International Journal of Scientific Engineering and Science
Volume 5, Issue 5, pp. 51-57, 2021. ISSN (Online): 2456-7361

On the Effect of Blocking in Randomized Block Design: An application to Competition of Soybeans Varieties against Weeds

ONU Obineke Henry[1], NELSON Maxwell[2], AMAKURO Okuata Avula[3]
[1, 2]Department of Mathematics and Statistics, Ignatius Ajuru University of Education, Rumuolumeni, Port Harcourt, Rivers State, Nigeria
[3]Department of Mathematics, Isaac Jasper Boro College of Education, Sagbama, Bayelsa State, Nigeria
Corresponding email: onuobinekehenry @ gmail.com

Abstract— *The research considered the effects of blocking in ramdomized block design with application in the competition of the varieties of soybeans against weeds in Northern Nigeria. It focused on the Randomized Complete Block Designs and compared it with these Designs when block effects are not considered. Data of the competition between Soybeans varieties against weeds were considered and rearranged in four different tables seen as table 1, table 2, table 3, and table 4 each with sample sizes of 24, 48, 72 and 96 respectively as seen in the example in this paper. Each of these tables were analyzed with or without blockings. The study considered ANOVA, Mean Square Error, F-test, Coefficient of determination, Akaike Information criterion (AIC), Schwarz Bayesian criterion (SBC) and Relative Efficiency of a pair of design. The effect of blocking in a design was seen to be of great importance, since in most cases, the analysis with blocks showed to be superior to those without blocking. The effects of sample size were also observed, as it was seen that changes in sample sizes have no effect on the calculated F-statistic and the larger the sample size the larger the values of AIC and SBC.*

Keywords— *Randomized complete block; Blocking; No Blocking; ANOVA; R squared; Relative Efficiency; Akaike Information Criterion; Schwarz Bayesian criterion.*

I. Introduction

Randomized block design is defined according to Stat Trek, Statistics Dictionary as a design in which the experimenter splits subjects into subgroups that is called blocks, in such a way that the variability within blocks is less than that between blocks. In randomized block design, a factor is always of kin interest while other factors are seen as nuisance factors.

According to concise encyclopedia of Statistics, Randomized block design is a design of experiment in which the units of the experiment are in groups known as blocks. The treatments are assigned at random to the experimental units in each block. It is also stated in this Statistical encyclopaedia that a completely randomized block design is obtained when all treatments allocated are seen at least once in each block. If not, we are said to have an incomplete randomized block design. A design of this nature is used to reduce the effects of systematic error. If the experimenter centres its attention wholly on the differences that exists between treatments, then, the effects due to variations between the different blocks should be eradicated.

According to Gary (2010), Randomized Complete Block Design is the first design that uses some kind of restricted randomization. It is a design in which every treatment is used in every block. It is also known as two-way ANOVA without interaction.

Gary (2010) stated that, incomplete block design is a design in which not every treatment is used in every block.

In the words of Prof William M.K. Trochim, in research method knowledge base, Randomized Block Design is defined as a research design that is comparable to Stratified sampling at random. It is known that in stratified sampling, randomized block designs are constructed to reduce variance effect or the nuisance effect in the data. This is done such that the experimenter will divide the sample into comparatively homogeneous blocks or subgroups which is known as "strata" in stratified sampling. The experimental design one may want to carryout is done within each homogeneous subgroup or block. The main idea is that the variability within each block is less than the variability of the entire sample. As a result, each estimate of the treatment effect within a block is more efficient than the estimates across the entire sample and when we pool these more efficient estimates across blocks, we should get an overall more efficient estimate more than we would get, without blocking.

Designed experiments are conducted to demonstrate a cause-and-effect relation between one or more explanatory factors (or predictors) and a response variable.

As expressed in Breadcrumb Stat 505 Applied Multivariate Statistical Analysis that within a randomized block designs, there are two factors known as Blocks and Treatments.

It was stated in Breadcrumb that a randomized complete block design having m treatments and n blocks is built in two stages which include, the experimental units which are the units to which the treatments are to be applied are divided into n blocks, each comprising of m units and the treatments are randomly assigned to the experimental units in such a way that each treatment appears once in each block as also seen in concise encyclopedia of Statistics. Randomized block designs are often applied in agricultural settings as seen in this research in the illustration given in this work.

Blocking in the statistical theory of design of experiment (Wikipedia, the free encyclopedia) is the arrangement of experimental units in blocks or groups that show similarity with one another. Blocking decreases unexplained variability. It has principle that lies in the fact that variability which cannot be overcome is aliased or confounded with higher order interaction as to jettison its influence on the final product. This higher order interaction is always less importance; therefore, it is desirable to alias this variability with the higher interaction.

In this paper, we considered a Randomized Block Design with blocking effects and when the blocking effect is not considered. Ogunsanya and Mutiu (2016) studied the application of Randomized Completely Block Design to yield of Maize. Urip et al (2019) worked on Four Factors experiments for fixed models in CRD, the paper was prepared to find ANOVA table for four factors experiments for fixed models in RCD. They determined source of variation, degree of freedom, sum of squares, mean squares, expected values of mean square and F-Test statistic. Iruegbukpe and Mbegbu (2011) looked at the choice of a completely randomized designs with serious emphasis on D-Optimal criterion, they studied unconstrained CRD, Zero-sum constraint and the Baseline constraints using D-Optimal criterion to choose the one that will yield optimal result. These studies did not consider the analysis of design with blocking effect on that without blocking effect and the effect of sample sizes on the models used. We applied mean square error, F-test, R Squared, AIC, SBC and Relative Efficiency. The data is arranged in four (4) different tables, with table 1 containing four (4) blocks with 24 sample size, table 2 containing 8 blocks with 48 sample size, table 3 containing 12 blocks with 72 sample size and table 4 containing 16 blocks with 96 sample size, this indicates that table 1 differs from table 2 by four samples in that order. These four tables are each analyzed with or without blocking. Note that all these arrangements were obtained from table 4. Each of these tables are analyzed with or without blocking and the ANOVA tables are shown respectively as table 1a, 1b, table 2a, 2b, table 3a, 3b and table 4a, 4b where 1a, 2a, 3a and 4a represent the ANOVA table for blocking analysis for table 1-4 and 1b, 2b, 3b and 4b represent the ANOVA table for analysis without blocking for table 1-4.

The work is aimed at investigating the effect of blocking in Randomized Block Design by comparing designs with blocking with designs without blocking. This will be achieved by using the Mean square error, R squared, Akaike Information, Schwarz Bayesian criteria to see how the data with or without blocking and with different sample sizes will fare in relation to the statistical models used. We also applied Relative Efficiency between a pair of design, where these designs in the tables 1-4 for with or without blocking and for changes in sample sizes were compared using these criteria. The data used was the data of the competition of soybean varieties against weeds obtained from Bussan (1995).

II. Review of Related Literatures

Urip et al. (2019) worked on Four Factors experiments for fixed models in CRD, the result of this research found an ANOVA Table for Completely Randomized Factorial (CRF)-2222 Design for Fixed Model independently which are made up of 16 of SV, 16 of df, 16 of SS, 16 of MS, 16 of EMS, 15 of F_0, and 15 of table F.

Ogunsanya and Mutiu (2016) studied the application of Randomized Completely Block Design to yield of Maize, it was observed that there is significant difference in the effect of the fertilizer's proportions and the varieties of maize on the yield of maize in Ogun State in Nigeria. Multiple comparisons test that was conducted for fertilizer proportion shows that the significant difference is to be between 50kg and 200kg fertilizers and between 100kg and 200kg fertilizers.

Iruegbukpe and Mbegbu (2011) looked at the choice of a completely randomized designs with serious emphasis on D-Optimal criterion, it was noted that completely randomized design are designs in which treatments are probabilistically assigned to the experimental units. In their study, they considered designs like the zero-sum constraint CRD, the unconstrained CRD and the baseline constraint CRD. It was revealed that the baseline constraint CRD is the choice of designs that is more suitable for an experiment with five treatments and six replicates.

According to Barak and David (2010); the threats posed on statistical tests alongside the precision of treatments estimates as occasioned by the imbalances witnessed in simple random allocation can be reduced by Randomized Block Designs (RBD). This was done by sacrificing complete randomization in treatment allocation. All of these are in terms of covariant effects or in the variance between group size.

Ramani et al. (2005), considered the statistical analysis of modified complete randomized design. The study compared randomized complete block design (RCBD) with the split-plot design with two goals, which include, to show the consequence of constructing F-test on a mean square error used in testing the significance of the effects of treatments under restricted randomization and to design an alternative to the formal that is based on split-plot analysis of variance, in order to obtain designs with better power under the same condition. The restricted randomization according to this study simply means submitting the treatments in whole-plots of four runs, where these four runs are treated as a single unit of experiment. It was observed that the split-plot analysis resulted in a more powerful alternative to the data collected under restricted randomization, they noted that constraints of experiments can prevent complete randomization within each block, also, overlooking the effect of restricted randomization on inferences from RCBD analysis can lead to several spurious interaction effects as well as potentially type 1 or type 2 errors.

III. Materials and Methods

The model used in this research is a three parameter model given as

$$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \quad (1a)$$

Where μ is the overall or grand mean of the experiment,
α_i is the i^{th} treatment effect representing the St. Paul MN (STP) and the Rosemount, MN (R)

International Journal of Scientific Engineering and Science
Volume 5, Issue 5, pp. 51-57, 2021. ISSN (Online): 2456-7361

β_j is the j^{th} block effect representing the varieties of the Soybeans
ϵ_{ij} is the stochastic or random error or noise and
Y_{ij} is the i^{th} and j^{th} response variables which represents the weed biomass in this study.
Model (1) is without interaction term and is to be used for all the analysis in this study.
The model having no Block effect and no interaction which will also be used in this study is given as

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad (1b)$$

the block effect has been added to the error.
Here the error becomes

$$\epsilon_{ij} = \epsilon_{ij} + \beta_j \quad (1c)$$

The research considers a Randomized Complete Block Design symbolically given the table below.

	Herb.2weeks		Herb.4weeks		No herb.		
Variety	R	STP	R	STP	R	STP	$Y_{i.}$
Parker	A	B	C	D	E	F	sum parker
lambert	G	H	I	J	K	L	sum lambert
M89-792	M	N	O	P	Q	R	sum m89-79
Sturdy	S	T	U	V	W	X	sum sturdy
$Y_{.j}$	Sum1	sum2	Sum3	sum4	Sum5	Sum6	SUM total

Where the alphabets in capital letter represents the treatments, the sum of each rows are represented as sum parker, sum lambert, sum m89-792, sum sturdy respectively which represent varieties of Soybeans. While the sum of each column are represented as Sum1, Sum2, Sum3, Sum4, Sum5 and Sum6 respectively and these represent the replicates of the treatments on the weeds Biomass.
Note that;
sum parker + sum lambert + sum m89-792+sum sturdy=Sum1+Sum2+Sum3+Sum4+Sum5+Sum6=SUM total.
We apply the correction factor as seen

$$C = \frac{Y_{..}^2}{N} = \frac{(\text{SUM total})^2}{m \times n} \quad (2)$$

where m represents number of treatment and n represents number of blocks

We obtain sum of square total given as seen in Gary W. O., (2010) and Kutner et al. (2005) as

$$SS_{TOTAL} = \sum_{ij} Y_{ij}^2 - C \quad (3)$$

The sum of square Treatment is obtained as

$$SS_{Treat} = \sum_i \frac{Y_{i.}^2}{m} - C \quad (4)$$

The sum of square block is given as

$$SS_{Block} = \sum_j \frac{Y_{.j}^2}{n} - C \quad (5)$$

Hence, we obtain the error sum of square given as

$$SS_{Error} = SS_{Total} - SS_{Treat} - SS_{Block} \quad (6)$$

The ANOVA table is as shown bellow

ANOVA TABLE

SV	DF	SS	MS	F_{Cal}
Treatment	$i-1$	SS_{Treat}	$MS_{Treat}/_{i-1}$	MS_{Treat}/MS_{Error}
Block	$j-1$	SS_{Block}	$MS_{Block}/_{j-1}$	MS_{Block}/MS_{Error}
Error	$(i-1)(j-1)$	SS_{Error}	$MS_{Error}/_{(i-1)(j-1)}$	
Total	$ij-1$	SS_{Total}	$SS_{Total}/_{ij-1}$	

The analysis without blocking is carried out as seen
Consider the table

	Herb.2weeks		Herb.4weeks		No herb.		
Variety	R	STP	R	STP	R	STP	$Y_{i.}$
Parker	A	B	C	D	E	F	sum parker
lambert	G	H	I	J	K	L	sum lambert
M89-792	M	N	O	P	Q	R	sum m89-79
Sturdy	S	T	U	V	W	X	sum sturdy
							sum total

You can see that block effect which is the application of herbicides is not represented in the table, here we consider only the treatment effect and we obtain the following
Correction factor $C = \frac{Y_{..}^2}{N} = \frac{(\text{sum total})^2}{N}$, where N is the total entry in the table

We apply equation (2) and (3) to obtain sum of square Total and sum of square Treatment respectively. Note that there is no difference in what was obtained in the blocking design with what is to be obtained for this design without blocking in respect to sum of square Total and sum of square Treatment.

The sum of square Error of the analysis with blocking is added to the sum of square Block for the analysis with blocking to obtain the sum of square Error for the analysis without Blocking, also, the degree of freedom for analysis with blocking is added to the degree of freedom of the block for the analysis with blocking to obtain the Error degree of freedom for analysis without blocking. That is to say

$$SS_{Error}\text{without block} = SS_{Error}\text{with block} + SS_{Block} \quad (7)$$

Also,

$$\text{DF without Blocking} = \text{Error DF} + \text{DF withblocking} \quad (8)$$

Where DF represent degree of freedom
The hypothesis is stated as
H_0: Soyabeans varieties compete favorably well against weeds
H_A: Soya beans varieties do not compete favorably well against weeds
If $F_{Cal} > F_{Tab}$ we reject H_0.

The above hypothesis is considered with significant level of 0.05.

In order to test the performance of the data as arranged in table 1, 2, 3, 4 with the model in (1) we employ the following criteria
coefficient of determination which is given as

$$R^2 = \frac{SSR}{SSTotal} \quad (9)$$

$$= 1 - \frac{SSE}{SSTotal} \quad (10)$$

Note that coefficient of determination lies between 0 and 1, the more the value is close to one, the better the model fit on the data, while as the value becomes closer to zero, the poorer the model fit. It has been made known by Kutner et al (2005) that the shortcoming of R squared include the fact that it does not take into account the number of model parameters in the model and that the maximum number of R squared does not reduce with increasing model parameters. As a result, one use R squared adjusted to address such challenge which is not considered in this study. We present Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC) which are also presented as alternatives for R squared and R squared

Volume 5, Issue 5, pp. 51-57, 2021. ISSN (Online): 2456-7361

Adjusted and that penalizes adding more predictors to the model. These criteria seek for the model with small value of AIC and SBC. They are given as seen in Kutner et al (2005) as shown

AIC= $n \ln SS_{Error} - n \ln n + 2p$ (11)

SBC= $n \ln SS_{Error} - n \ln n + [\ln n]p$ (12)

Where p is the number of model parameters and n is the sample size.

Mean Square Error approach, which is given as

MSE= $\frac{SS_{Error}}{(i-1)(j-1)}$ (13)

Was also applied to further see the adequacy of the model on the respective data as seen in this study. The criterion believes that the smaller the MSE the better the model fits the data.

The adjusted R squared is given as

$R^2_{Adjusted=} = (\frac{n-1}{n-p})(\frac{SS_{Error}}{SS_{Total}})$ (14)

$= 1 - \frac{MSE}{\frac{SSTotal}{n-1}}$ (15)

To evaluate the effectiveness of one design in relation to another design we apply a measure called Relative Efficiency given as

Relative Efficiency(A,B)= $\frac{(DF_A+1)}{(DF_B+1)}\frac{MSE_B}{MSE_A}$ (16)

Design A is more efficient than Design B if R.E of A is greater than 1, if R.E of A is less than 1, then Design B is more efficient. Note that R.E is Relative Efficiency.

In this research, the design with blocking is Design A, while that without blocking is Design B

IV. RESULTS AND DISCUSSION

4.1 Results

Example

Consider the experiment conducted to determine how different soybeans varieties compete against weeds. There were sixteen varieties of soybeans and three weed treatments: No herbicide, apply herbicide 2 weeks after planting the soybeans and apply herbicide 4 weeks after planting soybeans. The measure response is weed biomass in kg/ha. There were two replicates of the experiment –one in St. Paul MN, and one in Rosemount, MN –for a total of 96 observations. The data is seen in table 4.

Rearranging the data of table 4 as seen in tables 1, 2, 3, this involves reducing the sample sizes.

Each of the tables are analyzed with or without blocking

Table 1 analysis with blocking for 24 sample size

	Herb.2weeks		Herb.4weeks		No herb.		
Variety	R	STP	R	STP	R	STP	$Y_{i\cdot}$
Parker	750	1440	1630	890	3590	740	9040
lambert	870	550	3430	2520	6850	1620	15840
M89-792	1090	130	2930	570	3710	3600	12030
Sturdy	1110	400	1310	2060	2680	1510	9070
$Y_{\cdot j}$	3820	2520	9300	6040	16830	7470	45980

Analysis with blocking in table 1

Correction factor C= $\frac{Y_{..}^2}{N}$, where $N = 24$

$= \frac{45980^2}{24}$

$= 88090016.7$

$SS_{Total} = \sum_{ij} Y_{ij}^2 - C$

$= 141895600 - 88090016.7$

$= 53805583.3$

$SS_{Treat} = \sum_i \frac{Y_{i\cdot}^2}{m} - C$, where $m = 6$

$= 93268833.3 - 88090016.7$

$= 5178816.6$

$SS_{Block} = \sum_j \frac{Y_{\cdot j}^2}{n} - C$, where $n = 4$

$= 120741050 - 88090016.7$

$= 32651033.3$

$SS_{Error} = 53805583.3 - 5178816.6 - 32651033.3$

$= 15975733.4$

ANOVA TABLE 1a with blocking

SV	DF	SS	MS	F_{Cal}	F_{Tab}
Treatment	5	5178816.6	1035763.3	0.973	2.90
Block	3	32651033.3	10883677.8	10.23	3.29
Error	15	15975733.4	1065048.9		
Total	23	53805583.3	2339373.2		

ANOVA TABLE 1b without blocking

SV	DF	SS	MS	F_{Cal}	F_{Tab}
Treatment	5	5178816.6	1035763.3	0.383	2.77
Error	18	48626766.7	2701487.0		
Total	23	53805583.3	2339373.2		

Table 2 analysis with blocking for 48 sample size

	Herb.2weeks		Herb.4weeks		No herb.		
Variety	R	STP	R	STP	R	STP	$Y_{i\cdot}$
Parker	750	1440	1630	890	3590	740	9040
lambert	870	550	3430	2520	6850	1620	15840
M89-792	1090	130	2930	570	3710	3600	12030
Sturdy	1110	400	1310	2060	2680	1510	9070
Ozzie	1150	370	1730	2420	4870	1700	12240
M89-1743	1210	430	6070	2790	4480	5070	20050
M89-794	1330	190	1700	1370	3740	610	8940
M90-1682	1630	200	2000	880	3330	3030	11070
$Y_{\cdot j}$	9140	3710	20800	13500	33250	17880	98280

Analysis with blocking in table 2

Correction factor C= $\frac{Y_{..}^2}{N}$, where $N = 48$

$= \frac{98280^2}{48}$

$= 201228300$

$SS_{Total} = \sum_{ij} Y_{ij}^2 - C$

$= 319537200 - 201228300$

$= 118308900$

$SS_{Treat} = \sum_i \frac{Y_{i\cdot}^2}{m} - C$, where $m = 6$

$= 218983600 - 201228300$

$= 17755300$

$SS_{Block} = \sum_j \frac{Y_{\cdot j}^2}{n} - C$, where $n = 8$

$= 267181325 - 201228300 = 65953025$

$SS_{Error} = 118308900 - 17755300 - 65953025$

$= 34600575$

ANOVA TABLE 2a with blocking

SV	DF	SS	MS	F_{Cal}	F_{Tab}
Treatment	5	17755300	3551060	36.02	2.45
Block	7	65953025	9421860.7	9.53	2.25
Error	35	34600575	988587.9		
Total	47	118308900	2517210.6		

ANOVA TABLE 2b without blocking

SV	DF	SS	MS	F_{Cal}	F_{Tab}
Treatment	5	17755300	3551060		
				1.483	2.45
Error	42	100553600	2394133.3		
Total	47	118308900			

Table 3 analysis with blocking for 72 sample size

	Herb.2weeks		Herb.4weeks		No herb.		
Variety	R	STP	R	STP	R	STP	$Y_{i\cdot}$
Parker	750	1440	1630	890	3590	740	9040
lambert	870	550	3430	2520	6850	1620	15840
M89-792	1090	130	2930	570	3710	3600	12030
Sturdy	1110	400	1310	2060	2680	1510	9070
Ozzie	1150	370	1730	2420	4870	1700	12240
M89-1743	1210	430	6070	2790	4480	5070	20050
M89-794	1330	190	1700	1370	3740	610	8940
M90-1682	1630	200	2000	880	3330	3030	11070
M89-1846	1660	230	2290	2210	3180	2640	12210
Archer	2210	1110	3070	2120	6980	2210	17700
M89-642	2290	220	530	390	3750	2590	9770
M90-317	2320	330	1760	680	2320	2700	10110
$Y_{\cdot j}$	17620	5600	28450	18900	49480	28020	148070

Analysis with blocking in table 3

Correction factor C= $\frac{Y_{..}^2}{N}$, where $N = 72$

$$= \frac{148070^2}{72}$$
$$= 304510008.1$$
$$SS_{Total} = \sum_{ij} Y_{ij}^2 - C$$
$$= 471415700 - 304510008.1$$
$$= 440964691.9$$
$$SS_{Treat} = \sum_i \frac{Y_{i\cdot}^2}{m} - C, \text{ where } m = 6$$
$$= 328990116.7 - 304510008.1$$
$$= 24480108.7$$
$$SS_{Block} = \sum_j \frac{Y_{\cdot j}^2}{n} - C, \text{ where } n = 12$$
$$= 395152308.3 - 304510008.1$$
$$= 364701300.2$$
$$SS_{Error} = 440964691.9 - 24480108.7 - 364701300.2$$
$$= 51783283$$

ANOVA TABLE 3a with blocking

SV	DF	SS	MS	F_{Cal}	F_{Tab}
Treatment	5	24480108.7	4896021.7	5.200	2.37
Block	11	364701300.2	33154663.7	35.21	2.17
Error	55	51783283	941514.2		
Total	71	440964691.9	6210770.3		

ANOVA TABLE 3b without blocking

SV	DF	SS	MS	F_{Cal}	F_{Tab}
Treatment	5	24480108.7	4896021.7	1.483	2.45
Error	66	416484583.2	6310372.5		
Total	71	440964691.9			

Table 4 Analysis with blocking for 96 sample size

	Herb.2weeks		Herb.4weeks		No herb.		
Variety	R	STP	R	STP	R	STP	$Y_{i\cdot}$
Parker	750	1440	1630	890	3590	740	9040
lambert	870	550	3430	2520	6850	1620	15840
M89-792	1090	130	2930	570	3710	3600	12030
Sturdy	1110	400	1310	2060	2680	1510	9070
Ozzie	1150	370	1730	2420	4870	1700	12240
M89-1743	1210	430	6070	2790	4480	5070	20050
M89-794	1330	190	1700	1370	3740	610	8940
M90-1682	1630	200	2000	880	3330	3030	11070
M89-1846	1660	230	2290	2210	3180	2640	12210
Archer	2210	1110	3070	2120	6980	2210	17700
M89-642	2290	220	530	390	3750	2590	9770
M90-317	2320	330	1760	680	2320	2700	10110
M90-610	2480	350	1360	1680	5240	1510	12620
M88-250	2480	350	1810	1020	6230	2420	14310
M89-1006	2430	280	2420	2350	5990	1590	15060
M89-1926	3120	260	1360	1840	5980	1590	14150
$Y_{\cdot j}$	28130	6840	36400	25790	72920	35130	205210

Analysis with blocking in table 4

Correction factor C= $\frac{Y_{..}^2}{N}$, where $N = 96$

$$= \frac{205210^2}{96}$$
$$= 438657751$$
$$SS_{Total} = \sum_{ij} Y_{ij}^2 - C$$
$$= 4919153100 - 438657751$$
$$= 4480495349$$
$$SS_{Treat} = \sum_i \frac{Y_{i\cdot}^2}{m} - C, \text{ where } m = 6$$
$$= 464257883.3 - 438657751$$
$$= 25600132.3$$
$$SS_{Block} = \sum_j \frac{Y_{\cdot j}^2}{n} - C, \text{ where } n = 16$$
$$= 586225618.8 - 438657751$$
$$= 147567867.8$$
$$SS_{Error} = 4480495349 - 25600132.3 - 147567867.8$$
$$= 4307327349$$

ANOVA TABLE 4a with blocking

SV	DF	SS	MS	F_{Cal}	F_{Tab}
Treatment	5	25600132.3	5120026.4	0.089	2.21
Block	15	147567867.8	9837857.9	0.171	2.01
Error	75	4307327349	57431031.3		
Total	95	4480495349	47163108.9		

ANOVA TABLE 4b without blocking

SV	DF	SS	MS	F_{Cal}	F_{Tab}
Treatment	5	25600132.3	5120026.4	0.103	2.29
Error	90	4454895217	49498835.7		
Total	95	4480495349			

We applied R squared, Aikaike Information Criterion and Schwarz Bayesian Criterion to see the validity of these data in tables 1,2,3 and 4 with the model under study.

Table 5 R squared, AIC and SBC

Sample size	Blocks	R square	AIC	SBC
24	4	0.7031	327.81	331.34
48	8	0.7075	653.43	656.96
72	12	0.8826	976.99	983.82
96	16	0.0387	1697.44	1705.03

Application of Relative efficiency to pairs of designs

Comparison of the Blocking and the No Blocking for 24 sample size

For 24 sample size with or without Blocking, we have

Relative Efficiency (Blocking, No Blocking)= $\frac{(3+1)(18+3)}{(18+1)(3+3)} \times \frac{2701487.0}{1065048.9}$

$$= 1.87 > 1$$

Comparison of the Blocking and the No Blocking for 48 sample size

For 48 sample size with or without Blocking, we have

International Journal of Scientific Engineering and Science
Volume 5, Issue 5, pp. 51-57, 2021. ISSN (Online): 2456-7361

Relative Efficiency (Blocking, No Blocking) $= \frac{(7+1)(42+3)}{(42+1)(7+3)} \times \frac{2394133.3}{988587.9}$

$$= 2.02 > 1$$

Comparison of the Blocking and the No Blocking for 72 sample size

For 72 sample size with or without Blocking, we have

Relative Efficiency (Blocking, No Blocking) $= \frac{(11+1)(66+3)}{(66+1)(11+3)} \times \frac{6310372.5}{941514.2}$

$$= 5.92 > 1$$

Comparison of the Blocking and the No Blocking for 96 sample size

For 96 sample size with or without Blocking, we have

Relative Efficiency (Blocking, No Blocking) $= \frac{(15+1)(90+3)}{(90+1)(15+3)} \times \frac{49498835.7}{57431031.3}$

$$= 0.783 < 1$$

Comparison between sample size of 24 and 48 in terms of Relative Efficiency without Blocking

Relative Efficiency (24, 48) $= \frac{(3+1)(7+3)}{(7+1)(3+3)} \times \frac{49498835.7}{57431031.3}$

$$= 0.773 < 1$$

Comparison between sample size of 48 and 72 in terms of Relative Efficiency without Blocking

Relative Efficiency (48, 72) $= \frac{(7+1)(11+3)}{(11+1)(7+3)} \times \frac{941514.2}{988587.9}$

$$= 0.888 < 1$$

Comparison between sample size of 72 and 96 in terms of Relative Efficiency without Blocking

Relative Efficiency (48, 72) $= \frac{(11+1)(15+3)}{(15+1)(11+3)} \times \frac{57431031.3}{941514.2}$

$$= 58.8 > 1$$

4.2. Discussion

4.2.1 Table 1 with or without blocking

Considering table 1 analyzed with or without blocking, it was observed that using blocks gave more powerful test in determining how well the soybeans compete with weeds than without blocks. With blocking, we accepted the null hypothesis, that Soybeans compete favorably well with weeds and that there is no significant difference among the treatments and the blocks in the competition. In the analysis without blocking, we also accepted the null hypothesis. It also reveals that there is no significant difference among the treatments in the competition.

4.2.2. Table 2 with or without blocking

Considering table2 analyzed with or without blocking, it was observed that inclusion of blocks is more powerful in determining how soybeans varieties compete with weeds than that without blocking, this is evident in the increased value of F in the analysis with blocking. For the analysis with blocking, we reject null hypothesis and accept that soybeans do not compete favorably well with weeds and that there is significant difference between treatments and blocks in the competition. For the analysis without blocking, we accepted the null hypothesis that soybeans compete favorably well with weeds and that there is significant difference among the treatments effects in the competition.

4.2.3. Table 3 with or without blocking

Considering table3 analyzed with or without blocking, it was observed that inclusion of blocking proved more powerful in determining how soybeans varieties compete with weeds than without blocking as revealed by the higher value of F in the analysis with blocking. For analysis with blocking, we reject null hypothesis and accept that soybeans varieties do not compete favorably well with weeds and that there is significant difference among the treatments and the blocks, for analysis without blocking, we accept null hypothesis that soybeans varieties compete well with weeds and that there is no significant difference among the treatments effects.

4.2.4. Table 4 with or without blocking

Considering table4 analyzed with or without blocking, it was observed that the analysis without blocking proved to be a more powerful test than that with blocking as evident in the value of F for the without blocking higher than the value of F with blocking. This proves different from what was observed in the smaller sample sizes. It suggests that at a certain higher sample size, analysis without blocking be considered. For analysis with blocking, we accepted the null hypothesis that the soybeans varieties do compete favorably well with weeds and that, there is no significant difference among the treatments and the blocks and this is also true for that without blocking.

Generally, it was observed that in table 1 and 4, we accepted the null hypothesis that soybeans varieties compete favorably with weeds for analysis with blocking, while for analysis without blocking, we rejected the null hypothesis that soybeans varieties do no compete favorably with weeds.

Also, for table 3 and 4 it was revealed that we rejected the null hypothesis that is to say we accepted that soybeans do no compete favorably with weeds for analysis with blocking, while that without blocking, we accepted the null hypothesis that soybeans varieties compete favorably with weeds.

In table5, it was observed that the AIC and SBC increases with increasing sample size and that the data in table3 with 72 sample size proves to be the best data for the model followed by table2 with 48 sample size and table1 with 24 sample size and the table4 proved to be the worst with this data. This is evident in the value of R squared. The small values of AIC and SBC shows that the data or the model are better. But this is not in agreement with the R squared because the AIC and SBC shows that table1 proves to be the best with the model followed by the table2 and then table3. R squared agreed with AIC and SBC table4. If the suggestions of AIC and SBC are true, it then means that the analysis with small sample and three parameters model without interaction is favored by these criteria.

As the sample sizes increases, the F value increases which agrees with AIC and SBC that the 24 sample size is the best for this model, followed by 48 sample size and then 72, where 96 proves to be the worst for this model.

In the application of Relative Efficiency between pairs of design, it was observed that for sample size of 24, the design with blocking is more efficient than the design without blocking. For the sample size of 48, it was revealed that the design with blocking appears more efficient than the design

Volume 5, Issue 5, pp. 51-57, 2021. ISSN (Online): 2456-7361

without blocking. Whereas for sample size of 72 the design with blocking is more efficient than that without blocking. For 96 sample size, the opposite was observed, where the design without blocking becomes more efficient than that with blocking.

In the application of relative efficiency between two sample sizes, it was observed that between 24 and 48 sample size, the sample size of 48 proves to be more efficient than that of 24 sample size, also between 48 and 72 sample size, the sample size of 72 is more efficient than that of 48 and finally, between 72 and 96 sample size, the opposite was observed, where the sample of 72 was found more efficient than that of 96. Following from the findings from R squared, that the sample size of 72 is the best followed by the 48 and then the 24, which confirms that of the Relative efficiency, which says that 72 sample size is the best design. Also the both criteria agreed that the sample size of 96 proves to be the worst for this model. Likewise, the AIC and SBC, which also said that 96 sample size is the worst for this model, but AIC and SBC disagreed with the other criteria in the since that they observed that the sample size of 24 is the best followed by the sample size of 48 and then that of 72. It was observed that AIC and SBC values increases as the sample size increases.

V. Summary Conclusion and Recommendation

Summary

This research looked at Randomized Block Design but particularly, Randomized Complete Block Design. The study compared designs with blocking and designs without blocking. It also compared the designs in relation with their sample sizes. It is of the view to know the effect of blocking in a design and to see how sample size affect designs in Randomized Complete Block Designs. The paper employed ANOVA model, Coefficient of Determination, Akaike Information Criterion, Schwarz Bayesian Criterion, Mean Square Error and Relative Efficiency of pair of Design in making the decision as seen in this paper.

Conclusion

It was found reasonable to state that the AIC and SBC values increases with increasing sample size in a design of these settings and that the value may not necessary represent the true nature of things base on the discrepancies observed between the values of R squared and those of AIC and SBC. As a result of this, in other to get the true nature of things, we employed other tests like the relative efficiency which agreed with the AIC and SBC on the ground that the worst design for this model proved to be the sample size of 96, but they disagreed all through other than this. On the application of analysis of variance revealed that from table 1 through 4, the blocking design proved to be the more powerful test than those without blockings. In terms of sample size, it was revealed that as the sample size increases the F value also increases which reveals that the smaller the sample size the more powerful the test, which agrees with the AIC and SBC.

Recommendation

It is recommended that in carrying out analysis in RCD, blocking design should be used, especially when the sample size is not too large, say, from sample size of 80 and below. But for a very large sample size, the analysis without blocking should be considered, say 81 and above.

References

[1] Barak A. and David P. F. (2010). *Randomized Block Designs, Handbook of quantitative Criminology, page 437-454, doi: 10.1007/978-0387-77650-721.*

[2] Breadcrumb Stat 505, *Applied Multivariate Statistical Analysis*

[3] Bussan A. J. (1995). *Selection for weed competitiveness among soybean genotypes, M.Sc. Thesis, University of Minnesta, St. Paul MN, alken Press. 108-116.*

[4] *Gary W. O. (2010). A first course in design and analysis of experiments.*

[5] Iruegbukpe D. O. and Mbegbu J. I. (2011). On the Choice of Completely Randomized Design with emphasis on the D-Optimality criterion, *African journals online (AJOL).*

[6] Kutner M. H. Nachtsheim C. J., Neter J. and Li W. (2005). Applied linear Statistical models, *McGraw-Hill Irwin, fifth edition.*

[7] Ogunsanya B.G. and Mutiu Sulaiman (2016). Application of Randomized Completely Block Design to the Yield of Maize, *International Journal of Research 3(12); 997-1002.*

[8] Ramani S., Pilla, David J., Kitska, Catherine L. (2005). Statistical Analysis of modified complete randomized designs: applications to chemo-orientation studies, *journal of experimental biology, page 1267-1276; doi:10.1242/jeb.01523.*

[9] Springer- Verlage (2008). Randomized Block Design in the Concise encyclopedia of statistics, *Springer, New York. NY. http//doi.org/10.1007/978-0-387-32833-1_344.*

[10] *Stat Trek, Statistics Dictionary.*

[11] Urip Tisngati, Nely Indra Meifiani, Dwi Cahyani Nur Apriyani and Martini (2019). Four factors experiments for fixed models in completely randomized design, *international conference on Advance and Scientific Innovation (ICASI), IOP Conf.Series; Journal of Physics:Conf.Series 1175.*

[12] Willian M. K. Trochim, (2010). *Research methods knowledge Base, hosted by conjointly.*

[13] Wikipedia, the free encyclopedia (2018). *Statistical theory of the design of experiment*