

Chapter 3

Complete Randomised Block Design



Complete Randomised Block Design

Felix Kutsanedzie^{1*}; Sylvester Achio¹;

Edmund Ameko¹; Victoria Ofori²;

Diaba Kwasi Selassie³

¹Accra Polytechnic, GP 561, Accra, Ghana

²Agricultural Engineering Department, KNUST, Ghana

³Anglican University College of Technology, Nkorazan, Sunyani, Ghana

Abstract

Complete Randomised Design (CRD) is used for experiments in which the experimental materials or units are homogeneous. However most experimental materials in life are heterogeneous in nature, hence the need to employ Complete Randomized Block Designs (CRBD) in designing and analysis of experiments with such nature. It is the most used of all the types of design, hence this chapter explains the underlying conditions and how it is used in analyzing experiments comprehensively for researchers and would-be experiment designers.

Keywords

Block, Homogeneous, Heterogeneous, Analysis, Randomized

3.1 Introduction

So long as researchers or individuals continue to design experiments to explain mechanisms, phenomenon etc. using heterogeneous experimental materials, the Complete Randomized Block Design (RCBD) would be required for analyses. Whereas Complete Randomized Design (CRD) is used appropriate for the analysis for experiments which involves the use of homogenous experimental materials, RCBD is used for experiments that require the use of heterogeneous materials. Generally when using animals as experimental materials in an experiment, when they are of the same species, they can be considered as homogeneous but when they are of different species then they are considered as heterogeneous. Assuming a researcher wants to use a plot of land to serve an experimental material for a study which involves the sowing of seeds to ascertain the germination rate, the plot of land needs to be divided into blocks, because it cannot be taken as a homogeneous unit due to the variability of nutrients at each point of the land. Thus the use of RCBD requires that when experimental materials which are not homogenous are to be used in an experiment, they must be divided into subgroups which are similar in nature and referred to as blocks or replicates before the design is employed for its analysis. The import of reducing the heterogeneous experimental materials to blocks or replicates is to make sure variations are minimized or reduced as much as possible so that all variations existing would be due to variability in the treatments applied.

In RCBD, randomization of treatments is done such that every block is restricted to a single treatment. However it should be noted that based on randomization designs are partially classified as follows: homogeneity of experimental material – Complete Randomized Design (CRD); heterogeneity of experimental material – Randomized Complete Block Design, that is a single restriction of treatment; Latin Square Design (LSD) and Cross Over Designs, designs with two restrictions of

treatment; Graeco-Latin Square Design, design with more than two restriction of treatment allocation; Incomplete Block Design (IBD) includes those not grouped into replications and those grouped into replications.

With RCBD, there is at least a single restriction of treatments per block, the treatment are randomly allocated at least once for each replicate or block. Also treatments are randomized separately for each block and have equal probability of being allocated to any experimental unit per block or replicate.

3.2 Merits of RCBD over CRD

There is preciseness of RCBD over CRD.

The species or objects experimental materials generally thought to be homogeneous though it may not be necessarily as they may differ in one way or the other when carefully examined. This exposes the idea of homogeneity of material in the case of CRD as a flaw should differences exist between the experimental materials or units. However the RCBD reduces this flaw as it is intended to cater for the heterogeneity of experimental materials. As regard restriction of treatment, there is at least a single restriction, thus is every treatment is expected to be allocated at least once per block or replicate.

3.3 Illustration of Randomized Complete Block Design

Assuming a researcher wants to design an experiment to manage waste materials such coconut fruit waste (shell and fibre), palm nut shells, waste plastic bottles, waste plastic water sachets and waste plastic packaging bags through pyrolysis for the recovery of other usable products and by-products. A design of such nature is to ascertain whether the amount of products such as oil,

gas or char produced subjected to different treatments (different weights of the various materials used) on equal weight per weight basis are significant or not. The design below is an illustration RCBD.

Treatments (weights of the waste types) to be Pyrolysed				
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]
Coconut waste fruits (shell and fibre) [A]	2.3	5.0	11.2	15.7
Palm nut shell [B]	1.6	4.8	9.2	14.0
Waste plastic bottles [C]	4.4	9.2	14.4	19.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2
Waste plastic bags [E]	4.1	8.5	11.9	16.9

The table above is a design for five different treatments in terms of weights of the various waste types being considered in the study. The data being considered is the weights of the oil generated from these waste types. It should however be noted that the values fielded in the table are all being assumed for explanation and not real.

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	2.3	5.0	11.2	15.7	34.2
Palm nut shell [B]	1.6	4.8	9.2	14.0	29.2
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	11.9	16.9	41.4
Treatment Totals	16.6	36.4	60	84.1	197.1

The first step: Stating of hypothesis

For Blocks:

H_0 : All the means of the blocks are equal

H_A : At least one of the means of the blocks (blk) is different

H_0 : $\mu_{blk_A} = \mu_{blk_B} = \mu_{blk_C} = \mu_{blk_D} = \mu_{blk_E}$

H_1 : $\mu_{blk_A} \neq \mu_{blk_B} = \mu_{blk_C} = \mu_{blk_D} = \mu_{blk_E}$

For Treatments:

H_0 : All the means of the treatments are equal

H_A : At least one of the means of the treatments (t) is different

H_0 : $\mu_{trt_A} = \mu_{trt_B} = \mu_{trt_C} = \mu_{trt_D} = \mu_{trt_E}$

H_1 : $\mu_{trt_A} \neq \mu_{trt_B} = \mu_{trt_C} = \mu_{trt_D} = \mu_{trt_E}$

The second step: Calculating the Correction Factor (CF)

$$CF = \frac{Y^2}{TB}$$

where Y^2 = square of the sum of all observations

where TB is the product of the number of treatments and number of blocks or number of replicates (r) when replicates are used rather instead of the blocks

$$CF = \frac{(2.3 + 5 + 11.2 \dots \dots + 16.9)^2}{4 \times 5}$$

$$CF = \frac{(197.1)^2}{4 \times 5}$$

$$CF = \frac{38848.41}{20}$$

$$CF = 1942.42$$

The third step: calculating the Total Sum of Squares (TSS)

$$TSS = \sum Y_{ij}^2 - CF$$

$$TSS = \sum (2.3^2 + 5.0^2 + 11.2^2 \dots \dots + 16.9^2) - 1942.42$$

$$TSS = 2513.57 - 1942.42 = 571.15$$

Y_{ij}^2 is the square of all observations.

The fourth step: calculating the Block Sum of Squares (BLKSS)

$$BLKSS = \sum \frac{BLKSS^2}{nTrt} - CF$$

$$BLKSS = \sum \frac{34.2^2 + 29.6^2 + 47.3^2 \dots + 41.4^2}{4} - 1942.42$$

$$BLKSS = \frac{1169.64 + 876.16 + 2237.29 + \dots + 1713.96}{5} - 1942.42$$

$$BLKSS = \frac{7986.21}{4} - 1942.42$$

$$BLKSS = 1996.55 - 1942.42$$

$$BLKSS = 54.13$$

The fifth step: calculating the Treatment Sum of Squares (TRTSS)

$$TRTSS = \sum \frac{TRTSS^2}{nBLK} - CF$$

$$TRTSS = \sum \frac{16.6^2 + 36.4^2 + \dots + 84.1^2}{5} - 1942.42$$

$$TRTSS = \frac{275.56 + 1324.96 + 3600 \dots \dots + 7072.81}{5} - 1942.42$$

$$\text{TRTSS} = \frac{12273.33}{5} - 1942.42$$

$$\text{TRTSS} = 2454.67 - 1942.42$$

$$\text{TRTSS} = 512.25$$

The sixth step: calculating the Error Sum of Squares (ERRSS)

$$\text{ERRSS} = \text{TSS} - (\text{BLKS} + \text{TRTSS})$$

$$\text{ERRSS} = 571.15 - (54.13 + 512.25)$$

$$\text{ERRSS} = 571.15 - 566.38$$

$$\text{ERRSS} = 4.77$$

Where nTrt = Number of Treatment, nBLK = Number of Blocks.

The seventh step: completing the ANOVA Table

ANOVA TABLE					
Sources of Variation	df	SS	MS	Fcal.	Fcrit.
block	blk - 1 = 5-1 = 4	54.13	13.53	33.83	
treatment	trt - 1 = 4 -1 = 3	512.25	170.75	426.88	
error	(blk-1)(trt -1) = 4 x 3 = 12	4.77	33.67		
total	blk x trt - 1 = 20 -1 = 19	571.15			

For block: Fcrit (5%) @ df of 4, 12 = 3.26; Fcrit (1%) @ df of 4, 12 =5.41.

For treatment: Fcrit (5%) @ df of 3, 12 =3.49; Fcrit (1%) @ df of 3, 12 = 5.95.

The tenth step: looking up the F-critical table to find the F-critical values:

For the block:

Fcrit. or tab. (5%) at df of 4,12 =3.26

Fcrit. or tab. (1%) at df of 4,12 =5.41

For the treatment:

Fcrit. or tab. (5%) at df of 3,12 = 3.49

$$F_{crit. \text{ or tab. } (1\%) \text{ at } df \text{ of } 3,12} = 5.95$$

The eighth step: making the decision or conclusion:

Decision on block:

$$F_{cal. (33.83)} > (3.26), F_{crit. \text{ or tab. } (5\%) \text{ at } df \text{ of } 4,12}$$

$$F_{cal. (33.83)} > (5.41), F_{crit. \text{ or tab. } (1\%) \text{ at } df \text{ of } 4,12}$$

At both 5% and 1% levels of significance for the block, we reject the null hypotheses because the F-calculated values are greater than the F-critical values; thus we conclude that there is enough evidence provided by the data collected in the experiment to reject the null hypotheses. It thus implies that the blocks are significantly different, hence there was a need to block or in other words blocking can be justified. It means RCBD is the right design adopted for the study.

Decision on treatment:

$$F_{cal. (426.88)} > (5.95), F_{crit. \text{ or tab. } (1\%) \text{ at } df \text{ of } 3,12}$$

At 1%, since the F-calculated value is greater than the F-critical value, we reject the null hypothesis and conclude that the data collected provides enough evidence that the treatments are significantly different.

$$F_{cal. (426.88)} > (3.49), F_{crit. \text{ or tab. } (5\%) \text{ at } df \text{ of } 3,12}$$

Again at 5% level of significance for the treatment, we reject the null hypothesis on the basis that the F-calculated value is greater than the F-critical value and thus conclude that the treatments are significantly different, therefore the need to find out the treatments that are significantly different.

3.3.1 Finding the Treatments that are Significantly Different

Once the test reveals that there exist significance differences between the treatments, we proceed to find out which of the treatments significantly differ. This procedure of finding out is referred to as the *pair comparisons*.

Since there are four treatments and we are to do a pair comparison, it presupposes that we will have C_2^4 i.e. 4 combination 2 which means there would be 6 paired comparisons possible. These are as follow:

AB, AC, AD, BC, BD and CD. There are two statistical methods used in doing the pair comparisons, these are namely: the Least Significance Difference Test (LSD) and the Duncan Multiple Range Test (DMRT).

3.3.2 The Fisher's LSD Test

With regards to the LSD (Least Significant Difference) test, an LSD value is calculated at a prescribed level of significance either 5% or 1%, which serves as a boundary for the classification between whether one treatment is significantly or not significantly different from another when their means are compared. This means that if the means difference of any two treatments compared exceed the LSD computed at a prescribed significance level, then we conclude that the two treatments are significantly different or otherwise not. It is used or valid when used for independent (orthogonal) comparison and used when the treatment size is less i.e. less than six treatments. The generalized version of LSD is given by the formula:

$$LSD = \frac{|\bar{x}_i - \bar{x}_j|}{\sqrt{wms \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}$$

$$\text{However } sd = \left(\sqrt{WMS^2 \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \right)$$

Since $n_i = n_j$,

$$sd = \left(\sqrt{WMS^2 \left(\frac{2}{n} \right)} \right)$$

$$sd = \sqrt{\frac{2WMS^2}{n}}$$

where WMS = Within (Residual or unexplained) Mean Square, n =number of sample per each treatment or replication; sd =standard deviation.

$$LSD_{\alpha} = t_{\alpha} \times sd$$

However, $t_{\frac{\alpha}{2}}$ is used to in order to a two sided hypothesis, therefore

$$LSD_{\alpha} = t_{\frac{\alpha}{2}} \times sd$$

Taking the example being handled under the RCBD design for instance, we used the following data:

Treatments (weights of the waste types) to be Pyrolysed				
Blocks	A	B	C	D
Coconut waste fruits (shell and fibre)	2.3	5	11.2	15.7
Palm nut shell	1.6	4.8	9.2	14
Waste plastic bottles	4.4	9.2	14.4	19.3
Waste plastic satchets	4.2	8.9	13.3	18.2
Waste plastic bags	4.1	8.5	11.9	16.9
Treatment Means	3.32	7.28	12	16.82

$$LSD_{\alpha} = t_{\alpha} \times sd$$

$$sd = \sqrt{\frac{2WMS^2}{n}}$$

WMS from the ANOVA table = 33.67; $n = 4$; Error or residual or unexplained $df = 12$

$$sd = \sqrt{\frac{2(33.67)^2}{5}}$$

$$sd = \sqrt{\frac{2 \times 1133.89}{5}}$$

$$sd = \sqrt{\frac{2267.79}{5}}$$

$$sd = \sqrt{453.56}$$

$$sd = 21.30$$

We therefore proceed to read the t-critical value from the two-tailed table at the significance level at which the design revealed that the treatments were significant (5% or 0.05).

Therefore from the two tailed t-critical at 5% level of significance table we obtain:

$$\frac{t_{0.05}}{2} = t_{0.025}, df \text{ of the residual or unexplained} = 2.18$$

$$t - \text{crit}(\text{two} - \text{tailed}) \text{ at } df 12 = 2.18$$

$$LSD_{\alpha} = 2.18 \times 21.30$$

$$LSD_{0.05} = 46.43$$

The value t-crit (two-tailed) at $df = 12$ can be read from the $\frac{t_{0.05}}{2} = t_{0.025}$ critical table.

α (1 tail)	0.05	0.025
α (2 tail)	0.1	0.05
df		
1	6.3138	12.7065
2	2.9200	4.3026
3	2.3534	3.1824
4	2.1319	2.7764
5	2.0150	2.5706
6	1.9432	2.4469
7	1.8946	2.3646
8	1.8595	2.3060
9	1.8331	2.2621
10	1.8124	2.2282
11	1.7959	2.2010
12	1.7823	2.1788
13	1.7709	2.1604

Means Difference Table				
	A(\bar{x}_A)	B(\bar{x}_B)	C(\bar{x}_C)	D(\bar{x}_D)
A (\bar{x}_A)	-			
B (\bar{x}_B)	$(\bar{x}_B - \bar{x}_A)$	-		
C (\bar{x}_C)	$(\bar{x}_C - \bar{x}_A)$	$(\bar{x}_C - \bar{x}_B)$	-	
D (\bar{x}_D)	$(\bar{x}_D - \bar{x}_A)$	$(\bar{x}_D - \bar{x}_B)$	$(\bar{x}_D - \bar{x}_C)$	-

Means difference table				
	A (3.32)	B (7.28)	C(12.00)	D(16.82)
A (3.32)	-			
B(7.28)	3.96	-		
C(12.00)	8.68	4.72	-	
D(16.82)	13.5	9.54	4.82	-

Since all the mean differences for all the treatments are (A, B, C, D) are less than (<) the LSD value obtained (45.39), hence the LSD test confirms they are not significantly different.

$$(\bar{x}_B - \bar{x}_A), (\bar{x}_C - \bar{x}_A), (\bar{x}_D - \bar{x}_A), (\bar{x}_C - \bar{x}_B), (\bar{x}_D - \bar{x}_B), (\bar{x}_D - \bar{x}_C) < LSD_{0.05}$$

The Duncan Multiple Range Test

This test is also used for paired comparisons for a larger treatment size than the LSD. With the DMRT, the sample means of the treatments are ranked from the lowest to the highest and then the steps apart denoted by (r) is derived and used with the total degree of freedom (df) to read the q tabulated value from the studentized range table.

When using the DMRT, two population means are significantly different if the absolute value of their sample differences exceed W, where W is defined as below:

$$W = q(r, df \text{ of residual or unexplained}) \times \sqrt{\frac{WMS}{n}}$$

Where n =number of samples or observations per treatment group, r =the number of steps from the lowest treatment mean to the highest treatment mean when the treatment means are arranged in ascending order, df =degree of freedom of the residual or unexplained or error, WMS =within residual or unexplained or error mean square derived from the ANOVA table.

Arrangement of the treatment means in ascending order to determine the steps is done as follows:

$$\bar{x}_A = 3.32 < \bar{x}_B = 7.28 < \bar{x}_C = 12.0 < \bar{x}_D = 16.82$$

Since the distance between \bar{x}_A and \bar{x}_D is 4, thus moving from \bar{x}_A to \bar{x}_D is 4 steps then \bar{x}_D has $r = 4$, \bar{x}_C has $r = 3$, \bar{x}_B has $r = 2$

Therefore the table below can be constructed to aid the calculation of W:

	\bar{x}_B	\bar{x}_C	\bar{x}_D
	$r = 2$	$r = 3$	$r = 4$
q(r, df of error = 12)	3.082	3.773	4.199
$W = q(r, df\ of\ error = 12) \times \sqrt{\frac{WMS}{n}}$	$3.082 \times \sqrt{\frac{33.67}{5}}$	$3.773 \times \sqrt{\frac{33.67}{5}}$	$4.199 \times \sqrt{\frac{33.67}{5}}$
W	7.98	9.77	10.88

The values of q can be read from the critical values of the Studentized Range (0.05) presented.

Critical Values of the Studentized Range (0.05 level)

dfc	2	3	4	5	6	7	8	9
2	6.0849	8.3308	9.7980	10.8810	11.7340	12.4345	13.0266	13.5381
3	4.5007	5.9096	6.8245	7.5016	8.0370	8.4780	8.8521	9.1766
4	3.9265	5.0403	5.7571	6.2870	6.7065	7.0528	7.3465	7.6015
5	3.6354	4.6017	5.2185	5.6731	6.0329	6.3299	6.5823	6.8014
6	3.4605	4.3390	4.8956	5.3049	5.6285	5.8953	6.1222	6.3192
7	3.3439	4.1648	4.6812	5.0601	5.3591	5.6058	5.8154	5.9975
8	3.2612	4.0410	4.5288	4.8858	5.1672	5.3991	5.5962	5.7673
9	3.1991	3.9485	4.4149	4.7554	5.0235	5.2444	5.4319	5.5947
10	3.1511	3.8768	4.3266	4.6543	4.9120	5.1242	5.3042	5.4605
11	3.1127	3.8195	4.2561	4.5736	4.8229	5.0281	5.2021	5.3531
12	3.0813	3.4428	4.1985	4.5076	4.7477	4.9469	5.1159	5.2625
13	3.0553	3.7341	4.1509	4.4529	4.6897	4.8841	5.0490	5.1920
14	3.0332	3.7014	4.1105	4.4066	4.6385	4.8290	4.9903	5.1300

We can now proceed to conclude based on the computation of the W-value and the mean difference:

Means difference table				
	A (3.32)	B (7.28)	C(12.00)	D(16.82)
A (3.32)	-			
B(7.28)	3.96	-		
C(12.00)	8.68	4.72	-	
D(16.82)	13.5*	9.54	4.82	-

$(\bar{x}_B - \bar{x}_A) = 3.96 < 7.98$, there is no significant difference between A and B.

$(\bar{x}_C - \bar{x}_A) = 8.68 < 9.77$, there is no significant difference between C and A.

$(\bar{x}_C - \bar{x}_B) = 4.72 < 9.77$, there is no significant difference between C and B.

$(\bar{x}_D - \bar{x}_A) = 13.5 > 10.88$, there is significant difference between D and A.

$(\bar{x}_D - \bar{x}_B) = 9.54 < 10.88$, there is no significant difference between D and B.

$(\bar{x}_D - \bar{x}_C) = 4.82 < 10.88$, there is no significant difference between D and C.

3.4 Missing Data Handling

There is bound to be some data missing while conducting the experiment based on accidents such as breakage, death of an animal, spilling of a substance, destruction of a treatment on the experimental material by any extraneous subject or object (human or animal). When the experiment is started and these accidents occur, then the data on some experimental units cannot be obtained and thus referred to as 'missing data'. The experiment cannot be halted but continued and the missing data estimated after the experiment. In order to estimate the missing data the formula below can be used:

$$Y_{ij} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

where Y_{ij} =the missing data value, r =number of replicates or blocks, B =block or replicate total for block or replicate with the missing data, t =number of treatment, T =treatment total for treatment with missing data, $G = \sum_{ij}^n Y_{ij}$ = grand total of experimental units or observation units.

3.4.1 Handling a Single Missing Data

Therefore using the same experiment presented for the RCBD design with one of the data taken out as a missing data in the table below, the missing data value is estimated as follows:

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	2.3	5.0	11.2	15.7	34.2
Palm nut shell [B]	1.6	4.8	9.2	14.0	29.2
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	Y_{ij}	13.3	18.2	35.7
Waste plastic bags [E]	4.1	8.5	11.9	16.9	41.4
Treatment Totals	16.6	27.5	60	84.1	188.2

$$Y_{ij} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

$$r = 5, t = 4, B = 44.6, T = 36.4, G = 188.2$$

$$Y_{ij} = \frac{(5 * 35.7 + 4 * 27.5 - 188.2)}{(5 - 1)(4 - 1)}$$

$$Y_{ij} = \frac{(178.5 + 110 - 188.2)}{(4)(3)}$$

$$Y_{ij} = \frac{288.5 - 188.2}{12}$$

$$Y_{ij} = \frac{100.3}{12}$$

$$Y_{ij} = 8.4$$

where r =number of replication, t =number of treatments, B =Block or replicate total with missing data, T =Treatment total with missing data, G = grand total of all observations.

After the estimation of the missing data, the bias is computed using the formulae:

$$\beta_1 = \frac{[B_0 - (t - 1)x]^2}{t(t - 1)}$$

$$\beta_1 = \frac{[35.7 - (4 - 1)8.4]^2}{4(4 - 1)}$$

$$\beta_1 = \frac{[35.7 - (3)8.4]^2}{4(3)}$$

$$\beta_1 = \frac{[35.7 - 25.2]^2}{12}$$

$$\beta_1 = \frac{[10.5]^2}{12}$$

$$\beta_1 = \frac{(110.25)}{12}$$

$$\beta_1 = 9.2$$

Therefore the estimated bias $(\beta_1) = 9.2$

Adjustment in the Analysis of using the bias:

When the estimated bias has been computed, the adjustment of the analysis is done by subtracting the bias value from only the TSS (Total Sum of Squares) i.e.

$[571.15 - 9.2 = 561.95]$ and TrtSS (Treatment Sum of Squares) i.e. $[512.25 - 9.2 = 503.05]$

Once the missing data value is estimated, it can be put in the table in order to complete the ANOVA table. It must however be noted that one df (degree of freedom) is lost from the error i.e. $[((5-1)(4-1) - 1) = (12 - 1) = 11]$ and also the total degrees of freedom because of one missing data. Therefore the ANOVA would look as below:

Complete Anova Table					
Sources of Variation	df	SS	MS	Fcal.	Fcrit.
Block	$\text{blk} - 1 = 5 - 1 = 4$	54.13	13.53	31.21	
Treatment	$\text{trt} - 1 = 4 - 1 = 3$	503.05	167.68	386.69	
Error	$(\text{blk} - 1)(\text{trt} - 1) - 1 = [4 \times 3] - 1 = 11$	4.77	0.43		
Total	$\text{blk} \times \text{trt} - 1 - 1 = 20 - 1 - 1 = 18$	561.95			

From the ANOVA table, the decisions to be taken on the blocks and treatments remains the same, after the missing data had been computed and the ANOVA table completed for the data.

3.4.2 Handling More than One Missing Data Under RCBD

Assuming there are more than one missing data or values to be estimated under the RCBD, how this can be done is explained using the RCBD table below:

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	Y_{11}	5.0	11.2	15.7	31.9
Palm nut shell [B]	1.6	4.8	9.2	Y_{24}	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	Y_{43}	16.9	29.5
Treatment Totals	14.3	36.4	48.1	70.1	168.9

From the above table, it can be seen that three data values (Y_{11} , Y_{24} and Y_{43}) are conspicuously missing and which must be estimated in order to complete the table for the analysis of the data recorded.

The first procedure adopted in estimating the values is to use the formula given below to find the average estimated values of the missing data in the design:

$$Y_{ij} = \frac{(T_m + B_m)}{2}$$

This formula can apply only to RCBD when estimating average missing data.

Where Y_{ij} = estimated average missing data value, T_m = the mean value of treatment with the missing data, B_m = the mean value of treatment with the missing data.

Thus applying this formula, one can obtain the estimated average values of the missing data: Y_{11} , Y_{24} and Y_{43} .

$$Y_{ij} = \frac{(T_m + B_m)}{2}$$

$$\text{For } Y_{11} = \frac{(T_m + B_m)}{2}$$

$$T_m = \frac{14.3}{5} = 2.86, B_m = \frac{31.9}{4} = 7.98$$

$$Y_{11} = \frac{(2.86 + 7.98)}{2} = \frac{10.84}{2} = 5.42$$

$$\text{For } Y_{24} = \frac{(T_m + B_m)}{2}$$

$$T_m = \frac{70.1}{5} = 14.02, B_m = \frac{15}{4} = 3.75$$

$$Y_{24} = \frac{(14.02 + 3.75)}{2} = \frac{17.77}{2} = 8.89$$

$$\text{For } Y_{43} = \frac{(T_m + B_m)}{2}$$

$$T_m = \frac{48.1}{5} = 9.62, B_m = \frac{29.5}{4} = 7.38$$

$$Y_{43} = \frac{(9.62 + 7.38)}{2} = \frac{17.0}{2} = 8.50$$

Now that we have estimated the averages of missing data for Y_{11} , Y_{24} and Y_{43} , two of these estimated values can be substituted into the table for the formula given below to be used:

$$Y_{ij} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

for computing the estimated values of each of the missing data in turns, and replacing the average estimated values with the new computed values until no change in the new computed values occur. When no change occurs in the estimated values, it means the accurate missing data values have been found.

Using the stated procedures, we proceed to apply them.

Complete the table with the average estimated values (italicized in the table) leaving only the one whose estimated value one wants to determine first and in this case we start with Y_{11}

Blocks	Treatments (weights of the waste types) to be Pyrolysed				Block Totals
	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	
Coconut waste fruits (shell and fibre) [A]	Y_{11}	5.0	11.2	15.7	31.9
Palm nut shell [B]	1.6	4.8	9.2	8.89	24.49
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	8.50	16.9	38.0
Treatment Totals	14.3	36.4	56.6	78.99	186.29

Compute for the estimated value of Y_{11} using the formula given as:

$$Y_{ij} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

$$Y_{11} = \frac{(5(31.9) + 4(14.3) - 186.29)}{(5 - 1)(4 - 1)}$$

$$Y_{11} = \frac{((159.5 + 57.2) - 186.29)}{(4)(3)}$$

$$Y_{11} = \frac{(216.70 - 186.29)}{12}$$

$$Y_{11} = \frac{30.41}{12} = 2.53$$

Therefore substitute the value of $Y_{11} = 2.53$ into the table and compute for the estimated value of the next missing data, Y_{24} . The new table is given below:

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	2.53	5.0	11.2	15.7	34.43
Palm nut shell [B]	1.6	4.8	9.2	Y_{24}	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	8.50	16.9	38.0
Treatment Totals	16.83	36.4	56.6	70.1	179.93

$$Y_{ij} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 179.93)}{(5 - 1)(4 - 1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 179.93)}{(4)(3)}$$

$$Y_{24} = \frac{((78 + 280.4) - 179.93)}{12}$$

$$Y_{24} = \frac{((358.4 - 179.93))}{12}$$

$$Y_{24} = \frac{178.47}{12} = 14.87$$

The estimated missing data value for $Y_{24} = 14.87$. This value is then be substituted into the table for the subsequent estimation of the value for the missing data (Y_{43}).

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	2.53	5.0	11.2	15.7	34.43
Palm nut shell [B]	1.6	4.8	9.2	14.9	30.47
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	Y_{43}	16.9	29.5
Treatment Totals	16.83	36.4	48.1	84.97	186.3

$$Y_{ij} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

$$Y_{43} = \frac{(5(29.5) + 4(48.1) - 186.3)}{(5 - 1)(4 - 1)}$$

$$Y_{43} = \frac{((147.5 + 192.4) - 186.3)}{(4)(3)}$$

$$Y_{43} = \frac{(339.9 - 186.3)}{12}$$

$$Y_{43} = \frac{(339.9 - 186.3)}{12}$$

$$Y_{43} = \frac{(153.6)}{12} = 12.8$$

The estimated value for the missing data $Y_{43} = 12.8$. This is therefore substituted in the table and the value of the first estimated value Y_{11} taken out from the table and freshly computed for to ascertain whether the value will change or not.

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	Y_{11}	5.0	11.2	15.7	31.9
Palm nut shell [B]	1.6	4.8	9.2	14.9	30.47
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	12.8	16.9	42.3
Treatment Totals	14.3	36.4	60.9	84.97	196.57

$$Y_{ij} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

$$Y_{11} = \frac{(5(31.9) + 4(14.3) - 196.57)}{(5 - 1)(4 - 1)}$$

$$Y_{11} = \frac{((159.5 + 57.2) - 196.57)}{12}$$

$$Y_{11} = \frac{(216.70 - 196.57)}{12}$$

$$Y_{11} = \frac{20.13}{12} = 1.7$$

Since Y_{11} has changed from 2.53 to 1.7, it means one must continue computing until constant values are obtained.

So we continue to compute for Y_{24} by replacing the value of $Y_{11} = 1.7$

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	1.7	5.0	11.2	15.7	33.6
Palm nut shell [B]	1.6	4.8	9.2	Y_{24}	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	12.8	16.9	42.3
Treatment Totals	16.0	36.4	60.9	70.1	183.4

$$Y_{ij} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 183.4)}{(5 - 1)(4 - 1)}$$

$$Y_{24} = \frac{((78 + 280.4) - 183.4)}{(4)(3)}$$

$$Y_{24} = \frac{(358.4 - 183.4)}{12}$$

$$Y_{24} = \frac{175}{12} = 14.58$$

The new value for $Y_{24} = 14.58$, which has also changed from 14.9 to 14.58, so we substitute this value in the table:

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	1.7	5.0	11.2	15.7	33.6
Palm nut shell [B]	1.6	4.8	9.2	14.58	30.18
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	Y_{43}	16.9	29.5
Treatment Totals	16.0	36.4	48.1	84.68	185.18

$$Y_{ij} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

$$Y_{43} = \frac{(5(29.5) + 4(48.10) - 185.18)}{(5 - 1)(4 - 1)}$$

$$Y_{43} = \frac{((147.5 + 192.4) - 185.18)}{(4)(3)}$$

$$Y_{43} = \frac{(339.9 - 185.18)}{12}$$

$$Y_{43} = \frac{(154.72)}{12} = 12.9$$

Therefore the missing data for $Y_{43} = 12.89$, which is change from 12.8 to 12.9 appropriately equal.

We proceed to compute for the Y_{11} by substituting the value of $Y_{43} = 12.9$

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	Y_{11}	5.0	11.2	15.7	31.9
Palm nut shell [B]	1.6	4.8	9.2	14.58	30.18
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	12.9	16.9	42.4
Treatment Totals	14.3	36.4	61	84.68	196.38

$$Y_{11} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

$$Y_{11} = \frac{(5(31.9) + 4(14.3) - 196.38)}{(5 - 1)(4 - 1)}$$

$$Y_{11} = \frac{((159.5 + 57.2) - 196.38)}{(4)(3)}$$

$$Y_{11} = \frac{(216.7 - 196.38)}{12}$$

$$Y_{11} = \frac{20.32}{12} = 1.7$$

Therefore $Y_{11} = 1.7$, This value equals the previous hence suggest the right value for $Y_{11} = 1.7$, we then check for Y_{24} to see whether the estimated missing value remains the same as the previous value computed by substituting the value $Y_{11} = 1.7$.

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	1.7	5.0	11.2	15.7	33.6
Palm nut shell [B]	1.6	4.8	9.2	Y_{24}	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	12.9	16.9	42.4
Treatment Totals	16	36.4	61	70.1	183.5

$$Y_{24} = \frac{(rB + tT - G)}{(r - 1)(t - 1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 183.5)}{(5 - 1)(4 - 1)}$$

$$Y_{24} = \frac{(5(15.6) + 4(70.1) - 183.5)}{(4)(3)}$$

$$Y_{24} = \frac{((78 + 280.4) - 183.5)}{12}$$

$$Y_{24} = \frac{((358.4 - 183.5))}{12}$$

$$Y_{24} = \frac{(174.9)}{12} = 14.58$$

Since the value of the missing data $Y_{24} = 14.58$, same as the previous computed value for Y_{24} . It thus confirms that all the missing data values have been accurately estimated hence the table can now be completed and used for the Analysis of Variance (ANOVA).

Thus the completed table is shown below:

Treatments (weights of the waste types) to be Pyrolysed					
Blocks	10 kg [A]	20kg [B]	30kg [C]	40kg [D]	Block Totals
Coconut waste fruits (shell and fibre) [A]	1.7	5.0	11.2	15.7	33.6
Palm nut shell [B]	1.6	4.8	9.2	14.58	15.6
Waste plastic bottles [C]	4.4	9.2	14.4	19.3	47.3
Waste plastic satchets [D]	4.2	8.9	13.3	18.2	44.6
Waste plastic bags [E]	4.1	8.5	12.9	16.9	42.4
Treatment Totals	16	36.4	61	70.1	183.5

Bibliography

- [1] Brownlee, K. A. (1960). *Statistical theory and methodology in science and engineering*. New York: Wiley.
- [2] Campbell, D., Stanley J. (1963). *Experimental and quasi-experimental designs for research and teaching*. In Gage (Ed.), *Handbook on research on teaching*. Chicago: Rand McNally & Co.
- [3] Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.
- [4] Cochran, W. G., Cox, G. M. (1957). *Experimental Designs*. John Willey & Sons, N.Y..
- [5] Le Clerg, E. L., Leonard, W. H., Clark A. G. (1966). *Field Plot Technique*. Burgess Pub. Co., Minn., USA.