Statistical Performance Analysis of Complete and Incomplete Block Designs: a Comparison of RCBD, Lattice Design and Alpha-Lattice Designs under SARI Field Conditions

By

Ashenafi Abebe

A Thesis Submitted to the Department of Statistics, School of Graduate Studies, College of Natural Science, Jimma University
In Partial Fulfillment for the Requirements of Masters of Science (MSc) Degree in Biostatistics

October, 2011

Jimma, Ethiopia

Statistical Performance Analysis of Complete and Incomplete Block Designs: a Comparison of RCBD, Lattice Design and Alpha-Lattice Designs under SARI Filed Conditions

M.Sc. Thesis

Ashenafi Abebe Gaenamo

October, 2011

Jimma, Ethiopia

DEPARTMENT OF STATISTICS, SCHOOL OF GRADUATE STUDIES

JIMMA UNIVERSITY

As thesis advisors, we hereby certify that we had read and evaluated the thesis prepared by Ashenafi Abebe under our guidance, which was entitled statistical ***"Statistical Performance Analysis of Complete and Incomplete Block Designs: a Comparison of RCBD, Lattice Design and Alpha-Lattice Designs under SARI Field Conditions".*** We recommend that the thesis be submitted as it fulfills the requirements for the degree of Master of Science in Biostatistics.

Yehenew Getachew (Asst Prof, PhD Scholar) _____ _____

Major advisor Signature Date

Legesse Negash (PhD Scholar) _____ _____

Co- advisor Signature Date

As the members of the board of examiners of Msc thesis open defense examination of Ashenafi Abebe Gaenamo, we certify that we have read and evaluated the thesis and examined the candidate. We recommend that the thesis be accepted as it fulfills the requirements for the degree of Master of Science in Biostatistics.

Name of Chairman Signature Date

Name of Investigator Signature Date

Name of Internal Examiner Signature Date

Name of External Examiner Signature Date

Department Head Signature Date

DEDICATION

"This thesis is dedicated to my father **Abebe Gaenamo"**

STATEMENT OF THE AUTHOR

I declare that this thesis is a result of my genuine work and all sources of materials used for the thesis have been duly acknowledged. I have submitted thesis to Jimma University in partial fulfillment for the degree of Master of Science. The thesis can be deposited in the library of the university to be made available for people as reference.

Brief quotations from this thesis are followed without requiring special permission provided that an accurate acknowledgement of the course is made. Request for extended quotations for the reproduction of the thesis in part or in whole may be granted by the head of the department of statistics when in his or her judgment the proposed use of the material is for a scholarly interest. In all other instances, however, permission must be obtained from the author.

Name: **Ashenafi Abebe** Signature……

Place: Jimma University

Date of submission: ……..

Acknowledgement

First and foremost, I would like to acknowledge the mercy of God on me.

My heartfelt thanks goes to **Mr. Yehenew Getachew (Asst Prof, PhD Scholar)** my major advisor and instructor of the courses Survival Data analysis and Experimental Design and Analysis who did his best for the establishment of the program to run from scratch without him this journey would never have begun and my co-advisor **Mr**. **Legesse Negash (PhD Scholar)** for his friendly and considerable contribution and direction of this thesis and support of various kinds through some difficult times.

I am highly indebted to **Mr. Zenebe Fikrie Banjaw,** head of Department of Statistics, Jimma University for his kind assistance in many ways since the start of the program.

I would like to gratefully acknowledge **Mr. Solomon Admasu** and other staff members of South Agricultural Research Institute for helping me in getting data for my thesis work.

I am very much beholden to my best friends **Abdi, keke**, **Ephi**, **Ashenafi Y., Tesmek**, **Dr. Adane Desta**, **Ashu**, **Antush**, **Girmish** and **Serawit** for their ceaseless support and encouragement throughout my stay at study. God bless the rest of their life abundantly.

My warm thanks goes to my beloved family father **Abebe Gaenamo**, mother **Almaz Fisiha**, Sister **Adu** and Brother **Sadi** whose endless encouragement, moral and financial support were the sources of inspiration to me during my entire academic career as well as in my postgraduate study.

Last but not least I would like to thank my fiancé **Hirut Woldeyohannes (Heroye)** for her patience and intensive assistance throughout my stay at school.

Abstract

This study was conducted with the overall purpose of comparing the performance of commonly used incomplete block designs over that of the classical RCBD. Among the incomplete block designs, Lattice design and alpha lattice designs were employed. The comparison was statistically done mainly based on mean square errors and their corresponding CVs for each design. For this purpose, three datasets obtained from SARI were analyzed using CRD, RCBD, lattice and alpha lattice designs. The results of the soybean variety trial data containing 8 treatments having two factors with 3 replications at five different locations were used to assess the performance of RCBD over CRD. The result showed that 31, 3, 53, and 13% precision increased with RCBD over CRD for four sites namely, Hawassa, Areka, Gofa and Bonga, respectively. The CV for CRD is 25.9, 19.2, 7.3 and 12.9% for the four sites above, respectively. While that of RCBD is 22.6, 18.8, 5.9 and 12.3% respectively. This again confirms that RCBD is more efficient than CRD under those tested sites. The implication of the insignificant block effect is there is no need of block for this site. The results of the maize variety trial data containing 25 treatments with 4 replications shown that 0.4, 6.2, 15.0, 0.1 and 10.3% precision increased with Lattice design over classical RCBD for the five research sites namely, Hawassa, Areka, Bonga, Jinka and Arba Minch sub Center, respectively. The CV for lattice design was 26.4, 20.9, 15.7, 21.7 and 18.9% for the above five sites, respectively. While that of the RCBD was 28.22, 25.0006, 21.8115, 26.291 and 20.5045%, respectively. This proves the increased efficiency of Lattice design over that of classical RCBD under SARI field condition. For the maize variety trial dataset containing 81 treatments with 3 replications, alpha lattice design is found to be more efficient than RCBD having relative efficiency of 18.8%. The CV of alpha lattice design is 21.1% while that of RCBD 22.9%. The relative efficiencies of three datasets and their corresponding CVs respectively signify that the precision of experiment increased significantly using incomplete block designs instead of completely blocked designs mainly when the numbers of treatments are increased tremendously.

Based on the results of this study, under SARI field setup, we conclude that RCBD is more efficient than CRD, lattice design and alpha lattice designs are more efficient than classical RCBD. In order to increase the precision of agricultural field experiments researchers are advised to use RCBD for small number of treatments; lattice and alpha lattice designs whenever there are large number of treatments taking into considerations the nature of field conditions.

List of Acronyms

- ANOVA: Analysis Of Variance
- BIBD: Balanced Incomplete Block Design
- BLD : Balanced Lattice Besign
- CRD: Completely Randomized Design
- CV: Coefficient of Variation
- df: Degree of freedom
- IBD: Incomplete Block Design
- LSD: Least Square Deviation
- MSE: Mean squares Error
- MST: Mean Squares Treatment
- PBIBD: Partially balanced incomplete block design
- PBLD : Partially Balanced Lattice Design
- QQ plot: Quantile – Quantile plot
- RCBD: Randomized Complete Block Design
- R.E: Relative Efficiency
- SARI: Southern Agriculture Research Institute

Table of Contents

List of Tables

List of Figures

CHAPTER ONE: INTRODUCTION

1.1 Background of the study

Experimentation plays a momentous role in the field of agriculture. A good experiment is one which involves good planning, accurate data collection, proper data analysis and precise interpretation of the data. A statistician is supportive in drawing inferences and conclusions from the experiment. However, before that the researcher must properly define the objectives of the experiment. Agronomists would like to choose an experimental design that maximizes the amount of information that is obtained from a fixed number of observations. To determine the optimal design among a set of candidates, it is necessary to define some criteria which allow discrimination between possible designs. Experimental design could be considered as the crucial state of any experiment due to its aim to ensure that the experimenter is able to detect the treatment effects that are of interest by using the available resources to obtain the best possible precision.

Precision is the ability of an experiment to detect a true treatment effect. We can improve this precision by increasing the replication, proper allocation of treatments improved technique to reduce the variability among units treated alike, increasing the size of experimental units, the use of covariance, and the employment of a more efficient experimental design and method of analysis [27].

Design of experiments forms the backbone of any research endeavor in the discipline of agriculture and clinical trials. The foundations of the statistical approach to experimentation were laid by R.A. Fisher in the early 1930s. The subject evolved in agriculture but is now applicable in almost all sciences, engineering and arts**.** The aim of an experiment is to compare a number of treatments on the basis of the responses produced in the experimental material. The confidence and accuracy with which treatment differences can be assessed will depend to large extent on the size of the experiment and on the inherent variability in the experimental material. Hence, design of experiments is an essential component of research in agriculture. In order to make research globally competitive, it is essential that sound statistical methodologies be adopted in the data collection and analysis [40].

In any experimental design, treatments are likely to be administered on experimental units under the same condition. However, a difference among experimental units is inevitable to occur and this is called an experimental error (residual). This error is primarily the basis for deciding whether an observed difference is real or due to chance. In other words responses from each treatment are obtained from different units called replications and they are essential for the estimation of experimental error. Replications also help to improve precision of an experiment by reducing the standard error of a mean or of a difference between means. Replication together with randomization will provide a basis for estimating the error variance. The control of experimental error is another aspect of experiment that needs attention. Intuitively one can anticipate an increase in treatment difference produced if there is a sizeable reduction in experimental error. To realize this, ne way of controlling error is by blocking or putting together similar experimental units in the same group and randomly assigning all treatments into each block separately and independently. The purpose of randomization is to prevent systematic and personal biases from being introduced into the experiment by the experimenter.

The main technique adopted for the analysis and interpretation of data collected from an experiment is the Analysis Of Variance (ANOVA) technique that essentially consists of partitioning the total variation in an experiment into components assigned to different sources of variation due to the controlled factors and error. A design for agricultural trials must provide valid error terms and sufficient precision for the effects of interest. As Drane (1989) stated the manner in which the experiment is designed and executed determines what constitutes the experimental unit, the proper error terms in the ANOVA, and whether replication is either possible or desirable [11].

The designed experiments in this study are analyzed by ANOVA with the following two purposes.

- To partition or decompose total variation in the response variable into separate components, each component representing a different sources of variation, so that the relative importance of the different sources can be assessed.
- To give estimate of the underlying variation between experimental units in given treatment that provides the basis for inference about the effects of treatments.

This second purpose is a measure of experimental error which provides the basis for interval estimates and significance tests. The variance or more correctly, the mean square associated with each of the other sources of variation may be compared with the experimental mean square error. This comparison provides F statistic for testing the significance of the difference among means for the particular variance source. In addition, ANOVA provides information from which standard errors of means and differences may be computed, and from which interval estimates may be constructed.

Most popular method used to compare the performance of one design over the other design is relative efficiency. Efficiency is measured by the variance of the estimated treatment differences which depend on the design and the within-block variation, and it is estimated by the residual mean square. The efficiency of one design of experiment over another is usually measured in terms of reduced error variance, expected mean squared error, or average standard error of the difference between treatments means [12].

The efficiency of designs is compared in all locations to assess the efficiency of each design mainly their performance with respect to minimizing experimental error, coefficient of variation (CV) and mean squared error for yield.

The CV affects the degree of precision with which the treatments are compared and is a good index of the reliability of the experiment. It is an expression of the overall experimental error as percentage of the overall mean; thus, the higher the CV value, the lower is the reliability of the experiment.

1.2 Statement of the problem

Most of the time because of limited plot size, in field experiments, agricultural researchers will not use complete block designs, mainly when there are large number of treatments. As a result, most agronomists try to use different incomplete block designs such as lattice and alpha lattice designs.

In this study, we address the following research questions:

- What are the conditions to make choice among those incomplete block designs (IBD)?
- What will be the efficiency of those designs as compared to the classical RCBD under SARI field setup?
- What will be the efficiency of design when there are missing values in the dataset?
- Is the relative performance of such designs studied and documented for the case of SARI?

1.3 Objectives of the study

1.3.1 General objective

- The general objective of this study is to assess statistical performance of incomplete and complete block designs comparison of RCBD, lattice and Alpha lattice designs in field trials of Southern Agricultural Research Institute (SARI)

1.3.2 Specific objectives

The specific objectives of this study were:

- evaluating the performance of the three most commonly used experimental designs complete and incomplete block designs namely; RCBD, Lattice design and alpha lattice design in field setup of SARI.
- assessing ways of estimating missing values in RCBD
- to support the theoretical justifications mainly using different datasets from SARI, comparing different experimental designs.

1.4 Significance of the study

The result of the study will contribute to:

- identify the appropriate and efficient experimental designs for field experiments in field setup like SARI
- improve the precision of agricultural field experiments through using appropriate design and analysis.

CHAPTER TWO: LITRETURE REVIEW

Design of experiments had its origin in supplying layout plans of experiments for comparison among a number of experimental treatments in regard to some of their responses when these are applied to a set of experimental units under certain conditions. Mainly, the objective of experimental design is to select and group the experimental material so that the experimental error in the experiment is reduced.

The main purpose of conducting field experiment is to compare effectiveness of different treatments. Precision and accuracy are vital, but valid assessment of error is also crucial thing to be considered. This is why, for example, yield is influenced by non-treatment factors such as pests and soil fertility. If these factors are ignored, extraneous variation leads to erroneous comparisons. Proper field design and statistical analysis will also help minimize this problem. Classical methods for controlling such extraneous variation include replication, blocking, and randomization. Here, the first two replication and blocking help to increase precision in the experiment, while the last one randomization is used to decrease bias of the experimenter [21].

Most agronomic field experiments are being conducted using the concepts of replication, local control (blocking) and randomization [2].

Replication is used for the purpose of increasing in precision by reducing standard error and increases representation since wider area is used. Without replication there is no estimation of experimental error [3].

Randomization is used in field experiments in order to avoid systematic, selection, accidental biases and to avoid the subjective bias of the experimenter. It should be used whenever possible and practical so as to eliminate or at least reduce the possibility of confounding effects that could render an experiment practically. That is, randomization ensures that no treatment is consistently favored or discriminated being placed under best or unfavorable conditions, thereby avoiding

bias. It also ensures independence among observations, which is a necessary condition for validity of assumption to provide significance tests and confidence intervals.

Blocking is grouping of experimental units into blocks or groups of more or less uniform experimental units. So, experimental units within the same block are homogeneous.

Effective blocking not only yields more precise results than an experimental design of comparable size without blocking, but also increases the range of validity of the experimental results.

There are different experimental designs that are being used in agricultural field experiments. These include Complete Randomized Design (CRD), Randomized Complete Block Design (RCBD) and Incomplete Block Design; lattice designs are the most frequently used.

The most common type of experimental design for making inferences about treatment means is the completely randomized design (CRD), where all treatments under investigation are randomly allocated to the experimental units. CRD is appropriate for testing the equality of treatment effects when the experimental units are relatively homogeneous or the experiment is conducted under controlled environment. When the experimental units are heterogeneous, the notion of blocking is used to control the extraneous sources of variability. The major criteria of blocking are characteristics associated with the experimental material and the experimental setting [40].

As the size of block increases, variance per unit for variety contrast increases and ultimately leads to inefficient estimates of precision. Effective control of error variance usually requires relatively small blocks [13]. Under such circumstances, the use of RCBD becomes questionable.

RCBD is one of the most frequently used experimental designs, mainly due to the following merits: any number of treatments and replications can be included; the statistical analysis is easy and it provides information on the uniformity of experimental units.

Incomplete Block Design (IBD): if in a randomized block designs; the number of experimental units in a block is less than the number of treatments. Obviously in such designs one or more treatment block combinations are missing. The analysis of IBD is different from the analysis of Complete Block Designs in that comparisons among treatment effects and comparisons among block effects are no longer orthogonal to each other.

Incomplete block designs (IBD) occur as balanced or partially balanced. In balanced incomplete block designs all pairs of treatments occur together within a block the same number of times. Since each block does not contain all treatments, block and treatment effects are confounded[37]. Incomplete block designs such as lattice designs provide more precise estimates when the homogeneity condition does not hold, mainly when there is large number of treatments in the experiments.

Lattice designs are extensively used in agricultural field experiments especially for varietal trials. These designs are resolvable, but the requirement that the number of treatments be a complete square is a limitation. A block design is resolvable if the blocks can be partitioned into replicates, defined as sets of blocks with the property that each treatment is assigned to one unit in each set [46].

Yates (1936) reported that RCBD is the most popular design for field experiments.

Of the 414 agronomic field experiments in USA, the majority (72%) were implemented as RCBD [44]. They further described that the vast majority (96.7 %) of agronomic field experiments conducted by agronomists are implemented through RCBD for their simplicity and intuitive layout.

In the class of equally replicated designs with ***v*** treatments, ***b*** blocks and a common block size ***k***, a balanced incomplete block (BIB) design whenever existent, is the most efficient design for making tests versus control comparisons according to various efficiency criteria.

In a RCBD every treatment appears in every block precisely once. RCBD is the most efficient design because there is no loss of information in estimating treatment contrasts as well as block contrasts. RCBD is affordable when the block size contains small treatments.

Randomized block, Latin square, and other complete block types of experiments are inefficient for large number of treatments, because of their failure to adequately minimize the effect of experimental unit heterogeneity [30]. Generally, the greater the heterogeneity within blocks, the poorer the precision of variety effect estimates.

Additional improvement is possible through modeling field variability using spatial features of the field layout. It has been advocated that use of incomplete blocking is generally more effective in reducing the unexplained structured variation in comparison with complete blocking. They are more flexible than lattice designs and can accommodate any number of varieties. The advantage of alpha designs is that they are easy to construct, and can be constructed in cases where balanced incomplete block designs and lattice designs do not exist. The early alpha designs were aimed primarily at controlling variation down the columns of experimental units in the field. This is often adequate when experimental units are long and narrow [44].

Mandefro(2005) compared efficiency of alpha lattice design with RCBD and the results indicated that alpha lattice design improved the efficiency 8 to 9 percent as compared to RCBD mainly when there is large number of treatments [29].

Yates (1936) reported that the use of alpha lattice design in an international yield trials of different crops and found average efficiency 18 percent higher than the RCBD [46].

Gunjaca *et al* (2005) studied the efficiency of alpha lattice designs in Croatian variety trials of cereal and non-cereal variety trials composed of 152 data sets and found that the maximum relative efficiency of alpha lattice design compared to RCBD in cereal and non cereal varieties were 1.37 and 1.55 respectively. Here, the alpha lattice design increased the precision of the two variety trials by 37% and 55% respectively [18].

According to Snyder (1962) study, based on three data sets and they found that for their three data sets alpha lattice design increased the precision of the experiments by 26%, 17% and 55%, respectively. Alpha designs were used for field trials mainly because they provide better control on experimental variability among the experimental units under field conditions [43].

Hatfield (2000) showed that general lattice design (alpha lattice design) was on average more efficient than complete block analysis in reducing the mean square error when there are large number of treatments [21]. Alves *et al.* (2009) compared the efficiency of RCBD, alpha-design, and row-column design in genotypic mass selection. Their result indicated that greater efficiency for alpha-design and row-column design, enabling more precise estimates of genotypic variance, greater precision in the prediction of genetic gain and consequently greater efficiency in genotypic mass selection [1].

Patterson *et al.* (1976) reported the efficiency of alpha lattice designs relative to other incomplete block designs. Using a large collection of experiments, they have shown that alpha designs on the average produced a 30% gain in efficiency over designs which did not use incomplete block designs. They also reported that the use of generalized lattice designs (alpha lattice designs) instead of complete block designs. In 244 cereal variety trials grown in UK has resulted in average reduction of 30% in variances of varietal yield differences.

Historically agronomists have relied heavily on the CV as a measure of trial's reliability and thereby to see the efficiency of their designs.

But, it should be noted that the CV varies with the type of experiments and the characteristics measured. According to Gomez and Gomez (1984) the acceptable range of CV is: 6 to 8% for variety trials; 10 to 12% for fertilizer trials and 13 to 15% for insecticide and herbicide trials. Furthermore, they pointed out that in field experiment CV for yield is about 10%, that for tiller number is about 20%, and for plant height CV is about 3% [17].

CHAPTER THREE: STUDY METHODOLOGY

3.1 DATA

This study used data from South Agricultural Research Center (SARI) one soybean yield trial and two maize yield trials conducted at different locations. The trials were conducted at different research Centers of the region using RCB, lattice and alpha lattice designs.

The soybean trial was conducted using RCBD with three replications at five different locations; namely Hawassa, Areka, Gofa, Inseno and Bonga in 2007. The maize variety trial was conducted using 5×5 partially balanced lattice design with four replications at Hawassa, Areka, Bonga, Jinka and Arba-Minch centers of SARI in 2008/9. Maize variety trial was also conducted using alpha lattice design at Hawassa research center in 2008/9. The last experiment was laid out with 3 replications, 81 treatments, 9 blocks and 9 plots per block.

3.2 METHODOLOGY

In this part the methodologies for the data analysis using each design was discussed in detail.

3.2.1 Commonly used Experimental Designs under Ethiopian Context in Field Conditions

The commonly used experimental designs in National and Regional Agricultural Research Institutes are completely randomized design, randomized complete block design, lattice design and alpha lattice designs mainly for factorial and split-plot treatment structure [15, 28].

3.2.1.1 Completely Randomized Design (CRD)

This design is the simplest design from the standpoint of assignment of experimental units to treatments or treatment combinations. In this design, the treatments are allotted to experimental units entirely at random or by chance as the single group and the units forming the group should be homogeneous. So, this design is mostly recommended for controlled experiments such as laboratory or greenhouse experiments.

The ANOVA Model: $Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1,2,3,\dots,t \\ j = 1,2,3,\dots,r \end{cases}$

where, Y_{ij} is the i^{th} observation on j^{th} treatment; μ is the overall treatments average response to a mean; τ_i is the i^{th} treatment effect; ε_{ij} is the random error associated with the j^{th} experimental unit of the i^{th} treatment.

The model assumption for the ANOVA of CRD:

$E(\varepsilon_{ij})= 0$ observations within a treatment have the same mean for every i and j

$Var(\varepsilon_{ij})= \sigma^2$ all observations in different treatments have the same variance, namely, σ^2

Furthermore, we assume the ε_{ij} are uncorrelated.

Table 1: ANOVA table for CRD with t treatments

Source of variation	Degree of freedom	Sum of squares	Mean squares	F-value
Treatment	*t-1*	$\frac{\sum y_{i.}^2}{r} - C$	$\frac{Treatmant\ SS}{t-1}$	$\frac{MSTt}{MSE}$
Error	*n-t*	$\sum\sum y_{ij}^2 - \frac{\sum x_{i.}^2}{r}$	$\frac{SS\ error}{n-t}$	
Total	*n-1*	$\sum\sum y_{ij}^2 - C$		

Where, *t* = treatment number, *n* = total number of entries C= $\frac{G^2}{n}$ and G is the grand total of the treatment. The mean square for the treatment and error can be calculated by dividing the sum of squares of the corresponding variations by their df.

To find grand mean = $\frac{G}{n}$ and CV = $\frac{\sqrt{Mean\ square\ Error}}{Grand\ mean} * 100$

3.2.1.2 Randomized Complete Block Design (RCBD)

In agricultural research, the experimental units, often being plots of land or animals, will by their very nature be different from place to place or animal to animal etc.

RCBD is one of the most widely used experimental designs in agricultural research. It is the most common and extensively used block design when the treatments are the several levels of a factor and also it is the most efficient design because there is no loss of information in estimating

treatment contrasts as well as block contrasts. This design is a restricted randomization design in which the experimental units are first sorted into homogeneous groups, called *blocks* and the treatments are then assigned at random within blocks. The major reason for grouping plots (experimental units) into uniform blocks is to reduce plot to plot variation and to improve the precision of the experiment. Failure to adequately block a field can result in unacceptably large error variance and/or biased estimates of treatment effects [13].

The major advantages of this design are its accuracy of results, flexibility of design and ease of statistical analysis.

Blocking will increase treatment precision only if plots are blocked according to one or more varying external factors. If an experimental area is homogenous, blocking may actually decrease the precision of estimating treatment effects. This results from a larger mean square error (MSE) term in the ANOVA since error degrees of freedom are reduced without a comparable reduction in sum of square error (SSE). In this situation, (CRD) would give more precisely estimate of treatment effects than a RCBD [23].

The statistical model for RCBD is: $Y_{ijr} = \mu + \alpha_i + \beta_j + \varepsilon_{ijr}$

where i = 1,2,3,….,a , j =1,2,3….,b

Where Y_{ij} is the ith observation in the jth block and μ is an overall mean, α_i is the effect of the ith treatment, β_j is the effect of jth block, and ε_{ij} is a random error component.

Assumptions:

- The mathematical model $Y_{ijr} = \mu + \alpha_i + \beta_j + \varepsilon_{ijr}$ is additive
- α_i is the (additive) effect of the ith treatment and β_j is the (additive) effect of the jth block,
- μ and α_i are fixed parameters and β_j may be fixed or random effects.
- As usual, the treatment and block effects are subject to the restrictions that $\sum_{i=1}^{a} \alpha_i = 0$ and $\sum_{j=1}^{b} \beta_j = 0$, respectively

- ε_{ij} Distributed normally and independently with mean 0 and σ^2 i.e. $\varepsilon_{ij} \sim iid$ N(0, σ^2)

ANOVA table for RCBD with *t* number of treatments, *r* replications and *b* number of blocks is given as follow:

Table 2: ANOVA for RCBD

Source of variation	Df	SS	MS	F
Treatment	(t-1)	$\frac{\sum T_i^2}{b} - C$	$\frac{\text{Treatmant SS}}{\text{t} - 1}$	$\frac{MSTreatment}{MS\ Error}$
Block	(b-1)	$\frac{\sum T_i^2}{t} - C$	$\frac{\text{Block SS}}{(\text{n} - 1)(\text{t} - 1)}$	$\frac{MSblock}{MS\ Error}$
Error	(t-1)(b-1)	$\sum_{j=1}\sum_{i=1} Y_{ij}^2 - \frac{\sum T_i^2}{b}$	$\frac{\text{Error SS}}{(\text{n} - 1)(\text{t} - 1)}$	
Total	tb-1	$\sum_{j=1}\sum_{i=1} Y_{ij}^2 - C$		

Where, C = $\frac{G^2}{N}$, G is the grand treatment total where T_i is the treatment total for the i^{th} experimental unit

To find grand mean = $\frac{G}{n}$ and CV = $\frac{\sqrt{\text{Mean square Error}}}{\text{Grand mean}} * 100$

Relative efficiency of RCBD over CRD: R.E = $\frac{(\text{r}-1)\text{MS}_\text{B}^2+(\text{t}-1)\text{MS}_\text{E}^2}{(\text{rt}-1)\text{MS}_\text{E}^2}$

It should be noted that when the error dfs is less than 20, Fisher (1974) proposed an adjustment to account for the discrepancies in df. He suggests that the R.E parameter be multiplied by an adjustment factor as:

$$\text{R.E} = \frac{(\text{r}-1)\text{MS}_\text{B}^2+(\text{t}-1)\text{MS}_\text{E}^2}{(\text{rt}-1)\text{MS}_\text{E}^2} \text{ x } \frac{[(\text{r}-1)(\text{t}-1)+1][\text{t}(\text{r}-1)+3]}{[(\text{r}-1)(\text{t}-1)+3][\text{t}(\text{r}-1)+1]}$$

The analysis of incomplete block designs is different from the analysis of complete block designs in those comparisons among treatment effects and comparisons among block effects are no longer orthogonal to each other.

3.2.1.3 Lattice Designs

Historically, lattice designs were developed for large-scale agricultural experiments (Yates, 1936b) in which large numbers of varieties were to be compared. The main application since then has been and continues to be in agriculture.

Even though this limits the number of possible designs, lattice designs represent an important class of designs nevertheless, in particular when one is dealing with a large number of treatments. In certain types of agronomic experiments the number of treatments can easily be 100 or more, for example, in breeding experiments. These designs are referred to as quasi - factorial or lattice designs. These designs are the most commonly used in agricultural research when the number of treatments to be tested is significantly large. But, if small, (say less than ten), use of an ordinary RCBD or Latin square design may be appropriate according to the situation of the experiment. However, when the number of treatments tested is large, as is often the case with varietal trials or breeding experiments, use of RCBD may not be appropriate because of the increase in error variance due to the larger block size. IBD, including lattice design facilitates the comparison of a large number of treatments which are assigned to incomplete blocks within replications. In lattice design, the number of treatments must be an exact square and the number of units in each block is the square root of the number treatments. Lattice designs reasonably uses small block size in order to ensure that each block does not lose its homogeneity due to the large size. And also each block does not contain all treatments. The existing lattice designs can be classified according to: Number of treatments, t; Block size, k Number of different systems of confounding used; Number of restrictions imposed on randomization

Based on the above criterion, the two most commonly used lattice designs are: Balanced Lattice Besign (BLD) and Partially Balanced Lattice Design (PBLD).

3.2.1.3.1 Balanced Lattice Design (BLD)

In BLD, the number of treatments must be a perfect square and the block size is equal to the square root of the number of treatments. The number of replications in this design is one more

than the block size. Incomplete blocks are combined in groups to form separate replication. The special feature of this design, as distinguished from other lattices, is that every pair of treatments occurs once in the same incomplete block. Consequently, all pairs of treatments are compared with the same degree of freedom. However, if there are *k* blocks, there must be *k*+1replications to achieve the balance. This restriction in the number of replications and treatments makes the design less practical and more restrictive.

Computational Procedure of Balanced Lattice Design:

For *k* number of blocks, *k+1* replication df, SS and MS for each source of variation can be computed as: Replication = k= r-1, Treatment (Unadj.) = k^2-1 , Block (adj) = k^2-1

Intra-block error = (k-1)*(k^2-1) , Treatment (adj.) = k^2-1, Effective error = (k-1)(k^2-1)

Correction factor = $C = \frac{G^2}{(K)^2(K-1)}$ Total sum of squared= $\sum X_{ijk}^2 - C$

Replication sum of square = $\frac{\sum R^2}{K^2} - C$ Treatment(Unadj.) Sum of square= $\frac{\sum T^2}{(K+1)} - C$

Block(adj.)SS = $\frac{\sum W^2}{(K^3)(K+1)}$

Intrablock error SS = Total SS - [Replication SS+ Treatment (Unadj.) SS+ Block(adj.)SS]

Compute mean squares for the treatment, block(adj.) and Intrablock error as:

Treatment(unadj.)MS = $\frac{\text{Treatment (Unadj.) Sum of square}}{K^2-1}$

Block(adj.)MS= $\frac{\text{Block(adj.)SS}}{K^2-1}$, Intrablock error MS= $\frac{\text{Intrablock error sum of square}}{(k-1)(K^2-1)}$

Having obtained the values for the mean squares, we are now in the right position to compute the adjusted treatment total ($\dot{T}$)

$\dot{T} = T + \mu W$, where $\mu = \frac{\text{Block(adj.)MS} - \text{Intrablock error MS}}{K^2[\text{Block(adj.)MS}]}$

But, this computation is necessary only if the Intrablock error mean square is less than the block (adj.) mean square. In such conditions, the adjusted treatment totals ($\dot{T}$) for all treatments and the

effective error mean square should be computed. They will in turn be used in performing the effective error mean square is as follows.

Treatment (adj.)MS = $\left[\frac{1}{(k-1)(K^2-1)}\right]\left[\sum T^2 - \frac{G^2}{K^2}\right]$

Effective error MS = Intrablock error MS(1+kμ) ; F = $\frac{\text{Treatment (adj.)MS}}{\text{Effective error MS}}$

If the Intrablock error mean square, on the other hand, is greater than the block (adj.) mean square, the value of μ is taken to be 0 and, therefore there are no further adjustments necessary to the treatments. The F-test of significance is computed as the ratio of treatment(unadj.) mean squares to the Intrablock error mean square.

Comparing the F with tabular F value we can conclude that whether there is significant difference among the treatments or not.

We can determine the degree of precision with which the treatments are compared by computing:

CV = $\frac{\sqrt{Intrablock\ error\ MS}}{Grand\ Mean}$X100

Relative efficiency to estimate the precision relative to RCBD is computed as:

R. E = $\frac{\text{Block(adj.)SS+Intrablock error SS}}{(k-1)(K^2-1)(\text{Intrablock MS})}$X100

3.2.1.3.2 Partially Balanced Lattice Design

Partially balanced lattice Design is developed by Bose and Nair (1939) to overcome the problems associated with the restrictive assumptions of the balanced lattice design [40].

The number of replications required for balanced lattice becomes very large as the number of treatments increases. For this reason it is not usually practical to use balanced lattices for blocks with more than about seven units per block. In the interest of economy, then, the scientist is forced to accept a partially balanced design with fewer replications than would be required for full balance.

In partially balanced lattice designs, the number of replications is not restricted, but the number of treatment must be a perfect square and the block size is equal to the square root of the number of treatments. However, not all treatments occur together in the same block. This leads to

differences in precision with which some comparisons are made relative to other comparisons. The names of the sub categories of partially balanced lattice design follow the number of replications. For example, the balanced lattice with two replications is called *simple lattice*, with three replication *triple lattice*, with four replications *quadruple lattice* and so on. The pattern of statistical analysis is the same for simple, triple, and quadruple lattices.

Table 3: ANOVA table for Partially Balanced Lattice Design with *r* replications, *k* block size and $t = k^2$ treatments

Source of variation	Df	SS	MS	F
Replication	r-1	Replication SS	Replication MS	$\frac{\text{Replication MS}}{\text{Intrablock MS error}}$
Block (adj.)	r(k-1)	Block (adj.) SS	Block (adj.) MS	$\frac{Block\ (adj.)\ MS}{Intrablock\ MS\ error}$
Treatment (unadj.)	k^2-1	Treatment (unadj.) SS	Treatment (unadj.) MS	$\frac{\text{Treatment (unadj.) MS}}{Intrablock\ MS\ error}$
Intrablock error	(k-1)(rk-k-1)	Intrablock SS error	Intrablock MS error	-
Total	rk^2-1	Total SS		-

The sum of squares for total, replication, treatment and error are computed as in any other designs. The sum of squares due to block is a new statistic to be computed in lattice designs.

Correction factor: C.F = $\frac{(GT)^2}{rq^2}$ Total SS= $\sum\sum X_{ij(l)}^2 - C.F$, SS of replication= $\frac{\sum R_j^2}{q^2} - C.F$

SS of block(adj.) = $\frac{\sum\sum C_{ij}^2}{qr(r-1)} - \frac{C_i^2}{q^2r(r-1)}$, SS of treatment = $\frac{\sum T_i^2}{r} - C.F$

Intrablock error SS = Total SS – SS of replication- SS of block - SS of treatment

The mean square of the block and error are computed as usual dividing the sum of squares of block and error by their respective degrees of freedom.

$E_b = \frac{SS_B}{r(q-1)}$ and $E_e = \frac{SS_E}{(q-1)(rq-q-1)}$

Then, these two mean squares are compared either to go for adjustment factor or not. If $E_b \leq E_e$, then adjustment for block has no effect. This will lead us to ignore the blocking restriction and analyze the data as if the design had been a randomized block design with replications as blocks. If $E_b > E_e$, an adjustment factor μ, is computed for the design: $\mu = \frac{E_b - Ee}{qE_b}$

Finally, the effective error mean square that can be used in calculating t-test and interval estimates is calculated as: $E'_e = \frac{1+rq\mu}{(q+1)E_e}$

Adjusted treatment mean square is computed to test whether there is significant difference among adjusted treatment means or not. In order to do that, it is necessary to compute first the unadjusted blocks within replications sum of squares ($SS_{B(unadj)}$).

$$SS_{B(unadj)} = \frac{\sum\sum B_{il}^2}{q} - C.F - SS_R$$

Then the adjusted treatment sum of squares, ($SS_{t(adj)}$), is computed as

$$SS_{t(adj)} = SS_{t(unadj)} - \mu q \left(\frac{2}{1+q\mu}\right)\left[SS_{B(unadj)} \Big/ (r-1)(1+\mu q) - (SS_{B(adj.)})\right]$$

Computing the mean square of treatment: $MS_t = \frac{SS_{t(adj)}}{t-1}$ we can compute F= $\frac{MS_t}{E'_e}$

To find the relative precision over RCBD: $MS_{in\ RCBD} = \frac{SS_{B(adj)} + SS_E}{block\ d.f + error\ d.f}$

Effective $MS_E = \frac{E_e}{r} + \left[\frac{1+k\mu}{(k+1)}\right]$

Relative efficiency (R.E) of lattice design over RCBD: R.E = $\frac{MS_{in\ RCBD}}{\text{Effective } MS_E}$,

% Efficiency = (R.E)*100

3.2.1.4 Alpha-Lattice Design

These designs, called α-designs, were introduced by Patterson and Williams (1976) and further developed by John and Williams (1995) to be used mainly in the setting of variety trials in agronomy. Alpha lattice designs are available for many (*r*,*k*,*s*) combinations where *r* is the number of replicates, *k* is the block size and *s* is the number of blocks per replicate (the number of treatments *t*=*ks*). Efficient alpha designs exist for some combinations for which conventional lattices do not exist. It can also accommodate unequal block sizes. This design bridges the gap between RCBD and lattice designs. That is it has an additional feature in that the number of treatments should not necessarily be a perfect square. Thus, the development of alpha-lattice designs removed the restrictions on the number of treatments to be considered and its relation with block size required for lattice designs.

The linear model of observations in alpha design is of the form:

$\boldsymbol{y}_{ijk} = \boldsymbol{\mu} + \mathbf{t}_i + \mathbf{r}_j + \mathbf{b}_{jk} + \boldsymbol{e}_{ijk}$

Where y_{ijk} denotes the value of the observed trait for i^{th} treatment received in the k^{th} block within j^{th} replicate (superblock)

$\mathbf{t}_i$ is the fixed effect of the i^{th} treatment ($i = 1,2,\ldots,t$)

$\mathbf{r}_j$ is the effect of the j^{th} replicate (superblock) ($j = 1,2,\ldots,r$)

$\mathbf{b}_{jk}$ is the effect of the k^{th} incomplete block within the j^{th} replicate ($k = 1,2,\ldots s$).

$\mathbf{e}_{ijk}$ is an experimental error associated with the observation of the i^{th} treatment in the k^{th} incomplete block within the j^{th} complete replicate.

ANOVA for an alpha-lattice design with *t* number of treatments, *b* number of blocks within replication, and *r* number replications is given in the following table:

Table 4 ANOVA for an Alpha-lattice Design

Source of variation	Df	SS	MS	F
Replication	r-1	Replication SS	MS Replication	$\frac{\text{MS Replication}}{\text{MS Error}}$
Block(replication)	rb-1	Block(replication) SS	MS Block (Replication)	$\frac{\text{MS block(Replication)}}{\text{MS Error}}$
Treatment (adj.)	t-1	Treatment (adj.) SS	Treatment (adj.) MS	$\frac{\text{Treatment (adj.) MS}}{\text{MS Error}}$
Error	rt-rb-t+1	SS error	MS Error	
Total	rt-1	SS total		

The procedure to compute the SS and MS for the different source of variations in the alpha lattice designs is almost the same as of the lattice designs using their corresponding df.

3.2.2 Estimating Missing Data in RCBD

In RCBD, sometimes an observation in one of the blocks might be missing. This may happen due to carelessness or error for reasons beyond our control such as unavoidable damage to

experimental units by rodent, water lodging...etc. A missing observation introduces a new problem into the analysis since treatments are no longer orthogonal to blocks; that is, every treatment does not occur in every block. There are two general approaches to missing values analysis: Approximate analysis and exact analysis [BIBD]

Approximate analysis

Suppose one observation (x) is missing and let

$x'_{..}$ = grand total with missed value, $x'_{i.}$ = treatment total with missed value

$x'_{.j}$ = Block total with missed value

x is estimated by minimizing the contribution of x to sum of squares error (SSE)

$\hat{x} = \frac{ax'_{i.}+bx'_{.j}-x'_{..}}{(ab-a-b+1)} = \frac{ax'_{i.}+bx'_{.j}-x'_{..}}{(a-1)(b-1)}$

Where $\hat{x}$ is estimate of the missing observation, $x'_{i.}$ is the sum of the remaining observations on the treatment with the missing value, $x'_{.j}$ is the sum of the remaining values in the block with the missing observation and $x'_{..}$ is the grand total of all available observations; a and b are the numbers of replicates and treatments respectively.

When there are several missing values, for units $x_1, x_2, x_3, x_4 \ldots$, we first assign initial values for $x_2, x_3, x_3, \ldots$, and use formula above to find an approximation for x_1.

Using this approximated value and the values previously assumed for $x_3, x_4, x_5, \ldots$, we again use formula to insert an approximation for x_2.

Exact analysis [BIBD]

Model: $Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$, $\begin{cases} i = 1,2,3, \ldots., \text{a} \\ j = 1,2,3, \ldots., \text{b} \end{cases}$

- Because of incompleteness, all Y_{ij} don't exist.
- Var$(\tau_i - \tau_i)$ is constant or homogeneity of variance assumption

- Additive effect of block due to block, i.e. no interaction effect
- Usual treatment and block restrictions $\sum_{i=1}^{a} \tau_i = 0$ And $\sum_{j=1}^{b} \beta_j = 0$
- Non – Orthogonality of blocks and treatments

Table 5: ANOVA of BIBD for RCBD

Source of variation	df	SS	MS	F
Block(Unadj.)	(b-1)	SSBlock(unadj.)	MS_{Block}	$\frac{MSBlock}{MSError}$
Treatment(Unadj.)	(t-1)	SSTreatmentunadj.	$MS_{Treatment}$	$\frac{MSTreatment}{MSError}$
Block(adj.)	(b-1)	SSBlock(adj.)	$MS_{block(adj)}$	$\frac{MS_{block(adj)}}{MSError}$
Treatment(adj.)	(t-1)	SSTreatment(adj.)	$MS_{Treatment(adj.)}$	$\frac{MSTreatment(adj.)}{MSError}$
Error	N-t-b+1	SSError	MS_{Error}	
Total	N-1	SSTotal		

3.2.3 Combined Analysis of Several Experiments

Combined analysis is done for experiments repeated at several locations such as the case at hand in SARI dataset. The basic steps in the combined analysis of data and from experiments repeated in both time and space are similar for those designs discussed earlier.

Individual analysis of variance is computed for each location in each season. The error variances across the locations are checked for their heterogeneity.

Finally, an appropriate combined analysis is completed and interpreted. The error mean square in the ANOVA is the sample estimates S^2 of the error variance for these trials. These estimates provide the data for examining the homogeneity of variance.

The first approach is the quick test developed by Hartley (1950) used to test the homogeneity of variance is provided by the ratio of the largest to the smallest S^2 in the set. It is often possible to draw a conclusion regarding homogeneity of variance without further testing.

The test statistic is calculated as: $F=\frac{S^2_{max}}{S^2_{min}}$ and this ratio statistic can be compared with the tabulated value of F_{max} with $a\ and\ \gamma$ dfs. Then, the decision on null hypothesis of homogeneity of variance will be made [22].

An alternative procedure which is more sensitive than the ratio test is the Bartlett's test of homogeneity of variance (Bartlett, 1937). This test based on the natural logarithm of the sample variances, has been described by Snedecor and Cochran (1980). To perform this test, let:

γ: Error degrees of freedom for the individual trial; S_i^2: Error means square at location i

a: Number of locations Then, $M = \gamma[a(\ln S^2) - \sum_i \ln S_i^2]$, $S^2 = \frac{\sum_i S_i^2}{a}$ and $C = 1 + \frac{a+1}{3a\gamma}$

The ratio $\frac{M}{C}$ is a test statistic for the null hypothesis that each S^2 with an estimate of σ^2. The ratio $\frac{M}{C}$ is distributed as X^2 with $a - 1$ df.

With this analysis we look at the magnitude of among-location variation, the variation among treatments, and in particular, the location X treatment interaction. The test of location X treatment interaction gives indication of whether or not the treatments behave the same from one location to another. A significant interaction means that the effects of treatment vary from location to location. But, in this case, the combined analysis of data from all observation has little meaning.

A non-significant location X treatment interaction on the other hand doesn't necessarily mean that all of the meaningful comparisons among treatments are independent of location.

3.2.4 Design Efficiency

In testing treatment differences, several alternative experimental designs may be used. However, the several designs that may be equally valid for testing treatment effects are rarely equally efficient. A commonly used index for comparing the efficiency of two different designs is the inverse ratio of the variance per unit, i.e., the MSE's. Since different designs may have different

degrees of freedom for error, a correction factor, suggested by Fisher (1937), which multiplies the inverse ratio of variances, will give a better measure of the R.E.

The success of blocking is best measured by the relative efficiency of the RCBD as compared with that of the CRD. The most widely used measure of R.E is the relative precision defined as follows. The R.E of the CRD relative to a classical RCBD is computed as:

$$\text{R.E} = \frac{\text{Mean square error in CRD}}{\text{Mean square error in RCBD}} \times 100$$

The R.E of the PBLD relative to a comparable RCBD is computed as [17]

$$\text{R.E} = \left[\frac{Block(adj.)SS + Intrablock\ error\ SS}{r(k-1)+(k-1)(rk-k-1)}\right]\left[\frac{100}{MSE}\right]$$

Where SS = Sum of squares, MSE= Mean squared Error, *r* is number of replications and *k* is the block size.

The R.E of an alpha lattice design compared with a RCBD is estimated [30] as:

$$\text{R.E} = \frac{Mean\ square\ error\ in\ RCBD}{Mean\ square\ error\ in\ Alpha\ lattice\ Design} \times 100$$

If the resulting value of R.E is greater than 1.0, the later design is more precise than the former. And if the resulting value of R.E is less than 1.0, the former design is more precise than the later.

3.2.5 ANOVA Model Diagnostic Tests

The interpretation of data based on analysis of variance models is valid only when the assumptions of the models are satisfied. As a result, it is necessary to detect any assumption deviations and apply the appropriate remedial measures.

3.2.5.1 Normality Assumption

The normality assumption implies that the distribution of the response variable there by the residual and to be analyzed by ANOVA is normal in the population from which units are sampled. Shapiro-Wilk test and Kolmogorov-Smirnov test are the formal tests of normality. Since the Kolmogorov-Smirnov test is appropriate for only large data (sample), Shapiro-Wilk test will be used in this study.

The null-hypothesis of the Shapiro-Wilk test is that the residuals are normally distributed, therefore p-values that are larger than 0.05 indicate that values are normally distributed at the 5% level of significance. If the test is significant, the assumption of normality is violated. In this case, transforming the data will frequently correct the problem. Among such transformations are logarithmic, square root, inverse square root and reciprocal transformations will be appropriate depending on the nature of the data set.

Also, the simplest check for normality involves plotting the empirical quantiles of the residuals against the expected quantiles. This is known as the normal QQ-plot. Thus, QQ-plots are useful for diagnosing violations of the normality assumption. In this method, observed value and expected value are plotted on a graph. If the scatter plots deviates from a straight line, then the data are not normally distributed. The normal, lognormal, exponential, and Weibull distributions can be used in the plot.

But, if the data normality could not be stabilized by the transformation technique, still there is one approach which is to consider the non-parametric statistical methods.

To test the assumption of normality, we have to look carefully at the error terms associated with each observation to determine whether they are randomly distributed or not.

3.2.5.2 Homoscedasticity Assumption

It is prudent to assess the equal variance assumption before conducting any ANOVA procedure.

This is because ANOVA assumes the variability of observations (measured as the standard deviation or variance) is the same in all populations.

There are several tests for heteroscedasticity. These include the F-ratio test (limited to testing the variances in two groups), Bartlett's test and Levene's test. The F-ratio test and Bartlett's test required the populations being compared to be Normal, or approximately so. However, unlike t tests and ANOVA, they are not robust when conditions of non-Normality and are not aided by

Central Limit Theorem. Levene's test is much less dependent on conditions of Normality in the population.

Bartlett (1937) introduced the homogeneity of variance test that involves comparing a statistic whose sampling distribution is closely approximated by the Chi-square distribution with k-1 degrees of freedom. The test is well established measure. However, it should be kept in mind that the test is a bit sensitive to non-normality, especially if the trials of the distribution are too long. When this occurs, the test tends to show significance too often.

The test criterion, when k<2 independent estimate of variance S_i^2, and all have the same number of degrees of freedom γ using the logarithm to base e is:

$X^2 = \frac{M}{C}$ where, M = $\gamma[k(\ln S^2) - \sum_i lnS_i^2]$, $S^2 = \frac{\sum_i S_i^2}{a}$ and C = $1 + \frac{a+1}{3a\gamma}$

But, when the samples are unequal size, the test statistic is similar to the above one except computing M and C as:

M = $(\sum \gamma_i)\ln S^{-2} - \sum \gamma_i \ln S_i^2$ and C = $1 + \frac{1}{3(k-1)}\left[\sum \frac{1}{\gamma_i} - \frac{1}{\sum \gamma_i}\right]$ where, $S^{-2} = \frac{\sum \gamma_i S_i^2}{\gamma_i}$

Hartley (1950) proposed another measure for testing the homogeneity of variance.

This test is based on the ratio of the largest to the smallest within group variance. This is known as the test of homogeneity of variance.

If the ratio is non-significant, variances are said to be homogeneous. On the other hand, if the ratio is significant, variances are said to be heterogeneous.

The F-test for the homogeneity of variance is defined as: F=$\frac{S_1^2}{S_2^2}$ Where, S_1^2 - is the larger sample variance and S_2^2 - is the smallest sample variance. The null and alternatives for Bartlett's test is:

H0: ${\sigma_1}^2 = {\sigma_2}^2 = ... = {\sigma_k}^2$

H1: at least one population variance differs from another

3.2.5.3 Additivity Assumption

Randomized block ANOVA models assume additive block and treatment effects. That is, there is no treatment by block interaction. Tukey's test for non additivity is a one degree of freedom test of the hypothesis that there is a linear treatment by linear block interaction. Our null hypothesis to be tested is the significance of the Additivity of the data. If the result of F-value is large, or the p-value is less than the level of significance 5%, we conclude that the test for Additivity of the data is statistically significant. As a result of this, the data are non-additive. Else, the test of non additivity will not be significant, that is, there is no indication of a treatment by block interaction. The assumption of additivity, which states that the block effects are approximately the same for all treatments, can be examined by using Tukey's test. The test requires that we compute the sum of squares for non-additivity in the following manner:

$SS_{\text{non-additivity}} = \frac{N^2}{(\sum d_i^2)(\sum d_j^2)}$ where, N: each cell of the table of raw data multiplied by the corresponding treatment and block effect and the sum of all the products N=$\sum w_i d_i \ \sum\sum X_{ij}^2 \ d_i d_j$ $\sum d_i^2$ = sum of the treatment effect squared; $\sum d_j^2$ = sum of the block effect squared

Equation of $SS_{\text{non-additivity}}$ is basically, the contribution of non-additivity with one degree of freedom to the error sum of squares. This value is then tested against the remainder of the residual sum of squares to determine whether the hypothesis of additivity is correct or not.

CHAPTER FOUR: RESULTS AND DISCUSSION

4.1 ANOVA Model Diagnostics Tests

Before doing ANOVA for any designs, there are assumptions in which the response variable should follow among these assumptions:

4.1.1 Normality Test

In this study, for each data set in all research sites, the normality of the data was tested using the Shapiro-Wilks test and by QQ-plots. We tested the null hypothesis which states that the yield of variety is normally distributed against the alternative that the yield of variety is not normally distributed at 5% level of significance. The QQ plots of **figure1** reveals that generally for the first dataset on soybean conducted using RCBD at the five sites in 2007, the dots in QQ plots are close to the line. Hence, it is reasonable to infer that the data at each location follow a normal distribution.

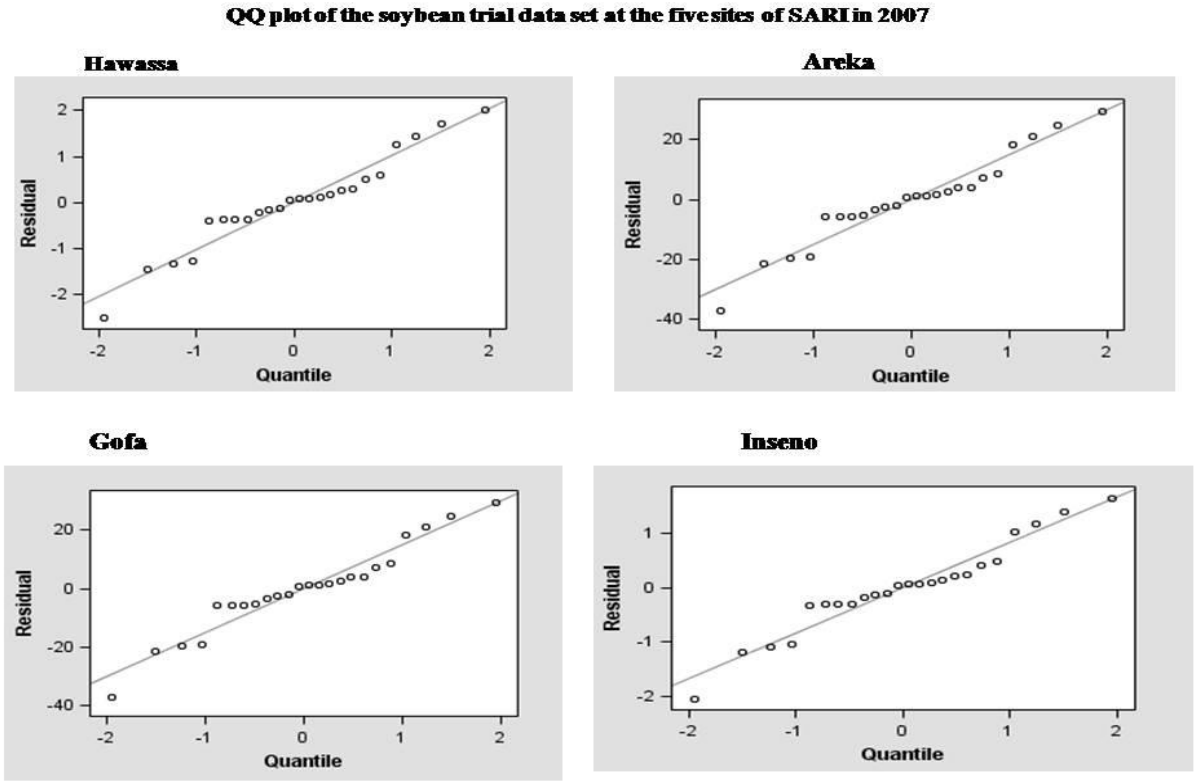
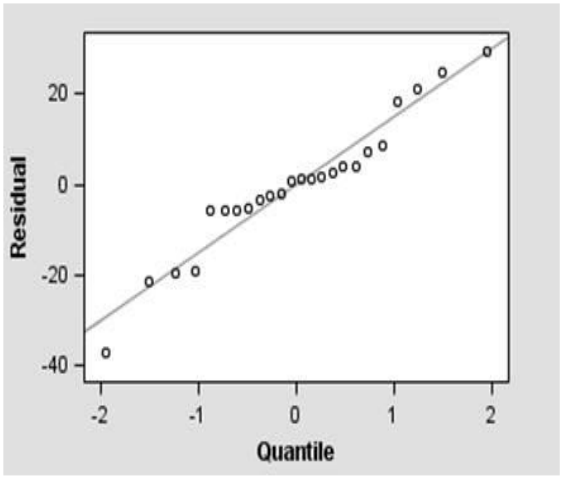


Figure 1: QQ plot of the soybean trial data set at the five sites of SARI in 2007

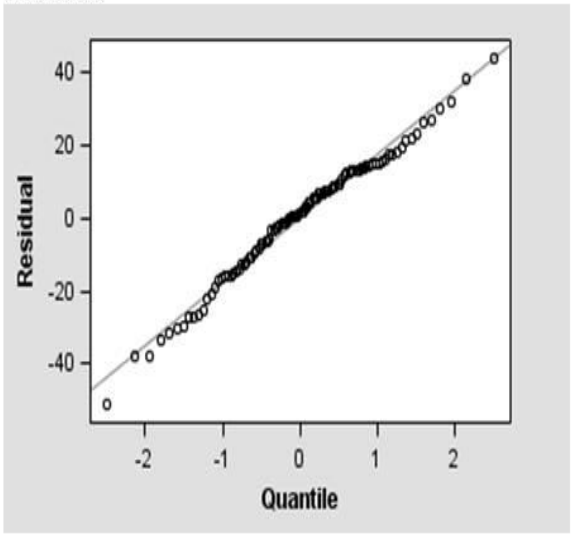
The QQ plot at **figure 2** shows that generally for the first dataset on maize variety trial dataset at five sites of SARI in 2008/9, the dots in QQ plots are close to the line. Hence, it is reasonable to infer that the data at each location follow a normal distribution.

Bonga

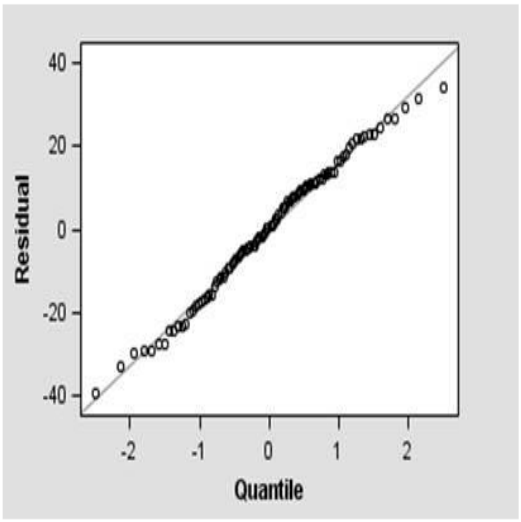


QQ plots for the Maize trial data set on the five sites of SARI in 2001/2

Hawassa



Areka



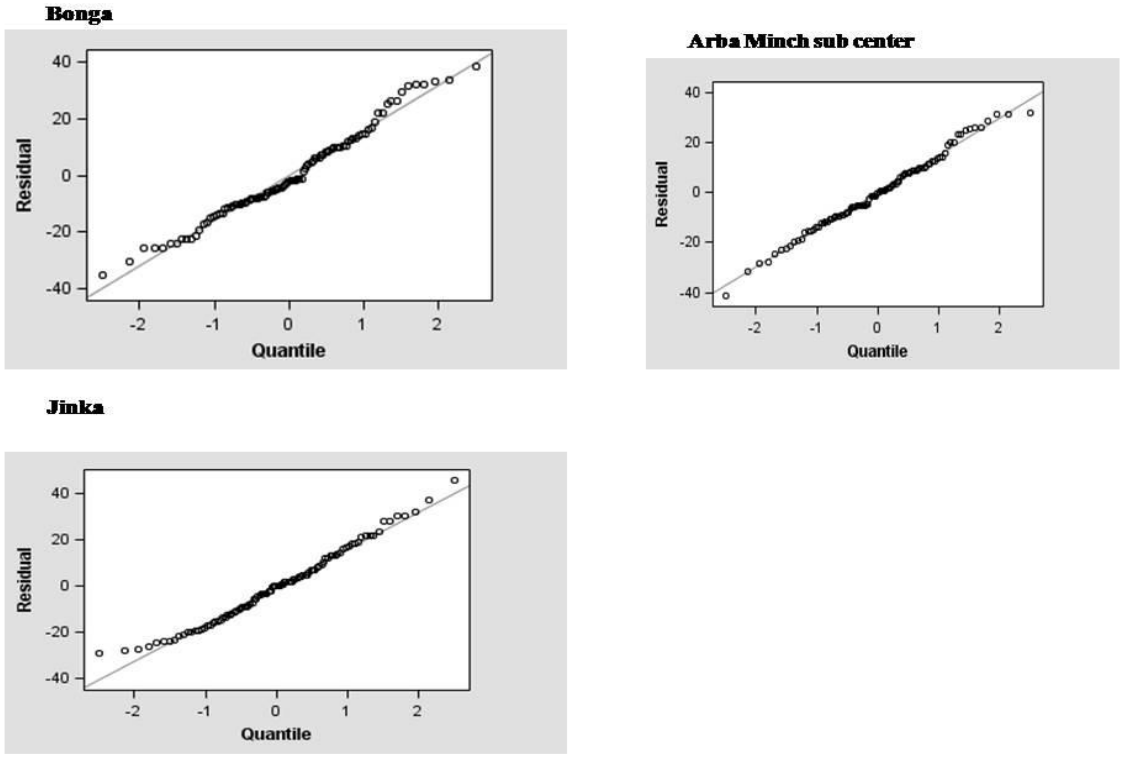


Figure 2: QQ plot of the Maize variety trial data set at the five sites of SARI in 2001/02

For the third dataset on maize variety trial conducted using alpha lattice design at Hawassa site of SARI in 2008/9, QQ plot at **figure 3** shows that the dots in QQ plots are close to the line. Hence, it is reasonable to infer that the data at this location follows a normal distribution.

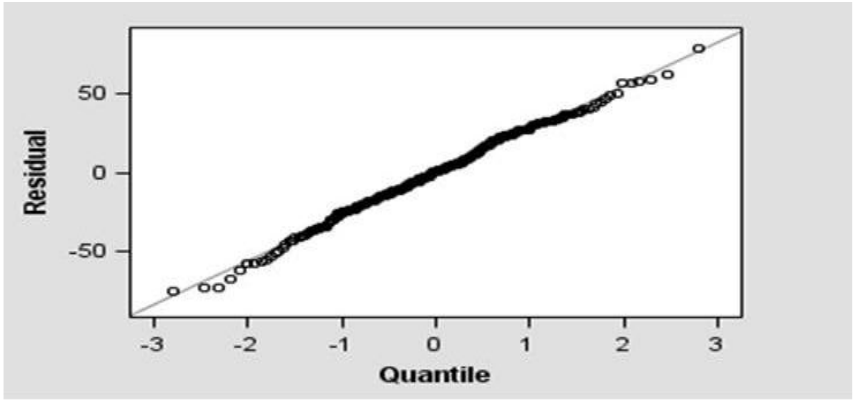


Figure 3: QQ plot of residuals of the Maize trial data set at the Hawassa site of SARI in 2008/9

The plot in figure 4 shows us that there is heterogeneity of variance across the five locations of SARI for the soybean variety trial dataset conducted using RCBD in 2007.

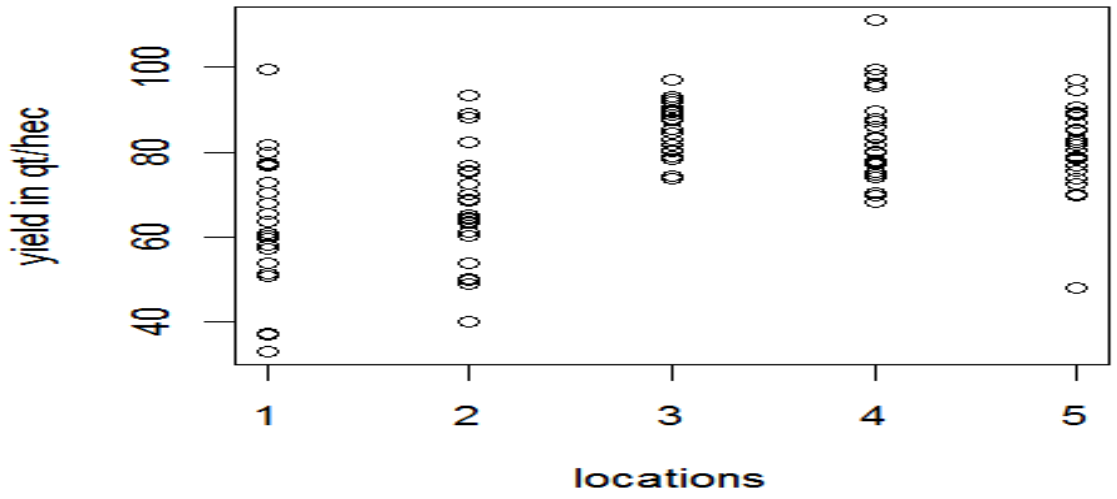


Figure 4: Plot of the soybean variety trial data across the five locations in 2007

From Table 6 for soybean variety data conducted using RCBD of all locations of 2007, the computed values of Shapiro-Wilks statistic (0.9667, 0.9487, 0.9587, 0.9430 and 0.9195) are not significant with their corresponding P-values (0.5888, 0.2542, 0.4131, 0.1903 and 0.0571) respectively. As a result, the normality assumption for Soybean variety trial data is satisfied. This shows that the soybean variety data in all the five locations was normally distributed.

Table 6: Normality test of the soybean variety trial data set in 2007

Research sites	Shapiro Normality test	
	W-statistic	P-value (Pr < W)
Hawassa	0.9667	0.5888
Areka	0.9487	0.2542
Gofa	0.9587	0.4131
Inseno	0.9430	0.1903
Bonga	0.9195	0.0571

From Table 7, we see that the computed values of the Shapiro-Wilk statistic (0.9762, 0.9897, 0.9921, 0.9827 and 0.9816) of Maize variety data conducted using Partially Balanced Lattice Design in all the five sites of SARI Hawassa, Areka, Bonga, Jinka and Arba Minch sub center are not significant with their corresponding P-values (0.0679, 0.6478, 0.8287, 0.2162 and

0.1775), respectively. Hence, the normality assumption of these data sets in all the five sites is satisfied.

Table 7: Normality test of the Maize variety trial data set 2008/9

Research sites	Shapiro Normality test	
	W-statistic	P-value (Pr < W)
Hawassa	0.976275	0.0679
Areka	0.989792	0.6478
Bonga	0.992104	0.8287
Jinka	0.982745	0.2162
Arba-Minch Sub center	0.981623	0.1775

Table 8 shows that the data set of the Maize variety trial conducted using Alpha lattice design at Hawassa site in 2008/9 the normality assumption is well satisfied having Shapiro-Wilk statistic and corresponding P-value (0.9861and 0.3104) respectively.

Table 8: Normality test of two Maize variety trial data sets at Hawassa site in 2008/9

Data set	Shapiro Normality test	
	W-statistic	P-value (Pr < W)
Maize data	0.9861	0.3104

4.1.2 Homoscedasticity Test

For the homogeneity of variance of each dataset in this study, Bartlett's test was used as the all the datasets satisfied normality assumption.

Looking for the results of the Bartlett's test at table 9, for Soybean trial dataset conducted using RCBD in 2007 the homogeneity of variance assumption at the five sites Hawassa, Areka, Gofa, Inseno and Bonga was satisfied having Bartlett's K-squared (3.8421, 0.9134, 1.638, 2.8627 and 5.3102) with corresponding P-values (0.1465, 0.6334, 0.4409, 0.239 and 0.07029) respectively.

Table 9: Homogeneity of variance test for Soybean trial dataset conducted using RCBD in 2007

Research Sites	Test for homogeneity of variance		
		Bartlett's K-squared	P value
Hawassa	Block	3.8421	0.1465
	Treatment	3.6754	0.8163
Areka	Block	0.9134	0.6334
	Treatment	4.1999	0.7565
Gofa	Block	1.638	0.4409
	Treatment	9.5311	0.2167
Inseno	Block	2.8627	0.239
	Treatment	4.9972	0.6603
Bonga	Block	5.3102	0.07029
	Treatment	5.1609	0.6403

From table 10, for the Maize variety trial dataset conducted using Lattice design in 2008/9 the homogeneity of variance assumption for the dataset of the five sites Hawassa, Areka, Bonga, Jinka and Arba Minch sub center is satisfied having Bartlett's K-squared (5.4493, 3.7744, 14.7715, 9.252 and 3.9258) and their corresponding P-values (0.2442, 0.2869, 0.05199, 0.0551 and 0.4161), respectively.

Table 10: Homogeneity of variance test for Maize variety trial dataset conducted using Lattice design in 2008/9

Research Sites	Test for homogeneity of variance		
		Bartlett's K-squared	P value
Hawassa	Block	5.4493	0.2442
	Treatment	32.5507	0.1139
Areka	Block	3.7744	0.2869
	Treatment	13.074	0.9649
Bonga	Block	14.7715	0.05199
	Treatment	23.9068	0.4669
Jinka	Block	9.252	0.0551
	Treatment	26.5343	0.3266
Arba Minch Sub Center	Block	3.9258	0.4161
	Treatment	29.3742	0.2064

For the third dataset of Maize variety trial conducted using Alpha Lattice design in 2008/9 at Hawassa site of SARI, the homogeneity of variance assumption is satisfied having Bartlett's K-squared 85.4416 and its corresponding P-value (0.8722) respectively as of the table 11.

Table 11: Homogeneity of variance test for Maize variety trial dataset conducted using Alpha Lattice design in 2008/9

Datasets	Test for homogeneity of variance		
		Bartlett's K-squared	P value
Maize data	Block	85.4416	0.8722
	Treatment	64.3705	0.8985

4.1.3 Additivity test

From the analysis of Tukey's additivity test for the datasets discussed so far, looking for the corresponding one df Tukey's P-value, we will infer whether there is an interaction between blocks and treatments. For the soybean variety trial data set at the five sites Hawassa, Areka, Gofa, Inseno and Bonga the additivity assumption is well satisfied having the corresponding P-value (0.07345,0.3292, 0.9686, 0.8181 and 0.2808) respectively (Table 12).

Table 12 Additivity test for Soybean trial dataset conducted using RCBD in 2007

Research Sites	Test for Additivity	
	F value	P value
Hawassa	3.7914	0.0734
Areka	1.0279	0.3292
Gofa	0.0016	0.9686
Inseno	0.0551	0.8181
Bonga	1.2666	0.2808

For this dataset, there is no evidence of interaction between blocks and treatments, so this model assumption is satisfied.

For the Maize variety trial dataset conducted using lattice design at five sites Hawassa, Areka, Bonga, Jinka and Arba Minch sub center the additivity assumption is well satisfied having the corresponding P-value (0.3613, 0.0731, 0.3042, 0.7076 and 0.06355) respectively as shown in the table 13. Again for this data, there is no evidence of interaction between blocks and treatments.

Table 13: Additivity test for Maize variety trial dataset conducted using Lattice design in 2008/9

Research Sites	Test for Additivity	
	F-value	P value
Hawassa	0.8441	0.3613
Areka	5.9395	0.0731
Bonga	1.0712	0.3042
Jinka	0.1418	0.7076
Arba Minch Sub Center	6.4125	0.06355

Looking table 14, for the last dataset on Maize variety trial at Hawassa site, the Additivity assumption is well satisfied having P-value (0.09541).

Table 14: Additivity test for Maize variety trial dataset conducted using Alpha Lattice design in 2008/9

Datasets	Test for Additivity	
	F-value	P value
Maize data	6.8847	0.09541

4.1.4 Combined analysis for CRD and RCBD analysis

4.1.4.1 Combined analysis of RCBD for soybean trial dataset at five locations

Hypothesis test for the variance homogeneity at five locations:

H0: $\sigma_1^2=\sigma_2^2 = \cdots = \sigma_5^2$ V_s H_1: at least one of the variance is different

Table 15: The MSEs of RCBD of the soybean variety trial data in five locations

Location	S^2	$\ln(S^2)$
1	196.524	5.280
2	159.271	5.070
3	25.562	3.241
4	116.074	4.754
5	96.749	4.572

Using the Ratio test:

A quick test of homogeneity of variance is provided by the ratio of the largest to the smallest S^2 in the set as:

$\frac{S^2_{max}}{S^2_{min}} = {196.524}/{25.562} = 7.688$, the upper 5% point for a=k=5, $\gamma = 23$ is 2.6399

The critical value is less than the test statistic; we reject the null hypothesis and conclude that there is heterogeneity of variance across the five locations.

Using the Bartlett's test:

Here, $\gamma = 23$, $a = 5$, $S^2 = \frac{\sum_i S_i^2}{a} = 118.836$, $\sum_i S_i^2 = 594.18$, $\sum_i lnS_i^2 = 22.9188$

$M = 23[5(4.777) - 22.9188] = 22.9188$, $C = 1+\frac{a+1}{3a\gamma} = 1.0174$ and $X^2 = \frac{M}{C} = 21.8399$

$X^2_{0.05}(5-1) = 9.49$

As we see the tabulated Chi-square value and comparing to the Statistic ratio X^2 , the test statistic is significantly greater than the critical value. Consequently, we reject the null hypothesis and conclude that there is heterogeneity of variance across locations.

So, for RCBD of the soybean variety trial dataset, as we observed so far, the results using the two approaches reveal us there is heterogeneity of variance across the locations.

4.1.4.2 Combined analysis of CRD for soybean trial dataset at five locations

Hypothesis test for the variance homogeneity at five locations:

H0: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_5^2$ V_s H_1: at least one of the variance is different

Table 16: The MSEs of CRD of the soybean variety trial data in five locations

Location	S^2	$\ln(S^2)$
1	258.621	5.555
2	165.526	5.109
3	39.111	3.666
4	112.613	4.723
5	108.961	4.690

Using the Ratio test:

$\frac{S^2_{max}}{S^2_{min}} = {258.621}/{39.111} = 6.612$, the upper 5% point for a=k=5, $\gamma = 23$ is 2.639

The critical value is less than the test statistic; we reject the null hypothesis and conclude that there is heterogeneity of variance across the locations.

Using the Bartlett's test

Here, $\gamma = 23$, $a = 5$, $S^2 = \frac{\sum_i S_i^2}{a} = 136.9664$ $\sum_i S_i^2 = 684.832$, $\sum_i \ln S_i^2 = 23.7458$

M = $23[5(4.9197) - 23.74584]$= 19.6162 and C = $1+\frac{a+1}{3a\gamma}$ = 1.0174,

Hence, $X^2 = \frac{M}{C}$ = 19.2807; $X^2_{0.05}$(5-1) = 9.49

Looking for the tabulated Chi-square value and comparing with ratio X^2, as the test statistic is greater than the critical value, we reject the null hypothesis and conclude that there is heterogeneity of variance across locations.

Since there is heterogeneity of variance across the five locations, we will be forced to do the analysis independently for each location of the two designs CRD and RCBD.

4.2 Randomized Complete Block Design (RCBD)

From Table 17, for the soybean variety trial data set conducted using RCBD at the five different sites of SARI, the block effects are significant for the four sites namely Hawassa, Areka, Gofa and Bonga having corresponding p-values (0.0377, 0.0270, 0.0491 and 0.0494) respectively. But, for the Inseno site there is not significant block effect with P-value (0.7248).

Table 17: ANOVA table for soybean variety trial data set conducted using RCBD in 2007

Location	Source	Degree of freedom	Mean square	F value	P >F
Hawassa	Block	2	786.4432	4.002	0.0377
	Treatment	23	4.8758	0.0247	0.8767
Areka	Block	2	895.1737	4.555	0.027
	Treatment	23	7.21540	0.05	0.8345
Gofa	Block	2	497.9182	2.53	0.0491
	Treatment	23	4.5989	0.18	0.6779
Inseno	Block	2	14.9769	0.13	0.7248
	Treatment	23	127.9942	1.10	0.3115
Bonga	Block	2	647.6100	3.296	0.0494
	Treatment	23	554.7767	5.73	0.0312

Table 18 shows that the relative efficiency of RCBD compared to CRD for the soybean data set in sites Hawassa, Areka, Gofa, Inseno and Bonga were 1.311597, 1.039272, 1.530044, 0.970182 and 1.126223 respectively.

Table 18: Summary for CRD and RCBD analysis of Soybean variety trial data in 2007

Sites	No of plots	No of varieties	No of blocks/ Replication	Mean square error		CV		Relative Efficiency
				CRD	RCBD	CRD	RCBD	
Hawassa	24	8	3	258.621	196.524	25.922	22.597	1.3115
Areka	24	8	3	165.526	159.271	19.138	18.772	1.0392
Gofa	24	8	3	39.111	25.562	7.2739	5.880	1.5300
Inseno	24	8	3	112.613	116.074	12.719	12.913	0.9701
Bonga	24	8	3	108.961	96.749	12.986	12.236	1.1262

This indicates that the use of RCBD for the sites Hawassa, Areka, Gofa and Bonga of soybean variety trial instead of a CRD increased experimental precision by 31, 3, 53, and 13 percent respectively. The relative efficiency of the RCBD compared to CRD for Inseno site is nearly one. This indicates that the efficiency of RCBD and CRD for this site is almost the same. Thus, blocking seems insignificant and unnecessary rather requiring extra cost.

For the sites Hawassa, Areka, Gofa and Bonga the MSE under RCBD (196.524, 159.271, 25.562 and 96.749) was smaller as compared to MSE of CRD (258.621, 165.526, 39.111 and 108.961) respectively. And moreover, it can also be noted that the CV of RCBD (22.597, 18.772, 5.880 and 12.236) was low as compared to CV of CRD (25.922, 19.138, 7.2739 and 12.986) respectively. But, for the Inseno site, there is slight increase in MSE and as well as at the CV in the RCBD which tells us there is no block effect in increasing the precision of design.

4.3 Partially Balanced Lattice Design

Table 19 shows the results of the ANOVA of RCBD and Lattice design with their corresponding Mean square error and Coefficient of variation for the maize variety trial data set in 2008/9 at five sites of SARI.

Table 19: Summary table for RCBD and Partially Balanced Lattice design analysis of Maize variety trials data in 2008/9

Sites	No of plots	No of Varieties	No of blocks Replication	Mean square error		CV		R.E
				RCBD	Lattice	RCBD	Lattice	
Hawassa	100	25	4	352.18	350.64	28.2200	26.4	1.0043
Areka	100	25	4	266.6	251.0	25.0006	20.9	1.0621
Bonga	100	25	4	165.47	143.79	21.8115	15.7	1.1507
Jinka	100	25	4	262.89	262.44	26.2911	21.7	1.0017
Arba Minch sub center	100	25	4	287.74	260.91	20.5045	18.9	1.1028

For the Maize variety trial data set, ANOVA for RCBD and Lattice design was performed. From the results of the two analysis, at the five sites of SARI in 2008/9 (Hawassa, Areka, Bonga, Jinka and Arba Minch Sub center), the MSE under lattice design (350.64, 251.0, 143.79, 262.44 and 260.91) was smaller as compared to MSE of RCBD (352.18, 266.6, 165.47, 262.89 and 287.74) respectively. Moreover, it is noted that the CV of Lattice design (26.4, 20.9, 15.7, 21.7 and 18.9) was low as compared to CV of RCBD (28.2200, 25.0006, 21.8114, 26.2911 and 20.5045) for all the five sites mentioned above. The relative efficiency of the RCBD relative to Lattice design is 1.0043, 1.0621, 1.1507, 1.0017 and 1.1028 for Hawassa, Areka, Bonga, Jinka and Arba Minch sub center respectively. Hence, the use of Lattice for the sites Hawassa, Areka, Bonga, Jinka and Arba Minch sub center of Maize variety trial data in 2008/9 instead of RCBD increased experimental precision by 0.44, 6.2, 15.07, 0.17 and 10.31 percent, respectively (Table 19).

4.4 Alpha Lattice Design

The significance of blocking within replication (group) in both designs for these data set indicates that blocking was effective in reducing experimental error and furthermore, increasing precision of design (Table 20, 21).

Table 20: ANOVA for RCBD of the Maize variety trial data at Hawassa site in 2008/9

Data set	Source	Df	SS	MS	F value	P>F	C.V
Maize data	Block	2	4595.7	2297.85	15.69	5.767e-11 ***	22.914
	Treatment	80	1467.1	18.334	0.1252	0.7047	
	Residuals	160	23429.2	146.43			

Table 21: ANOVA for Alpha lattice design of the Maize variety trial data at Hawassa site in 2008/9

Data set	Source of variation	Df	SS	MS	F value	P>F	C.V
Maize data	Replication	2	4388.8	2194.42	26.7225	1.629e-10 **	21.1 %
	trt.unadj	80	7345.0	91.81	1.1180	0.2814	
	replication:block.adj	24	5137.0	214.04	2.6065	0.0002 **	
	Residuals	136	11168.2	82.12			

Table 22 shows that the relative efficiency for Maize dataset was 118.851% implying that the use of alpha lattice design increased experimental precision by 18.851% compared to RCBD. The CV(21.1%) and MSE(82.12%) of alpha lattice design is low as compared to RCBD having CV=22.9136 and MSE = 97.6) respectively.

Table 22: Summary table for RCBD and alpha lattice design analysis of Maize variety trial data set at Hawassa site in 2008/9

Data set	No of plots	No of entries	No of blocks/ replication	Mean square error		CV		R.E
				RCBD	Alpha Lattice	RCBD	Alpha Lattice	
Maize data	243	81	3	97.6	82.12	22.913%	21.1%	1.1885

4.5 RCBD with Missing Values

Table 23 shows that the analysis of RCBD for the soybean variety trial data at Areka site with two missing values. This has been done in two approaches first using approximate analysis and replacing the estimated value then performing the usual ANOVA. The second approach is applying the concept of IBD. The MSE of RCBD with two missing values for missing estimate approach (197.044) is greater than the MSE of RCBD using IBD approach (176.641). Furthermore, the CV of RCBD with missing estimate approach (21.235%) is greater than the CV of RCBD using IBD approach (15.762%).

Table 23: ANOVA the soybean variety trial for the Areka data set with two missing values

Analysis of variance with estimated missing value						
Source	Df	SS	MS	F value	P>F	CV
Block	2	798.2745	114.0392	0.58	0.7609	21.235
Treatment	7	445.8454	222.9227	1.13	0.3547	
Residual	12	2364.5294	197.0441			
Analysis of variance for Incomplete Block						
Block	2	590.16	295.08	1.6705		15.763
Treatment	7	1236.9	176.70	1.0003	0.3799	
Residual	12	2119.7	176.641			

Table 24 shows the results of the second analysis for RCBD of the soybean variety trial data at Bonga site with three missing values. This again has been done using the two approaches first using approximate analysis and replacing the estimated value then performing the usual Two-way ANOVA. The second approach is applying the concept of IBD for RCBD. Using the two approaches, for the IBD approach, the block effect is significant with p-values (0.0346). The MSE of RCBD using missing estimate approach (113.074) is greater than to the MSE of RCBD using IBD (91.68). Additionally, the CV of RCBD with estimated missing values approach (15.51659%) is greater than the CV of RCBD using IBD approach (13.0228%).

Table 24: ANOVA the soybean variety trial for the Bonga site with three missing values

Analysis of variance with estimated missing value						
Source	Df	SS	MS	F value	P>F	CV
Block	2	717.3	358.65	3.1718	0.0413 *	15.5166
Treatment	7	2247.0	321.00	2.8388	0.0195 *	
Residual	11	1243.8	113.074			
Analysis of variance for Incomplete Block						
Block	2	796.71	398.36	4.3451	0.0346 *	13.0228
Treatment	7	1672.9	238.98	2.6067	0.0754	
Residual	11	1008.5	91.68			

4.6 DISCUSSION

For the analysis of RCBD on soybean variety trial datasets, in the year 2007 the block effect was significant for four sites Hawassa, Areka, Gofa and Bonga. But, for the Inseno site, the blocking effect was not significant due to many reasons. Moreover, the relative efficiency of RCBD over CRD for these sites, Hawassa, Areka, Gofa and Bonga was greater than one. The use of the RCBD for the sites Hawassa, Areka, Gofa and Bonga of soybean variety trial data instead of a CRD increased experimental precision by 31, 3, 53, and 13 percent respectively. But for the Inseno site, the relative efficiency of RCBD relative to CRD is approximately one. This means among these four sites, we would have required approximately about two times as many replicates with CRD to obtain the same efficiency as is obtained by RCBD for Hawassa, Gofa and Bonga. Therefore, RCBD is more efficient than CRD for the soybean variety trial datasets at the four sites, Hawassa, Areka, Gofa and Bonga. But, for the Inseno site, because of many factors, blocking doesn't seem to be appropriate. That is, experimental units seems to be homogeneous than the expected heterogeneous. Looking for the MSE and CV, for the sites Hawassa, Areka, Gofa and Bonga the MSE under RCBD was smaller as compared to CRD and it can also be noted that the CV of RCBD was low as compared to CRD.

But, for the Inseno site, there was slight increase in MSE in the RCBD as well as in CV, there was slight increase. This tells us there is no blocking effect in increasing the precision for this site. This may be due to technical problems like wrong block orientation and direction which render blocking ineffective (Girma, 2005).

It was also tried to see analysis of RCBD when there are missing values. Using the two approaches for the analysis of RCBD, the IBD approach revealed to be better approach with less MSE and lower CV as compared to the missing estimate approach. As a result using IBD is more efficient than the estimated missing approach of RCBD for the SARI field condition.

Results from the analysis of lattice design indicates that the CV and the MSE of Maize variety trial was calculated for Hawassa, Areka, Bonga, Jinka, and Arba Minch research centers and CV and MSE of RCBD were greater than that of lattice design. The efficiency of lattice design relative to RCBD shows that in those five sites considered, lattice design was more efficient than RCBD. The significance of blocking within replication (group) in both designs for the maize trial

dataset indicates that blocking was effective in reducing experimental error and furthermore, increasing precision of design.

For the Maize variety trial data at Hawassa site in 2008/9, the MSE for alpha lattice design was smaller as compared to RCBD. Moreover, it can be noted that the CV of Alpha Lattice design was lower as compared to RCBD.

The relative efficiency of Alpha Lattice design relative to RCBD is 1.1885 for the Maize data at Hawassa site. This implies that the use of Alpha Lattice for the Maize variety trial data set at Hawassa site in 2008/9 instead of RCBD increased experimental precision by 18.8504 percent.

From the experiment in the Maize data set (Table 22) smaller MSE was observed for alpha lattice design than RCBD. The CV of alpha lattice design is low as compared to RCBD. Lower CV for the alpha lattice design indicates a good index of reliability. The relative efficiency of greater than one indicates how much efficient alpha lattice design is as compared to RCBD. Because of the value of relative efficiency is greater than one, alpha lattice design results in smaller error variance. Alpha lattice design provided better control of within replicate variation and also better control to experimental error than RCBD. This study showed that alpha lattice design is more efficient than RCBD. This was similar with the findings of Patterson and Hunter (1983), Yau (1997), and Masood et al. (2008).

CHAPTER FIVE: CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

The primary objective of this study was to compare and identify the more efficient experimental design for agricultural trials under SARI field condition. Accordingly, we came up with the following conclusions:

For the first data set on soybean variety trial at the five sites of SARI in 2007, on average RCBD was found to be more efficient than CRD for four of the sites. Perhaps, due to technical problem on the proper orientation of blocks or any other potential reasons, blocking effect was low or insignificant at the Inseno site. It is therefore important to consider RCBD in the SARI trials. However, block orientation and direction should be carefully made according to the gradient of the field experiment.

For the Maize variety trial, lattice design was found to be more efficient than RCBD in increasing the precision of the field experiments under the SARI condition.

Also, for another Maize variety trail dataset, alpha lattice design was found being more efficient design than RCBD improving the precision of field experiments. Under heterogeneity of experimental units and in a situation where there are large number of varieties to be tested, the chance that family of incomplete block designs being used is quite high.

5.2RECOMMENDATIONS

Based on the results of the analysis of CRD, RCBD, Lattice and Alpha-lattice designs under the field conditions of SARI, the following possible recommendations are given:

- To control variability in field experiments, it is suggested that an experiment with an RCBD could be replaced with family of incomplete designs like lattice or alpha lattice design whenever the number of treatments in the experiment is quite large. This is mainly because, in such situation finding a homogeneous block is quite difficult and sometimes it is impossible.
- For a large number of missing values in RCBD, we recommend that agricultural researchers should opt for incomplete block design analysis approach instead of using missing-estimated approach.
- For some locations used in this study, may be blocking is not necessary. This can be due to technical issues such as orientation and direction of blocking. So, we recommend that further studies should be done in this area so that appropriate design will be recommended for future use.

REFERENCES

1. Alves E. G., St. Aubyn, A. and Martins, A. (2009): Experimental designs for evaluation of genetic variability and selection of ancient grapevine varieties: a simulation study. Heredity Journals, **104**(6):552-562.

2. Atkinson A.C. and Bailey, R.A. (2001): One Hundred Years of the Design of Experiments on and off the Pages of "Biometrika". Biometrica, **88**(1): 53-97.

3. Bartley, M.S. 1936. Square-root transformation in analysis of variance. J.R. Stat. Soc. Suppl. **3**:68-78

4. Blau G. and Han, S. (2007): Statistical Design of Experiments. http://www.Pharmahub.org/

5. Campbell B.T. and Bauer, P.J. (2007): Improving the precision of cotton performance trials conducted on highly variable soils of the South-Eastern USA coastal plain. Blackwell synergy-plant breeding, **126**, 622-627.

6. Cochran W.G. and Cox, G.M. (1957): Experimental Designs. John Wiley & Sons, Inc. New York.

7. Cox V.G. (1950): Experimental Design. Biometrics **6**, 301-302.

8. Cullis B.R. and Gleeson A.C. (1991): Spatial Analysis of Field Experiments. An Extension to Two Dimensions. Biometrics **47**, 1449-1460.

9. Das M.N. and Giri N.C. (1986). Design and analysis of experiments. New Age international Limited.

10. Day B.B. and Austin L. (1939): A Three-dimensional Lattice Design for Studies in Forest Genetics. J. Agr. Res., **59**, 101-119.

11. Drane J.W. (1989): Compromises and Statistical Designs for Grazing Experiments in Grazing Research. Crop Science Society of America and American Society of Agronomy: **16**, 69-83.

12. Edme S.J., Tai and J.D. Miller (2007): Relative Efficiency of Spatial Analysis for Non-replicated Early-stage Sugarcane Field Experiments. Journal of the American Society of Sugar Cane Technologists, **27**: 89-104.

13. Eskridge K.M. (2008): Field Trial Designs in Plant Breeding. University of Nebraska, Lincoln, NE.

14. Gharde Y. (2000): Analysis of Covariance. I.A.S.R.I., Library Avenue, New Delhi.

15. Girma T. A. (2002): Design and Analysis of Field Experiments in Agriculture. Ethiopian Agriculture Research Organization. Technical Manual No. 15, Addis Ababa.

16. Girma T. A. (2005): Using Spatial Modeling Techniques to Improve Data Analysis from Agricultural Field Trial. PhD Thesis of University of KwaZulu-Natal, Pietermaritzburg.

17. Gomez K.A. and Gomez, A.A. (1984): Statistical Procedures for Agricultural Research, 2nd ed. John Wiley and Sons, Inc.; New York

18. Gunjaca J., Renka H., Jambresic I. and Sindrak Z. (2005): Efficiency of Alpha Designs in Croatian Variety Trials. 27th Int. Conf. Information Technology Interfaces ITI, Cavtat, Croatia.

19. Gwowen Shieh, Show-Li Jan (2004): The effectiveness of randomized complete block design. Statistica Neerlandica: **58**(1): 111–124.

20. Hartley, H.O 1950. The Maxi mum F-ratio as a short-cut test for heterogeneity of variance. Biometrika: **37**:308-312

21. Hatfield J.L. (2000): Precision Agriculture and Environmental Quality: Challenges for Research and Education. U.S. Department of Agriculture, in Ames, Iowa.

22. Idrees N. and Khan M.I. (2009): Design improvement Using Uniformity Trials Experimental data. Pakistan Journal Agricultural Science. **46**, Faisalabad

23. Johansson H. (1963): Arrangement and Design of Field Experiments in Progeny Testing. World Consultation Forest. Genet. Tree Imp, FAO, Stockholm.

24. Karcher D.E., Richardson M.D., and Secks M.E (2003): Does Blocking Affect Experimental Efficiency on Sand-Based Putting Greens? Crop Science Society of America, **43**: 2295-2297.

25. Kempton R.A., Seraphin J.C. and Sword A.M. (1994): Statistical Analysis of Two Dimensional Variations in Variety Yield Trials. Journal of Agricultural Science, Cambridge, **122**, 335-342.

26. Kuehl R.O (1994): Statistical Principles of Research Design and Analysis. Duxbury Press, Belmont, CA.

27. Leonardo Journal of sciences (2009), Issue 15 p. 71-8: Statistical Approaches in Analysis of Variance, Romania.

28. Little T. M. and Hills, F. J. (1978): Agricultural Experimentation, Design and Analysis. John Wiley and Sons, Inc. New York.

29. Mandefro N. (2005): Statistical procedures for Designed Experiments. Ethiopian Agriculture Research Organization, Addis Ababa.

30. Masoon A. A., Mujahid Y., Khan M.I., and Abid S. (2006): Improving precisions of Agricultural field Experiments. Journal of Sustainable Development, **3**: 11-13.

31. Masood M. A., Farooq K., Mujahid Y. and Anwar M. Z. (2008): Improvement in Precision of Agricultural Field Experiments through Design and Analysis. Pak. J. Life Soc. Sci., **6**(2): 89-91.

32. Montgomery D. C. (1991): Design and Analysis of Experiments. 3rd ed. John Wiley and Sons, Inc. New York.

33. M. Kashifa M. I. , Khanb M. , Arifb M. Anwerc and M. Ijazc(2001): Journal of Scientific Research: **3** (1), 91-95: www.banglajol.info/index.php/JSR

34. Neter J., W. Wasserman and M. H. Kutner. (1990): Applied linear statistical models: regressions, analysis of variance, and experimental designs, 3rd ed. Irwin, Burr Ridge, Illinois.

35. Panse V.G. (1958): Status of Agricultural Experiments. F.A.O., Rome.

36. Patterson H.D. and Hunter E.A. (1983): The efficiency of incomplete block designs in National List and Recommended cereal variety trials. J. Agric. Sci., **101**, 427-433.

37. Patterson H.D. and Williams E.R. (1976): A New Class of Resolvable Incomplete Block Designs. Biometrika, **63**:83-92.

38. Pearce S.C. (1983): The Agricultural Field Experiments. A Statistical Examination of Theory and Practice. John Wiley and Sons, Inc., New York.

39. Raza I. and Masood M.A (2009): Efficiency of Lattice Design in Relation to Randomized Complete Block Design in Agricultural Field Experiments. Pakistan J. Agric. Res. **22**: 150-153.

40. Reza Hoshmand, A. (2006): Design of Experiments for Agriculture and the Natural Science, 2nd Edition: Daniel Webster College Nashua, New Hampshire, U.S.A.

41. Sharma V.K. (2000): Balanced Incomplete Block Designs. I.A.S.R.I., Library Avenue, New Delhi-110012.

42. Snedecor, G. and Cochran, W.G. 1980. Statistical Methods. 7th ed. Ames, IA: Iowa State University press.

43. Snyder E.B. (1962): Lattice and Compact Family Block Designs in Forest Genetics.U.S. Department of Agriculture, Gulfport, Mississippi.

44. Vas Es C.P. G. and Sellman M. (2007): Spatially-Balanced Complete Block Designs for field experiments. Cornell University, Ithaca, United States. Geoderma **140**, 346-352.

45. Wu T. and Dutilleul P. (1999): Validity and Efficiency of Neighbor Analysis in Comparison with Classical Complete and Incomplete Block Analysis of Field Experiments. Agron. J. **91**: 721-731.

46. Yates F. (1936): A New Method of Arranging Variety Trials Involving a Large Number of varieties. J. Agric. Sci. **26**, 424-455.

47. Yates F. (1939): The Recovery of Inter-block Information in Variety Trials Arranged in Three-dimensional Lattices. Ann. Eugen. **9**, 136-156.

48. Yau S.K. (1997): Efficiency of alpha-lattice designs in international variety yield trials of barley and wheat. The Journal of Agricultural Science, **128**, 5-9.