmented experimental design is any standard experimental design in standard (or control) treatments to which additional (new) treatments have been added.*** The additional treatments require enlargement of the complete blocks or incomplete blocks in a block design set up or rows or / and columns in a row - column design set up, etc. The groupings (or blocks) in an augmented design may be of unequal sizes.

Augmented designs can be run in 0-way and 2-way elimination of heterogeneity settings also. Augmented designs eliminating heterogeneity in one direction are called augmented block designs and augmented designs eliminating heterogeneity in two directions are called augmented row-column designs. Federer (1956, 1961) gave the analysis, randomization procedure and construction of these designs by adding the new treatments to the blocks of RCB design and balanced lattice designs in control treatments. Federer (1963) gave procedures and designs useful for screening material inspection and allocation with a bibliography. Federer

and Raghavarao (1975), who obtained augmented designs using RCB design and linked block design for one-way heterogeneity setting, gave a general theory of augmented designs. They also gave a method of construction of augmented row-column design using a Youden Square design and also provided formulae for standard errors of estimable treatment contrasts. Federer *et al.* (1975) gave systematic methods of construction of augmented row column design. A procedure of analysis of data generated from these designs has also been given. The estimable contrasts in such designs may be (i) among new varieties (test treatments), (ii) among check varieties (control treatments), and (iii) among all check and new varieties simultaneously. Indeed it may be possible to estimate the contrasts between check and new varieties. We shall concentrate on augmented designs for 1-way elimination of heterogeneity settings. In general, the randomization procedure for an augmented block design is:

1. Follow the standard randomization procedure for the known design in control treatments or check varieties.
2. Test treatments or new varieties are randomly allotted to the remaining experimental units.
3. If a new treatment appears more than once, assign the different entries of the treatment to a block at random with the provision that no treatment appears more than once in a block until that treatment appears once in each of the blocks.

The analysis of variance of the data generated from an augmented block design with $v = u + w$ treatments comprising of w tests and u controls arranged in b blocks having k_1 plots (experimental units) in block 1, k_2 plots (experimental units) in block 2, and so on, and k_b plots (experimental units) in block b , such that $k_1 + k_2 + \dots + k_b = n$, the total number of plots (experimental units) in the design, is sketched in Table 7.1.

Table 7.1: ANOVA table for augmented block design

Source	DF	SS	MS	F-value
Blocks (Eliminating treatments)	$b - 1$	$ASSB$	$MSSB$	$MSSB/MSE$
Treatments (Eliminating blocks)	$v - 1$	$ASST$		
Among Tests	$w - 1$	SST	$MSST$	$MSST/MSE$
Among Controls	$u - 1$	SSC	$MSSC$	$MSSC/MSE$
Tests vs Controls	1	$SSTC$	$MSSTC$	$MSSTC/MSE$
Error	$n - v - b + 1$	SSE	MSE	
Corrected Total	$n - 1$	TSS		

For making the exposition clear, we shall consider the Augmented Randomized Complete Block Design. Let us consider the experimental situation where w test treatments (t th test denoted by $N_t, t = 1, 2, \dots, w$) are to be compared with u control treatments (s th control denoted by $C_s, s = 1, 2, \dots, u$) via n experimental units arranged in b blocks such that j th block is of size $k_j (> u), j = 1, 2, \dots, b$. For an augmented randomized complete block design, each of the control treatments is replicated b times and occurs once in every block and test treatments occur only

once in any one of the b blocks. Therefore, it can be seen easily that in the j th block there will be $k_j - u = n_j$ test treatments, $j = 1, 2, \dots, b$. The randomization procedure is given as follows:

1. Randomly allot u controls to u of the k_j experimental units in each block, $j=1, 2, \dots, b$.
2. Randomly allot the w test treatments to the remaining experimental units.
3. If a new treatment appears more than once, assign the different entries of the treatment to a complete block at random with the provision that no treatment appears more than once in a complete block until that treatment occurs once in each of the complete blocks.

For augmented randomized complete block design standard errors for comparing mean differences are as follows

Estimated standard errors of the estimated difference

- (i) Between two control treatments, $SE(1) = \sqrt{\frac{2MSE}{b}}$
- (ii) Between two test treatments in the same block, $SE(2) = \sqrt{2MSE}$
- (iii) Between two test treatments not in the same block, $SE(3) = \sqrt{2MSE(1+1/u)}$
- (iv) Between a test treatment and a control treatment, $SE(4) = \sqrt{MSE(1+1/b+1/u-1/ub)}$

7.3 Example 1

An experiment was conducted with $w = 8$ new accessions (that were to be tested) denoted by $N_t, t = 1, \dots, 8$ and $u = 4$ control treatments denoted by $C_s, s = 1, \dots, 4$ of a genotype. There are 20 plots (experimental units) that could be arranged in three blocks ($b = 3$). There are 7 plots (4 for control treatments and 3 for new accessions) in the first and third block and 6 plots (4 for control treatments and 2 for new accessions) in the second block, *i.e.*, $k_1 = k_3 = 7; k_2 = 6$. For random allocation of these treatments in the experiment, we have to proceed as:

- (i) Allot the 4 control treatments to each block randomly. In this process, say following is the arrangement:

Blocks	Experimental units						
	1	2	3	4	5	6	7
1		C_3	C_4		C_1	C_2	
2	C_4	C_2	C_1	C_3			
3		C_3	C_1		C_4	C_2	

The 7th experimental unit is for blocks 1 and 3 and not for block 2. Of the 20 experimental units, 12 have been allotted to the control treatments. The remaining 8 will be allotted to the 8 new accessions.

8 new accessions are allotted randomly to the remaining experimental units of the 3 blocks. This way 4 controls and 8 new accessions randomly occupy 20 experimental units. The final arrangement looks as in Table 7.2.

Table 7.2: Data from an augmented block design

Blocks	Experimental units						
	1	2	3	4	5	6	7
1	N_8 (74)	C_3 (78)	C_4 (78)	N_3 (70)	C_1 (83)	C_2 (77)	N_7 (75)
2	C_4 (91)	C_2 (81)	C_1 (79)	C_3 (81)	N_1 (79)	N_5 (78)	
3	N_4 (96)	C_3 (87)	C_1 (92)	N_2 (89)	C_4 (81)	C_2 (79)	N_6 (82)

The figures in the parenthesis represent the observed value of the character under study from an experiment conducted in the above layout. Source for this data is Federer (1956). The analysis of the data has been carried out and the ANOVA Table 7.3 is given.

Table 7.3: Analysis results of the data in Table 7.2

ANOVA

SOURCE	DF	SS	MS	F-value	Prob > F
Blocks(eliminating treatments)	2	69.500	34.750	1.29	0.3424
Treatments(eliminating blocks)	11	285.095	25.918	0.96	0.5499
Among Tests	7	215.169	30.738	1.14	0.4447
Among Controls	3	52.917	17.639	0.650	0.6092
Tests vs Controls	1	15.047	15.042	0.56	0.4834
Error	6	161.833	26.972		

R^2	CV	Root MSE	Yield Mean
0.800	6.372	5.194	81.500

Estimated standard errors of the estimated difference

- (i) between two controls is 4.24.
- (ii) between two tests in the same block is 7.34.
- (iii) between two tests in different blocks is 8.21.
- (iv) between a control and a test is 6.36.

We can use the adjusted values/means of the test treatments for comparison purpose. All those treatments for which yield levels are up to the satisfaction of breeder can be selected for further national level trials. In these kinds of experiments, generally, multiple characteristics

are observed. It may, therefore, be desirable to perform multivariate analysis of variance and use other related multivariate analytic techniques like cluster analysis, discriminant analysis, etc.

Keeping in view the importance of this design and for the ease of Biological Research Workers Agarwal and Sapra (1995) developed a user friendly program AUGMENT1 at the Documentation Unit of National Bureau of Plant Genetic Resources, New Delhi, to analyze the data of Augmented RCB design. It is interesting to note that the augmented RCB design is variance balanced with respect to tests *vs* controls comparisons.

7.4 Optimum replication of controls in a block

A survey of the literature reveals that generally the experiments described above are conducted using an augmented randomized complete block design. However, the experimenters would often like to know how many times the control treatments be replicated in each of the blocks so as to maximize the efficiency per observation for making test treatments *vs* control treatments(s) comparisons? An answer to this question was obtained by Parsad and Gupta (2000) and is described in the sequence.

Suppose that there are w test treatments which occur only once in the design and each of the u controls occurs in each of the b blocks, then to maximize the efficiency per observation the number of times each control appears in each of the blocks is

$$r = \frac{\sqrt{u+b-1}\sqrt{w}}{ub}$$

provided $b + u - 1 \leq w$. For example, consider the problem of obtaining the optimum number of

replications of the controls in an experiment with $w = 24$, $u = 3$, $b = 4$. We have $r = \frac{1}{3} \sqrt{\frac{6 \times 24}{4 \times 4}} = 1$.

Similarly, for $w = 98$, $u = 2$, $b = 7$, we have $r = \frac{1}{2} \sqrt{\frac{8 \times 98}{7 \times 7}} = 2$.

Remark 7.1 For a single control situation, *i.e.* $u = 1$, the above expression reduces to $r = \sqrt{\frac{w}{b}}$

and it can easily be seen that for $u = 1$, the condition $b + u - 1 \leq w$ becomes $b \leq w$, which is always true.

There may, however, arise many combinations of w , u and b for which the above expression of r does not yield a positive integer value of r . In such situations, a question that arises is as to what integer value of r should be taken? To answer this question, the efficiency per observation has been calculated for $w \leq 100$, $b \leq 25$ and $u \leq 10$ such that $b + u - 1 \leq w$ and r has been taken as $r^* = \text{int}(r)$ and $\text{int}(r) + 1$ besides taking $r = 1$. A close scrutiny reveals that if value of $r > \#.42$ then

take $r^* = \text{int}(r) + 1$ and for values of r smaller than or equal to #.42 take $r^* = \text{int}(r)$ for $u \geq 2$. For $u = 1$, the same rule applies but the value of r is taken as #.45 instead of #.42. Here # is the integral

$$\text{part of } r = \frac{\sqrt{u+b-1}\sqrt{w}}{ub}.$$

7.5 Statistical package for augmented designs

A user friendly, menu driven, graphic user interface (GUI) based Statistical Package called “Statistical Package for Augmented Design” (SPAD) has been developed at IASRI, New Delhi. The package generates randomized layout of augmented designs and performs the analysis of data generated. For given number of test treatments, number of control treatments and number of blocks, it computes the optimum replication number of each control treatment in every block of the design such that the efficiency per observation of the test treatments vs control treatment(s) comparisons is maximum. The user may choose the optimum replication number. However, the package provides flexibility in choosing the replication number of each control treatment in every block. The user can choose the replication of each control treatment in every block according to the resources available. It also asks the user to give the block sizes. One can have unequal block sizes as well. Once the user defines the number of test treatments, number of control treatments, and number of blocks in the design, the randomized layout of the design is generated. The package also provides the analysis of the data generated from augmented designs. A null hypothesis on any user-defined contrast can also be tested. This software is available at Design Resources Server. The URL is <http://www.iasri.res.in/design/AugmentedDesigns/home.htm>.

7.6 Example 2

An experiment was conducted at Directorate of Wheat Research during 2002-03 to compare 54 new accessions with 4 check varieties to see whether any of the check varieties can be replaced by any of the new accessions. The experiment was conducted using an augmented randomized complete block design with 6 blocks each of size 13 such that each of 4 check varieties are allocated in each of the six blocks and accessions are allocated only once in the design. The data on (i) days to 75% SE, (ii) FLL in centimeters and (iii) 1000 grain weight in grams is given in Table 7.4.

Table 7.4: Experiment data from an augmented RCB design

Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)	Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)
IC-028532	1	85	21.8	36.7	IC-079026	4	85	22.8	28.6
IC-028661	1	88	22.98	31.6	C-3	4	86	19.8	33.1
IC-028696	1	85	21	22.7	IC-079008	4	86	26.8	25.4

Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)	Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)
IC-028741	1	85	22.4	30.2	IC-079027	4	86	23.1	19.9
C-4	1	84	23.3	36	C-1	4	87	17.5	28.2
IC-028764	1	81	22.2	31.4	IC-079034	4	84	26.1	20.9
IC-028794	1	82	22.6	29.5	IC-079037	4	88	27.3	25.5
IC-028835	1	85	22.22	28.3	IC-079047	4	82	23.8	27.9
C-1	1	86	19.34	27.2	C-4	4	85	23.3	34.9
IC-028843	1	85	20.8	34.7	IC-079048	4	92	23.7	27
IC-028847	1	85	21	33.4	IC-079050	4	87	22.8	22.8
C-2	1	86	22.8	29.6	IC-079007	4	88	25.3	25.5
C-3	1	88	22.88	24.4	C-2	4	87	28.7	31.2
IC-036882	2	85	23.9	25.7	C-3	5	87	15.2	30.9
C-1	2	88	20.2	29.3	IC-082330	5	85	20.7	18.3
IC-036875	2	87	22.7	23.6	IC-082335	5	90	21.8	31.2
IC-042408	2	79	24.3	35.9	IC-082336	5	85	22.2	29
C-3	2	86	13.7	37.9	IC-082338	5	88	19.2	27.9
IC-036885	2	84	23.34	16.6	C-1	5	86	15.5	34.4
IC-041405	2	92	24.9	24.9	IC-082343	5	88	19.9	23.5
C-4	2	82	23.7	35.8	IC-082351	5	90	19.6	27.5
IC-036884	2	85	28.4	28.3	C-2	5	85	23.7	36.3
IC-042458	2	80	25.1	28.7	IC-082352	5	85	20.8	27.9
IC-036871	2	89	25.9	24.9	IC-082362	5	84	14.6	28.9
C-2	2	81	23.1	38.1	C-4	5	86	20.7	36.9
IC-042343	2	83	25.5	26.1	IC-082326	5	83	21.2	18.5
C-4	3	88	26.9	35.9	IC-104612	6	83	26.4	39.9
IC-060221	3	83	22.24	33.5	IC-104601	6	87	19.7	36.5
IC-073491	3	82	25.36	34.6	C-4	6	85	18.8	30.1
IC-063947	3	82	21.6	19.9	C-2	6	87	22	36.5
IC-066518	3	85	21.04	26.9	IC-104609	6	87	20.4	32.5
IC-073214	3	81	22.9	36.9	IC-104607	6	85	21.4	39.2
C-1	3	88	17.4	33.5	IC-104611	6	87	21.4	34.4
C-2	3	85	23.6	24.8	C-3	6	86	17.7	36.5
IC-073207	3	90	21.32	21.4	IC-104613	6	86	21.7	24.3
IC-073493	3	86	21.6	18.9	IC-104614	6	84	21.3	34.6
IC-060218	3	82	19.9	23	C-1	6	87	18.4	31.1

Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)	Accession No.	Block	Days to 75 % SE	FLL (cm)	1000 Grain Weight (gm)
C-3	3	88	18.2	21.6	IC-104610	6	88	22.8	29.2
IC-060997	3	83	20.52	22.9	IC-104604	6	87	23.1	24.8

In what follows, the data are analyzed (a) to test the homogeneity of all the 58 treatment effects, (b) to test the equality of (i) all check varieties, (ii) all new accessions, and (iii) all new accessions with all check varieties, (c) to make all possible pairwise comparisons of check varieties with each of the new accessions to identify the best performing accession.

Remark 7.2 Using the online package, the following augmented design may be generated. It is indeed possible to generate another randomized layout of this design. The optimum replication of control in each block works out to one in this case. C# denotes the label of the control treatment and T# denotes the label of the new accession.

B1: (T4, T24, T46, T13, T17, C4, T34, T23, C1, T3, C2, C3, T16)

B2: (T12, T1, C4, C3, T54, C1, T10, T49, T37, C2, T43, T25, T30)

B3: (T21, T50, C3, T26, T47, C2, T48, C4, T20, C1, T40, T28, T29)

B4: (C2, T53, T22, T39, T5, T52, T11, C3, T18, C4, C1, T6, T42)

B5: (T14, T27, C4, T31, T9, C1, T45, C3, C2, T32, T36, T38, T2)

B6: (C4, C2, T8, T35, C3, T41, T15, C1, T7, T51, T44, T33, T19)

Remark 7.3 For preparing the data file for SAS, the treatment labels have to be given as numerals 1,2,3, For the Example 2 in Section 7.6, the numerals 1, 2, 3, 4 denote the controls (or check varieties) and the numerals 5, 6, 7, 8, . . . , 54, 55, 56, 57, 58 denote the 54 tests (or new accessions). While writing down the contrasts also, this has to be borne in mind.

7.6.1 Analysis of data

The parameters of the augmented design are given as:

Number of tests, $w = 54$; Number of controls, $u = 4$; Block sizes, $k_1 = k_2 = k_3 = k_4 = k_5 = k_6 = 13$

Replication of controls, $r_0 = 6$, Replication of tests, $r = 1$, Total number of observations, $n = 78$. The data has been analyzed using SAS software. The commands and the data preparation are given in the sequel.

```
DATA augmented;
INPUT block trt SE FLL GW;
CARDS;
```

Statistical Analysis of Agricultural Experiments

1	1	86	19.34	27.2
1	2	86	22.8	29.6
1	3	88	22.88	24.4
1	4	84	23.3	36.0
1	5	85	21.8	36.7
1	6	88	22.98	31.6
1	7	85	21	22.7
1	8	85	22.4	30.2
1	9	81	22.2	31.4
1	10	82	22.6	29.5
1	11	85	22.22	28.3
1	12	85	20.8	34.7
1	13	85	21	33.4
2	1	88	20.2	29.3
2	2	81	23.1	38.1
2	3	86	13.7	37.9
2	4	82	23.7	35.8
2	14	85	23.9	25.7
2	15	87	22.7	23.6
2	16	79	24.3	35.9
2	17	84	23.34	16.6
2	18	92	24.9	24.9
2	19	85	28.4	28.3
2	20	80	25.1	28.7
2	21	89	25.9	24.9
2	22	83	25.5	26.1
3	1	88	17.4	33.5
3	2	85	23.6	24.8
3	3	88	18.2	21.6
3	4	88	26.9	35.9
3	23	83	22.24	33.5
3	24	82	25.36	34.6
3	25	82	21.6	19.9
3	26	85	21.04	26.9
3	27	81	22.9	36.9
3	28	90	21.32	21.4
3	29	86	21.6	18.9
3	30	82	19.9	23.0
3	31	83	20.52	22.9
4	1	87	17.5	28.2
4	2	87	28.7	31.2
4	3	86	19.8	33.1
4	4	85	23.3	34.9

4	32	85	22.8	28.6
4	33	86	26.8	25.4
4	34	86	23.1	19.9
4	35	84	26.1	20.9
4	36	88	27.3	25.5
4	37	82	23.8	27.9
4	38	92	23.7	27.0
4	39	87	22.8	22.8
4	40	88	25.3	25.5
5	1	86	15.5	34.4
5	2	85	23.7	36.3
5	3	87	15.2	30.9
5	4	86	20.7	36.9
5	41	85	20.7	18.3
5	42	90	21.8	31.2
5	43	85	22.2	29.0
5	44	88	19.2	27.9
5	45	88	19.9	23.5
5	46	90	19.6	27.5
5	47	85	20.8	27.9
5	48	84	14.6	28.9
5	49	83	21.2	18.5
6	1	87	18.4	31.1
6	2	87	22	36.5
6	3	86	17.7	36.5
6	4	85	18.8	30.1
6	50	83	26.4	39.9
6	51	87	19.7	36.5
6	52	87	20.4	32.5
6	53	85	21.4	39.2
6	54	87	21.4	34.4
6	55	86	21.7	24.3
6	56	84	21.3	34.6
6	57	88	22.8	29.2
6	58	87	23.1	24.8

```

;
ODS RTF FILE='Model1.rtf';
PROC GLM;
CLASS trt block;
MODEL SE FLL GW = trt block;
LSMEANS trt/PDIFF LINES;

```


Note: It may appear difficult to generate the contrasts for each experimental situation. Therefore, a SAS Macro has been developed where user has only to enter the data file and variable names and with that information all other steps are generated automatically. This macro is available at <http://www.iasri.res.in/sscnars/augblkdsjn.aspx>.

7.6.2 Output of analysis

The results obtained from the analysis are given in Table 7.5.

Table 7.5: Results for the character SE
ANOVA for the character “Days to 75 % SE”

Source	DF	Type III SS	MS	F-Value	Prob > F
Treatments	57	432.564	7.5889	3.28	0.0069
Among New Accessions	53	405.251	7.646	3.31	0.0068
Among Controls	3	20.333	6.778	2.93	0.0676
Controls vs New Accessions	1	6.980	6.980	3.02	0.1027
Blocks	5	19.000	3.800	1.64	0.2087
Error	15	34.667	2.311		
Corrected Total	77	507.295			

R-Square	CV	Root MSE	SE Mean
0.932	1.777	1.520	85.551

It may be noted that the model explains about 93 percent of the total variability in the data pertaining to “Days to 75 % SE.” The CV is also very low (CV = 1.78). The treatment effects are significantly different (p -value = 0.0069), but the block effects are not significant. The effect of new accessions is also significantly different (p -value = 0.0068), but the effects of controls and new accessions vs controls are not significantly different.

The pairwise treatment comparisons using LS MEANS produce Table 7.6.

Table 7.6: t comparison lines for least squares means of treatments

LS-means with the same letter are not significantly different								
							SE LSMEAN	LSMEAN of Treatment Number
				A			93.750	18
	B			A			91.750	38
	B			A		C	90.750	21
	B	D		A		C	90.000	42
	B	D		A		C	90.000	46

LS-means with the same letter are not significantly different									
								SE LSMEAN	LSMEAN of Treatment Number
E	B	D		A		C		88.750	15
E	B	D		A		C		88.750	28
E	B	D				C		88.000	45
E	B	D				C		88.000	44
E	B	D		F		C		88.000	6
E	B	D		F		C		87.750	57
E	B	D		F		C	G	87.750	40
E	B	D		F		C	G	87.750	36
E	B	D		F		C	G	87.000	1
E	B	D		F		C	G	86.833	3
E	B	D		F	H	C	G	86.750	19
E	B	D	I	F	H	C	G	86.750	39
E	B	D	I	F	H	C	G	86.750	51
E	B	D	I	F	H	C	G	86.750	52
E	B	D	I	F	H	C	G	86.750	14
E	B	D	I	F	H	C	G	86.750	54
E	B	D	I	F	H	C	G	86.750	58
E	J	D	I	F	H	C	G	85.750	17
E	J	D	I	F	H	C	G	85.750	33
E	J	D	I	F	H	C	G	85.750	55
E	J	D	I	F	H	C	G	85.750	34
E	J	D	I	F	H		G	85.167	2
E	J	D	I	F	H		G	85.000	5
E	J	D	I	F	H		G	85.000	11
E	J	D	I	F	H		G	85.000	13
E	J	D	I	F	H		G	85.000	47
E	J	D	I	F	H		G	85.000	8
E	J	D	I	F	H		G	85.000	41
E	J	D	I	F	H		G	85.000	4
E	J	D	I	F	H		G	85.000	43
E	J	D	I	F	H		G	85.000	12
E	J	D	I	F	H		G	85.000	7
E	J		I	F	H	K	G	84.750	32
E	J		I	F	H	K	G	84.750	53

LS-means with the same letter are not significantly different									
								SE LSMEAN	LSMEAN of Treatment Number
E	J		I	F	H	K	G	84.750	22
E	J		I	F	H	K	G	84.750	29
E	J		I	F	H	K	G	84.000	48
E	J		I	F	H	K	G	83.750	35
E	J		I	F	H	K	G	83.750	56
E	J		I	F	H	K	G	83.750	26
	J		I	F	H	K	G	83.000	49
	J		I		H	K	G	82.750	50
	J		I		H	K		82.000	10
	J		I		H	K		81.750	31
	J		I		H	K		81.750	37
	J		I		H	K		81.750	23
	J		I			K		81.750	20
	J					K		81.000	9
	J					K		80.750	16
	J					K		80.750	25
	J					K		80.750	24
	J					K		80.750	30
						K		79.750	27

The LINES display does not reflect all significant comparisons. The following additional pairs are significantly different: (18, 15) (38, 1) (38, 3) (38, 39) (21, 3) (21, 17) (42, 2) (42, 47) (42, 41) (42, 4) (42, 43) (46, 2) (46, 47) (46, 41) (46, 4) (46, 43) (28, 26) (1, 4) (1, 49) (1, 50) (3, 49) (3, 50) (39, 37) (14, 20) (17, 16) (2, 9) (2, 16) (2, 25) (2, 24) (2, 30) (4, 9) (4, 16) (4, 25) (4, 24) (4, 30) (29, 27)

Note: While interpreting the results care need to be taken to convert the treatment numbers back to new accessions and control varieties. The control varieties are labeled 1 – 4 and the new accessions are labeled 5 – 58. This means that treatment number 5 is actually new strain 1; treatment number 58 is actually new strain 54, and so on.

The estimated standard errors of various estimated comparisons can be obtained by using the online portal “Strengthening Statistical Computing for NARS” at www.iasri.res.in/sscnars/. The estimated standard errors are given below:

Estimated standard errors of the estimated difference

- (i) between two controls is 0.878 and Tukey’s HSD at 5 % is 6.172.
- (ii) between two new accessions in the same block is 2.150 and Tukey’s HSD at 5 % is 15.119.

- (iii) between two new accessions in different blocks is 2.404 and Tukey’s HSD at 5 % is 16.904.
- (iv) between a control and a new accession is 1.783 and Tukey’s HSD at 5 % is 12.536.

Table 7.7: Results for the character FLL
ANOVA for the character “FLL (cm)”

Source	DF	Type III SS	MS	F-Value	Prob > F
Treatments	57	425.265	7.461	1.26	0.3196
Among New Accessions	53	188.509	3.557	0.60	0.9116
Among Controls	3	179.234	59.745	10.10	0.0007
Control vs New Accessions	1	57.523	57.523	9.73	0.0070
Blocks	5	45.524	9.105	1.54	0.2366
Error	15	88.698	5.913		
Corrected Total	77	672.516			

R-Square	CV	Root MSE	SE Mean
0.868	11.067	2.432	21.972

It may be noted that the model explains about 87 percent of the total variability in the data pertaining to “FLL.” The CV is little high compared to the one obtained for the character “days to 75 % SE.” (CV = 11.067). The treatment effects are not significantly different (p -value = 0.3196); similarly the block effects are also not significant (p -value = 0.2366). The effect of new accessions is also not significantly different (p -value = 0.9116), but the effects of controls and controls vs new accessions are highly significant (p -values = 0.0007 and 0.0070, respectively).

Estimated standard errors of the estimated difference

- i) between two controls is 1.404.
- (ii) between two new accessions in the same block is 3.434.
- (iii) between two new accessions in different blocks is 3.845.
- (iv) between a control and a new accession is 2.851.

Note: While interpreting the results care need to be taken to convert the treatment numbers back to new accessions and control varieties. The control varieties are number 1 – 4 and the new accessions are numbered 5 – 58. This means that treatment number 5 is actually new strain 1; treatment number 58 is actually new strain 54, and so on.

**Table 7.8: Results for the character 1000 grain weight
ANOVA for the character “1000 grain weight (gm)”**

Source	DF	Type III SS	MS	F-Value	Prob > F
Treatments	57	1907.634	33.467	1.85	0.0946
Among New Accessions	53	1507.241	28.439	1.57	0.1694
Among Controls	3	74.508	24.836	1.37	0.2899
Controls vs New Accessions	1	325.884	325.884	17.98	0.0007
Blocks	5	144.933	28.987	1.60	0.2202
Error	15	271.817	18.121		
Corrected Total	77	2512.795			

R-Square	CV	Root MSE	GW Mean
0.892	14.582	4.257	29.192

It may be noted that the model explains about 89 percent of the total variability in the data pertaining to “1000 grain weight.” The CV is little high compared to the one obtained for the character “days to 75 % SE.” (CV = 14.582). The treatment effects are not significantly different (p -value = 0.0946); similarly the block effects are also not significant (p -value = 0.2202). The effect of new accessions is also not significantly different (p -value = 0.1694), similarly, the effect of controls is also not significantly different (p -value = 0.2899). However, the effect of controls vs new accessions is highly significant (p -value = 0.0007).

Estimated standard errors of the estimated difference

- (i) between two controls is 2.458.
- (ii) between two new accessions in the same block is 6.020.
- (iii) between two new accessions in different blocks is 6.731.
- (iv) between a control and a new accession is 4.992.

7.6.3 Analysis using R

In the sequence are give the R code for analysis of data generated from an augmented design. The results obtained are not given to avoid duplication.

R code

```
d12=read.table(“augmented.txt”,header=TRUE)
attach(d12)
names(d12)
#anova
trt=factor(trt)
block=factor(block)
```

```
lm1=lm(SE~trt+block)
#anova(lm1)
library(car)
Anova(lm1,type="III")
library(lsmeans)
lsm=lsmeans(lm1,"trt")
#to provide letters for groups, install and then load multcompView
library(multcompView)
cld(lsm,Letters="abcdefghij")
#generating the contrasts, trts 1-4 are the controls and 5-58 are new accessions
contrast.mat1=matrix(0,53,58)
for (i in 1:53)
{
  contrast.mat1[i,(5:(4+i))]=1
  contrast.mat1[i,(5+i)]=-i
}
contrast.mat2=matrix(0,3,58)
for (i in 1:3)
{
  contrast.mat2[i,1]=1
  contrast.mat2[i,(1+i)]=-1
}
controls.vs.newaccessions=contrast(lsm,list(con1=c(rep(-27,4),rep(2,54))))
Among.New.Accessions=contrast(lsm,list(apply(contrast.mat1,1,list)))
Among.controls=contrast(lsm,list(apply(contrast.mat2,1,list)))
lht(lm1,Among.New.Accessions@linfct)
lht(lm1,Among.controls@linfct)
lht(lm1,controls.vs.newaccessions@linfct)
lm2=lm(FLL~trt+block)
Anova(lm2,type="III")
lsm2=lsmeans(lm2,"trt")
controls.vs.newaccessions=contrast(lsm2,list(con1=c(rep(-27,4),rep(2,54))))
Among.New.Accessions=contrast(lsm2,list(apply(contrast.mat1,1,list)))
Among.controls=contrast(lsm2,list(apply(contrast.mat2,1,list)))
lht(lm2,Among.New.Accessions@linfct)
lht(lm2,Among.controls@linfct)
lht(lm2,controls.vs.newaccessions@linfct)
cld(lsm2,Letters="abcdefghij")
lm3=lm(GW~trt+block)
Anova(lm3,type="III")
lsm3=lsmeans(lm3,"trt")
controls.vs.newaccessions=contrast(lsm3,list(con1=c(rep(-27,4),rep(2,54))))
Among.New.Accessions=contrast(lsm3,list(apply(contrast.mat1,1,list)))
```

```

Among.controls=contrast(lsm3,list(apply(contrast.mat2,1,list)))
lht(lm3,Among.New.Accessions@linfct)
lht(lm3,Among.controls@linfct)
lht(lm3,controls.vs.newaccessions@linfct)
cld(lsm3,Letters="abcdefghij")
detach(d12 )

```

Remark 7.4 This Chapter has focused essentially on augmented design or Category A designs in which the test treatments have single replication. Category B designs are the ones in which both the test treatments and control treatments are replicated. In Category B a standard design (RCB design, BIB design, Latin Square, nested, etc) in test treatments is supplemented with the additional control treatments. Generally all the controls appear together. The analysis of such designs has been described at several places in the book. It is for this reason that the analysis of these designs has not been discussed separately in this Chapter. For instance in Section 2.3.2 in Chapter 2, the Example described is actually a Category B design. The analysis steps are almost same as category A experiments or augmented randomized complete block designs except that formulae for standard error of pairwise comparisons for tests, controls and test *vs* controls will change as per the design adopted.