PREDICTION MODELS FOR THE PERCENTAGE EXTRACT OF OXALIC ACID FROM THE BARK OF EUCALYPTUS CAMALDULENSIS

BY

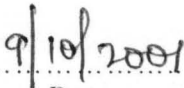## Isah Abubakar Garba

M.ENG/SEET/99/2000/415

A PROJECT SUBMITTED TO THE DEPARTMENT OF CHEMICAL
ENGINEERING, FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA,
IN PARTIAL FULFILLMENT FOR THE AWARD OF MASTERS DEGREE
(M.ENG) IN CHEMICAL ENGINEERING

JULY, 2001

DECLARATION

I, *ISAH ABUBAKAR GARBA (M.ENG/SEET/99/2000/415)* declare that this thesis titled Prediction model for the percentage extract of oxalic acid from the bark of eucalyptus camaldulensis, presented for award of Masters of Engineering in the Department of Chemical Engineering has not been presented for any other degree elsewhere.
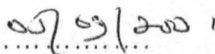
....
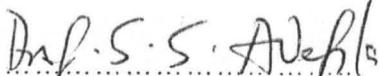*Signature*

.........9|10|2001
Date

CERTIFICATION

This is to certify that this project work was supervised and approved by the following persons on behalf of Chemical Engineering Department ,Federal University of Technology Minna.

....
DR. K.R. ONIFADE
(Project Supervisor)
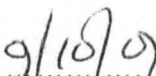
....
Date

....
DR. J.O. ODIGURE
(Head of department)

....
Date

....
External Examiner

....
Date

DEDICATION

In the name of ALLAH the most gracious the most merciful. May the peace and blessings of ALLAH be on His noble Prophet MUHAMMAD (S.A.W.), the Opener and the Seal.

This work is dedicated to our beloved Prophet *MUHAMMAD (S.A.W.), His KHALIFA, SHEIKH TIJJANI (R.A.) and Their KHALIFA, SHEIKH IBRAHIM INYAAS (R.A.).*

ACKNOWLEDGEMENT

v

ABSTRACT

The main aim of this research work was to use 2^k factorial analysis to predict the percentage extract of oxalic acid obtained from the bark of eucalyptus camaldulensis as a function of three operating parameters. Three 2^3 models were used and only two satisfy the G-test. The obtained equations are:

First Model:

Version 1: $y = 2.022 - 0.434x_1 - 0.411x_2 - 0.423x_3 - 0.342x1x_2 + 0.338x_1x_3$

Version: 2: $y = 2.022 + 0.434x_1 + 0.411x_2 + 0.423x_3 - 0.342x_1x_2 + 0.338x_1x_3$

Where x_1= mass, x_2= size and x_3= time.

Second Model:

$$y = 1.327 + 0.260x_1 - 0.438x_3 - 0.314x_1x_2$$

Where x_1= concentration, x_2= time and x_3= size.

It was found out that for the first model mass, time and size are inversely proportional to the percentage extract. Nevertheless, for the second model, concentration is directly proportional to the percentage extract while time and size are inversely proportional to the percentage extract oxalic acid obtained.

The two models that satisfy the G-test were adequate but the first model was more accurate than the second. During the course of the work, a computer program that can implement the model and prediction calculation was developed. The program can then be used to get a model and predict the percentage extract for three operating variables, which falls within the range of the data used for the 2^k designs. The program is an interactive Pascal program that can be used for 2^2, 2^3 and 2^4 factorial analyses.

TABLE OF CONTENTS

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

LIST OF TABLES

LIST OF FIGURES

NOTATIONS AND ABBREVIATIONS

α -   Level of significance

r -   No of replicates

N -   No of combinations i.e. 2^k

k -   No of factors

λ -   Insignificant coefficients

$w_1, w_2, \ldots w_k$ -   Natural factors

xj -   Coded factors

bj -   Coefficients of regression factors

$/bj/$ -   Absolute value of coefficient of regression factors.

Ri -   Response of ith combination

yi -   Percentage (%) extract of the ith combination

Su -   Experimental error

Su^2 -   Mean squared error

yr_1 -   Percentage (%) extract for replicate 1

yr_2 -   Percentage (%) extract for replicate 2

CHAPTER ONE

1.0   INTRODUCTION

The production of useful raw materials from the parts of trees like the bark, stem, twig, culm, leaves etc. is as old as humanity itself. Even the ancient man extracted raw materials like herbal-medicine and fibres from parts of trees. Nowadays most of the raw materials for manufacturing drugs are extracted from parts of trees. With recent developments more raw materials are found to be present in the bark and other parts of trees.

Eucalyptus is a large genus of trees and shrub native of Australia, New Zealand, Tasmania and Malaysia (Anonymous,1994). There are 650 – 700 species of eucalyptus, of which eucalyptus camaldulensis is among. Eucalyptus camaldulensis was introduced to Nigeria in 1916 and since then it has been cultivated in large number in Nigeria especially in savannah region of the country. Areas of Nigeria where eucalyptus is cultivated include the Jos Plateau, Kaduna, Kano and Niger State.

Oxalic acid is a colourless, crystalline, toxic organic compound belonging to the family of carboxylic acids. Oxalic acid is widely used as acid rinse in laundries, where it is effective in removing rust and ink stains because it converts most of insoluble ion compound into soluble complex ion. For the same reason, it is the chief constituent of many commercial preparations used for removing scales from automobile radiators. Oxalic acid was first prepared synthetically by a Swedish chemist, Carl Wilhelm Scheel, in 1776. It was manufactured by heating sodium formate in the presence of an alkali catalyst or by oxidizing carbohydrates with nitric acid (Anonymous,1982).

1

However, oxalic acid has wide range of industrial application, but most of the oxalic acid used in Nigeria are imported. As a result, over the years, research is on to produce oxalic acid locally. Investigation on eucalyptus camaldulensis revealed that the bark contains calcium oxalate (Shafii,1998). Because of the presence of calcium oxalate in the bark of eucalyptus camaldulensis, oxalic acid is produced by reacting the grinded bark with sulphuric acid as summarized by the equation below:

$$CaC_2O_4 + H_2SO_4 \rightarrow (CO_2H)_2 + CaSO_4$$
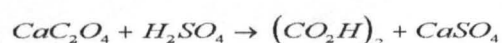
After reaction for a chosen time, oxalic acid is obtained by concentrating the resulting filtrate.

A lot of work has been done by final year students of this department on the production of oxalic acid from the bark of eucalyptus camaldulensis. Each student varied factors such as particle size, temperature, reaction time, acid concentration, volume of acid and mass of sample. It is the combination of these factors that I will use to analyze the available results.

The analysis is focused on using statistical investigation to analyze and predict the percentage extract of oxalic acid. There are several experiments, which require statistical investigation. These are characterized by the nature of treatments under investigation and also the nature of comparison required among them so as to meet the objective of the experiment. There are three main types of experiment: (i) varietal trials, (ii) factorial experiment, and (iii) bio-assays.

In the course of this work, factorial experimental analysis will be used to analyze and predict the percentage extract with 3 factors at 2 levels i.e 2^3 factorial. Factorial experiment involves simultaneously more than one factor each at two or more levels. If the number of levels of each factor in an experiment is the same, the experiment is called symmetrical factorial; otherwise, it is called asymmetrical factorial or sometimes mixed factorial. These experiments provide an opportunity to

2

study not only the individual effects of each factor but also their interactions. When the experiments are conducted factor by factor, changing the levels of one factor at a time and keeping the other factors at constant levels, the effect of interaction cannot be investigated. In many biological and clinical trials, factors are likely to have interactions. Therefore, factorial types of experiments are more informative in such investigations. They have the further advantage of economizing on experimental resources. When experiments are conducted factor by factor, much more resources are required for the same precision than when they are tried in factorial design (Das and Giri, 1979).

A 2^k Full Factorial experiment entails the choice of factors and their levels (high and low). The response variable is also clearly defined for any experimental data. After necessary statistical test, a model equation that describes the process results. The model equation can also be used to predict percentage extract for the variables within the range of the chosen levels.

Finally, a computer program will be developed which codes the steps of getting the model equation used to predict the percentage extract. And the results will be simulated by the program for comparison with the experimental results.

1.1 AIMS AND OBJECTIVES

1. To use 2^3 full factorial design equation in predicting the percentage extract of oxalic acid from the bark of eucalyptus camaldulensis.

2. To see the effect(s) of chosen factors (operating variables), on the percentage extract.

3. To produce a computer program that can be used for complementing 2^3 factorial design analysis.

4. To compare the result for hand calculation with that of the program.

5. To see how choice of level of factors affects the model equation.

3

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 EUCALYPTUS AND ITS SPECIES.

The word 'Eucalyptus' is of Greek origin meaning "well covered" (Keay, 1989). Eucalyptus is a large genus of trees and shrub, native of Australia, Tasmania, New Zealand and Malaysia and near by island (Anonymous, 1982). The height of eucalyptus tree varies from one specie to the other. The smallest may be less than 6.5 feet (2 meters) tall, while the tallest may be about 330 feet (100 meters) tall (Anonymous, 1994).

The leaves of eucalyptus are hard oblique or vertical, with leaves of many species containing aromatic oil called eucalyptus oil. The margins are smooth, but the edges of few species are wavy or slightly toothed. The fruit of eucalyptus are capsules surrounded by a wood, cap shaded receptacle and contains numerous minute seeds. The capsule is topped by a disc that breaks up into two to seven distinct valves depending on the specie. Some seeds are prominently winged while others are smooth or sculptured (Anonymous, 1994).

2.1.1 EUCALYPTUS SPECIES

There are 650- 700 species as well as subspecies of eucalyptus. Some species of eucalyptus are given below starting with the raw material of this work i.e. eucalyptus camaldulensis.

1. Eucalyptus cameldulensis: This specie is widely cultivated in plantation and from self sowed seed; variable and hybridizing with other species. The leaves are narrowly lanceolate up to 25cm long and 1.5cm broad while the bark is smooth, ash coloured and peeling.

4

2. Eucalyptus cloeziana: This is another specie with brown stingy bark. The leaves are lanceolate and up to 12cm long and 3cm broad.

3. Eucalyptus citriodora. This is widely grown in towns for decoration with smooth and polished bark. The leaves are lanceolate and are 16cm long and 2cm broad, they are also strongly lemon- scented.

4. Eucalyptus deglupta: Most of this species are large with smooth reddish bark. The leaves are ovate-lanceolate, they are up to 14cm long and 7cm broad.

5. Eucalyptus torelliana: They are usually planted in the towns for shade. The bark is scaly and persistent below, smooth and peeling above (Keay, 1989). These are just few out of the many species of eucalyptus trees.

2.1.2   ECONOMIC IMPORTANCE OF EUCALYPTUS

Eucalyptus Tree is of great economic importance because of the valuable products obtained from it. Some of the uses to which eucalyptus can be put in to are given below.

1. **Tannin**: This is usually contained as soluble materials in the bark of eucalyptus species. As the name implies tannins are used for tanning skin of animals to remove the hair and improve their quality. The tanning process converts raw hides to leather (Irwin, 1981). Tannins can be exploited for commercial purposes, sold to various tannaries and leather factories which at present depend on imported mineral tannins.

2. **Alkanoids**: They are obtained on commercial scale from the leaves and bark of eucalyptus as drug rutin. Alkanoids are also chemically basic compounds with an important physiological activity as analgesics in medicine.

5

3. **Charcoal**: This yet another raw material that could be derived from the bark of eucalyptus. Apart from the use in the production of bread by the supply of heat in most African countries, most eucalyptus trees carbonize easily providing good commercial charcoal. The charcoal from eucalyptus tree yields substantial amount of energy of about 7, 900 calories per kilogram against 4700 calories per kilogram from wood (Hill, 1979).

4. **Essential oil**: The essential oils found in eucalyptus species are extracted from the leaves of eucalyptus citriodora, eucalyptus robusta, eucalyptus smitti and eucalyptus globulus (Anonymous, 1982). The oil finds a wide application in our day-to-day life. Among the used of the oil is that it could be used to manufacture performs and soaps. The terpene and ketone derivatives in the oil are used as stain removers and also used in veterinary medicine while the piperitone is used in manufacture of synthetic thymol and menthol.

5. **Fuel wood**: Eucalyptus trees are also used as fuel wood. The wood of most eucalyptus species burns well when air-dried and leave little ash. Eucalyptus plantation and other forest resources can supply a substantial part of the household fuel needed. One of the most important advantage of using eucalyptus as a fuel is that it does not smoke (Hill, 1979).

6. **Timber**: Eucalyptus trees are also grown to serve as timber, which vary in properties and uses. Because of the height of the tree, it can also be used as electric and telephone poles. The trees are also used in making canoes. The important timber species are eucalyptus maculata and eucalyptus citriodora (Anonymous, 1994).

7. Pulp and paper: Many paper industries use the bark of eucalyptus species as raw material for the production of paper. The world production of pulp from eucalyptus is over one million tonnes annually.

8.   **Ornamentals**: As ornamentals, eucalyptus are grown for their attractive form and foliage, colourful flowers, or decorative bark. The scarlet-flowering gums (E. ficifolia and E. calophylla) are two widely planted trees with brightly coloured flowers. Eucalyptus citriodora is grown for its red flowers and attractive form. Species with decorative bark include the red iron bark (E. sideroxylon), which has black, furrowed bark; the spotted gum (E. maculata), which has smooth, mottled bark in shades of gray, tan, and reddish brown; and smooth white bark that trails in long ribbons from the branches (Anonymous, 1994).

Other uses

-    Local medicines

-    Production of local ink

-    Wind breaks etc.

2.2   OXALIC ACID

Oxalic acid otherwise known as ethanedioic acid is an important individual diacids. It exists usually as a crystal hydrate, $HOOC\text{-}COOH.2H_2O$, whose melting point is $101°C$ and anhydrous acids which melts at $189.5°c$. Oxalic acids are found naturally as oxalates in free state. Calcium oxalate is contained in all plants, while potassium acid salt, $KOOC\text{-}COOH$ in dock, oxalis. Oxalates of alkaline metals dissolve in H_2O, while calcium oxalates practically does not dissolve in cold H_2O. When metabolism is upset in human organism, calcium oxalates accumulates, causing the formation of stones in the liver, kidneys and urinary tracts.

The low solubility of calcium oxalate is used in analytical chemistry for quantitatively and qualitatively determination of calcium.

7

2.2.1 USES AND IMPORTANCE

Oxalic acid is used industrially for dyeing textiles and removing rust stains and ink blots from them, because it converts most insoluble ionic compounds into soluble complex ion. And for the same reason, it is the chief constituent of many commercial preparation used for removing scales from radiators of automobile (Anonymous, 1982).

Being a product of oxidation of many organic substances, oxalic acid itself is readily oxidized. For example, it is oxidized by potassium permanganate thus:

$$HOOC\text{-}COOH \xrightarrow[KMnO_4]{[O]} > 2CO_2 + H_2O \ldots\ldots\ldots\ldots 2.1$$
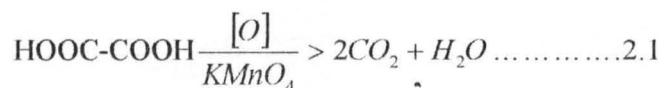
This property is used in Analytical chemistry for preparing standard solutions of potassium permanganate.

Oxalic acids is used for the manufacture of ink and for bleaching straw. It is also used in Leather industries in manufacture of dyes and in the preparation of related compounds such as glyoxalic acid and glycolic acid. (Anonymous, 1994).

2.2.2 PHYSICAL AND CHEMICAL PROPERTIES

Physical properties: `

It is a colourless crystalline solid, which sublimes at 160°C and is fairly soluble in water. It can also be obtained as crystalline dihydrate, $(COOH)_2.2H_2O$ (Arene and Kitwood, 1979). Oxalic acids are also soluble in ethanol but insoluble in $CHCl_3$.

Oxalic acid is very poisonous; it melts at 101.5°C when hydrated, while when anhydrous, the melting point is 189.5°C (Anonymous, 1982).

8

Chemical properties:

In most chemical reactions, oxalic acids shows the typical properties of carboxylic groups (Arene and Kitwood, 1979). This will be illustrated by some of the chemical process below:

i.   **Esterification**: Anhydrous oxalic acid refluxed with excess methanol or ethanol yields esters without need of stronger acid as catalyst:

$$(CO_2H)_2 + 2C_2H_5OH \rightarrow (CO_2C_2H_5)_2 + 2H_2O \ldots\ldots\ldots 2.2$$
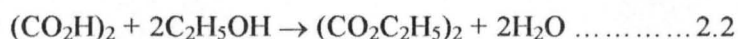
ii.   **Amide formation**: Ammonium oxalate yields some oxamide on heating

$$(CO_2NH_4)_2 \rightarrow (CONH_2)_2 + 2H_2O \ldots\ldots 2.3$$

iii.   **Acid chloride formation**: Oxâlyl chloride (b.p.64°C) is obtained by the action of Phosphorus pentachloride and similar reagents on oxalic acid.

$$(CO_2H)_2 + 2PCl_5 \rightarrow (COCl)_2 + 2POCl_3 + 2HCl \ldots\ldots 2.4$$
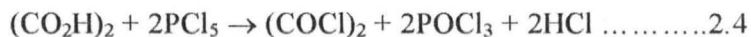
iv.   Alkali metal oxalates heated with Soda lime yield hydrogen and carbonate. Formally at any rate, this reaction is analogous with the decarboxylation of other carboxylic acid by this method. (Dazaley, 1979).

$$\begin{bmatrix} CO_2^{(-)} \\ CO_2^{(-)} \end{bmatrix} + 2OH^{(-)} \xrightarrow[dry]{heat} \quad 2CO\, CO_3^{2(-)} + H_2 \ldots\ldots\ldots\ldots 2.5$$
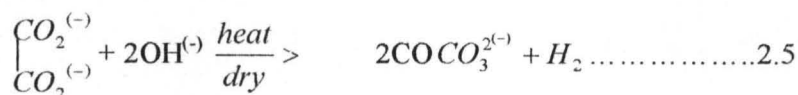
The above reactions are more or less typical carboxylic acid behaviour. Oxalic acid does not form an anhydrous.

v.   On heating with concentrated H_2SO_4, oxalic acid and oxalates are dehydrated to an equimolar mixture of carbon monoxide and dioxide;

$$(CO_2H)_2 \xrightarrow[heat]{Con.H_2SO_4} CO + CO_2 + H_2O \ldots\ldots\ldots\ldots 2.6$$
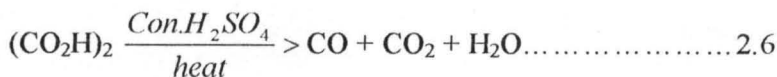
This behaviour is the familiar elementary test oxalic acid (Dazeley, 1979).

9

2.2.3 METHOD OF PRODUCTION

The usual laboratory method for preparing oxalic acid is by oxidizing sucrose with concentrated nitric acid in the presence of vanadium pentoxide. (Anonymous, 1994)

$$C_{12}H_{22}O_{11} \xrightarrow[V_2O_5]{HNO_3} > 6(CO_2H)_2 + 5H_2O\ldots\ldots\ldots\ldots\ldots.2.7$$
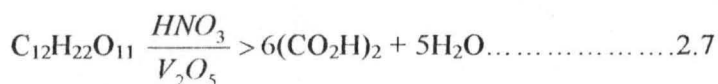
In this work the data is generated from the production of Oxalic acid from the bark of eucalyptus camaldulensis. This is possible because of the presence of calcium oxalate in the bark of the eucalyptus tree. Sulphuric acid is added to the grinded bark so that the calcium oxalate in the bark reacts with the sulphuric acid to give oxalic acid and calcium sulphate as a by-product, thus:

$$CaC_2O_4 \quad + \quad H_2SO_4 \quad \rightarrow \quad (CO_2H)_2 + \quad CaSO_4 \ldots\ldots\ldots.2.8$$

Calcium   Sulphuric   Oxalic   Calcium

oxalate   acid   acid   sulphate

The experimental procedure of the process is given below:

Experimental procedure

The bark of eucalyptus camaldulensis was soaked in water to extract the Tannin and other water-soluble contents and then dried. The bark was crushed in a mortar and screened into samples of different size ranges.

A known sample weight of each size range was put in a conical flask and a chosen volume of a known concentration of dilute H_2SO_4 acid solution was added to it. The mixture was agitated for a specified time at a specified temperature (TEMP 1) on a hot plate equipped with a magnetic stirrer.

The hot solution was filtered into a beaker. The filterate (FILT 1) obtained was weighed and then allowed to cool for an hour at room temperature. The thick precipitate of calcium sulphate formed after cooling was filtered off and weighed. This was called PREPT 1. The filterate obtained which constituted the second filterate (FILT 2) was also weighed. The second filterate was concentrated to about one-third of its original volume at a specified temperature (TEMP 2) in a moisture extraction oven, till the colour of the filtrate turns wine red. This filtrate was referred to as 'concentrated filtrate' (FILT 3).

The concentrated filtrate was allowed to cool at room temperature for another one hour, during which another precipitate (PREPT 2) of $CaSO_4$ was formed. This precipitate was filtered off to obtain the final filtrate (FILT 4). The final filtrate was weighed and kept in a refrigerator overnight to allow the oxalic acid contained in it to crystallize.

The supernatant liquid was decanted from the oxalic acid crystal which were carefully rinsed with very cold water (temperature below 4°C) to remove traces of H_2SO_4 adhering to the crystals. The low temperature was necessary to prevent dissolution of the oxalic acid. The percentage purity of the oxalic acid obtained was determined by volumetric analysis. The whole procedure with its set of operating variable constituted a run. The results were tabulated accordingly in Chapter four.

2.3 BASIC PRINCIPLE OF EXPERIMENTAL DESIGN

If an experiment is to be performed most efficiently, then a scientific approach to planning the experiment must be considered. By statistical design of experiments, we refer to the process of planning the experiment so that appropriate data will be collected, which may be analyzed by statistical methods resulting in valid objective

11

conclusions. The statistical approach to experimental design is necessary if we wish to draw meaningful conclusions from the data. When the problem involves data that are subjected to experimental errors, statistical methodology is the only objective approach to analysis. Thus, there are two aspects to any experimental problem, the design of the experiment and the statistical analysis of the data. These two subjects are closely related, since the method of analysis depends directly on the design employed.

The two basic principles of experimental design are replication and randomization. By replication, we mean a repetition of the basic experiment. If a treatment is allotted to 'r' experimental units in an experiment, it is said to be replicated r times. If in a design each of the treatments is replicated r times, the design is said to have r replications. Replication is necessary to increase the accuracy of estimates of the treatment effects. It also provides an estimate of the error variance which is a function of the difference among observations from experimental units under identical treatments. Although the more the number of replications, the better it is, so far as precision of estimates is concerned, it cannot be increased indefinitely as it increases cost of experimentation.

Randomization is the cornerstone underlying the use of statistical methods in experimental design. By randomization we mean both the allocation of the experimental material and the order in which the individual runs or trials of the experiment are to be performed are randomly determined. Statistical methods require that the observations (or errors) are independently distributed random variables. Randomization usually makes this assumption valid. By properly randomizing the experiment, we will also assist in 'averaging out' the effects of extraneous factors that may be present.

In order to use the statistical approach to designing and analyzing an experiment, it is necessary that everyone involved in the experiment has a clear idea in advance of exactly what is to be studied, how the data is to be collected, and at least a qualitative understanding of how this data is to be analyzed. An outline of the recommended procedure is as follows:

1.   **Recognition and statement of the problem**: This may seem to be a rather obvious point, but in practice it is often not simple to realize that a problem requiring experimentation exists, and to develop a clear and generally accepted statement of this problem. It is necessary to develop all ideas about the objectives of the experiment. A clear statement of the problem often contributes substantially to a better understanding of the phenomena and the final solution of the problem.

2.   **Choice of factors and levels:** The experimenter must select the independent variables or factors to be investigated in the experiment. The factors in an experiment may either be quantitative or qualitative. If they are quantitative, thought should be given as to how these factors are to be controlled at the desired values and measured. We must also select the values or levels of the factors to be used in the experiment. These levels may be chosen specifically, or selected at random from the set of all possible factors levels.

3.   **Selection of response variable:** In choosing a response or independent variable, the experimenter must be certain that the response to be measured really provides information about the problem under study. Thought must also be given to how the response will be measured, and the probable accuracy of the measurements.

13

4.   **Choice of experimental design:** This step is of primary importance in the experimental process. The experimenter must determine the difference in true response he wishes to detect and the magnitude of the risk he is willing to tolerate so that an appropriate sample size (number of replicates) may be chosen. He must also determine the order in which the data will be collected and the method of randomization to be employed. It is also necessary to maintain balance between statistical accuracy and cost. Most recommended experimental designs are both statistically efficient and economical, so that the experimenter's efforts to obtain statistical accuracy usually result in economic efficiency. A mathematical model for the experiment must also be proposed, so that statistical analysis of the data may be performed.

5.   **Performing the experiment:** This is the actual data collection process. The experimenter should carefully monitor the process of the experiment to ensure that it is proceeding according to the plan. Particular attention should be paid to randomization, measurement accuracy, and maintaining as uniform an experimental environment as possible.

6.   **Data analysis**: Statistical methods should be employed in analyzing the data from the experiment. Numerical accuracy is an important concern here, although present day computers have largely relieved the experimenter from this problem, and simultaneously reduced the computational burden. Graphical methods are also frequently useful in the analysis process.

7.   **Conclusions and recommendations**: Once the data have been analyzed, the experimenter may draw conclusions or inferences about his results. The statistical inferences must be physically interpreted, and the practical significance of those finding evaluated. Then recommendations concerning these finding must be made. These recommendations may include a further round of experiments, as experimentation is usually an iterative process, with one experiment answering some question and simultaneously posing others (Montgomery, 1976).

2.3.1   FACTORIAL EXPERIMENT

In factorial experiments, combination of two or more levels of more than one factor are the treatments. For example with two factors, we can have (i) nitrogen fertilizer at two levels, denoted by n_1, and n_2 and (ii) irrigation at two levels, l_1 and l_2 in an agricultural experiment. We can form the following four combinations taking one level from each factor, l_1n_1, l_1n_2, l_2n_1 and l_2n_2. Such combinations form treatments in factorial experiments. The comparison required in this type of experiments are not the pair comparison as in varietal trials but a special type of comparison called main effects and interactions (Das and Giri, 1979).

2.3.2 2^k FULL FACTORIAL EXPERIMENT

There are several special cases of general factorial deigns that are important because they are widely used in research work and also because they form the basis of other designs of considerable practical values. The first of these special cases is that of two factors, each at only two levels. These levels may be quantitative or qualitative, such as two values of temperatures, pressure, or time; or they may be qualitative such as two machines, two operators, the "high" and "low" levels of factors, or perhaps the presence and absence of a factor. Such design requires: 2 x 2 x

15

$2 \ldots \ldots x2 = 2^k$ observations and is called a 2^k factorial design. The special second case is that of k factors, each at three levels, which is called a 3^k factorial design (Montgomery, 1976).

A 2^k factorial design requires us to choose just two levels for each factor and then calls for simulation runs at each of the 2^k possible combinations of factor levels. Usually we use a minus sign with one level of a factor and a plus sign with the other level. Which sign is associated with which level is quite arbitrary, although for quantitative factors it is less confusing if we associate the minus sign with lower numerical value. No general prescription can be given for how one should specify the levels (Averill and Kelton, 1996).

2.3.3 A 2^2 DESIGN

The first design in the 2^k series is one with only two factors say A and B, each run at two levels. This design is the simplest case of 2^k series and is called a 2^2 factorial design. The levels of the factor may be arbitrarily called "low" and "high". The treatment combination and response of this design is displayed below:

Table 2.1: Treatment combination and responses for 2^2 design.

Treatment Combination	Response R.
A low, B low	R_1
A high, B low	R_2
A low, B high	R_3
A high, B high.	R_4.

The treatment combinations in this design are shown graphically in Figure 2.1. By convention we denote the effect of a factor by a capital Latin letter. Thus "A" refers to the effect of factor A. "B" refers to the effect of factor B and "AB" refers to the AB interaction. In the 2^2 design the low and high levels of A and B are denoted by 0 and 1, respectively on the A and B axes.

16

The coordinates of the vertices of the square also represent the four treatment combinations as follows: 00 represent both factors at low level, 10 represents A at the high level and B at the low level, 01 represents A at the low level and B at the high level, and 11 represents both factors at the high level. These treatment combinations are usually represented by lower case letters, as shown in Figure 2.1. We can see from the figure that the corresponding lower case letter denotes the high level of any factor in the treatment combination, and the low level of a factor in the treatment is denoted by the absence of the corresponding letter. Thus, "a" represents the treatment combination of A at the high level and B at the low level, "b" represents A at low level, and B at high level, and "ab" represents both factors at high level. By convention, "1" is used to denote both factors at the low level.

The average effect of a factor is the change in the response produced by a change in the level of the factor, averaged over the levels of the other factors.
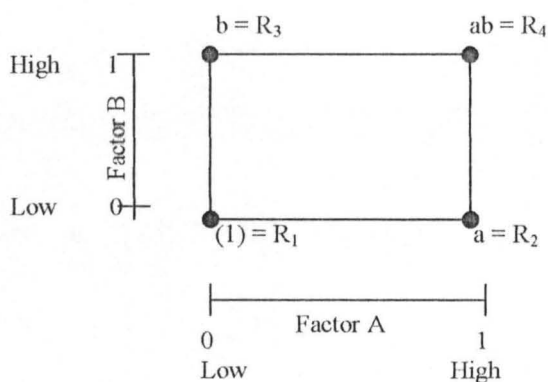
Fig. 2.1: **Treatment combination in a 2^2 design**

The lower case letters "1", "a", "b" and "ab" now represent the total of all the responses of all treatment combination, as illustrated in Fig 2.1. Now the simple effects of A at the low level of B is

$$[a-(1)] = R_2-R_1 \qquad \qquad \dots\dots\dots\dots\dots 2.9$$

and the simple effect of A at the high level of B is

$$[ab-b] = R_4-R_3 \qquad \qquad \dots\dots\dots\dots\dots 2.10$$

17

Averaging these two quantities yields

$$A = \frac{1}{2} \{ [ab-b]+[a-(1)] \} = \frac{1}{2} \{[R_4-R_3]+[R_2-R_1]\}$$

$$= \frac{1}{2} [ab+a-b-(1)] = \frac{1}{2} [R_4+R_2-R_3-R_1] \qquad \dots\dots\dots\dots\dots 2.11$$

The average B effect is found from the simple effect B at low level of A

i.e. $[b-(1)] = [R_3-R_1]$ $\qquad \dots\dots\dots\dots\dots\dots 2.12$

and at high level of A

i.e $[a b - a] = [R_4 - R_2]$ $\qquad \dots\dots\dots\dots\dots\dots 2.13$

so that

$$B = \frac{1}{2} \{[ab-a]+[b-(1)] = \frac{1}{2} \{[R_4-R_2]+[R_3-R_1]\}$$

$$= \frac{1}{2} [ab+b-a-(1)] = \frac{1}{2} [R_4+R_3-R_2-R_1] \qquad \dots\dots\dots\dots\dots 2.14$$

The average effect for the interaction AB is the difference between the effects

of A at the low level of B and the effect of A at low level of B. Thus:

$$AB = \frac{1}{2} \{ [ab-b]-[a-(1)] \} = \frac{1}{2} \{[R_4-R_3]-[R_2-R_1]\}$$

$$= \frac{1}{2} [ab+(1)-a-b] = \frac{1}{2} [R_4+R_1-R_2-R_3] \qquad \dots\dots\dots\dots\dots\dots 2.15$$

From Equation 2.11 contrast for estimating A is given by

Contrast A = a b+ a −b −(1)

$$= R_4-R_2-R_3-R_1 \qquad \dots\dots\dots\dots\dots\dots 2.16$$

All other contrasts are calculated in a similar manner Kempthorne (1952) and

Anderson and McLean (1974) call this contrast the total effect of A.

2.3.4 SIMPLE AND AVERAGE EFFECTS

To illustrate the simple effect of a factorial experiment, let us consider an experiment on sugar beet with 2 factors. These were without nitrogen (n_o) versus 3 cwt. sulphate of ammonia per acre, with nitrogen (n_1) and depth of winter (shallow ploughing (7in) versus deep ploughing (11in) took place in late January, the nitrogen was applied in the late April, and the seed was sown early in May. Since both factors occur at 2 levels of variations, the experiment is described as 2x2 factorial experiment. The 4 treatment combinations are shown below:

Table 2.4:   Treatment combination and yield of sugar (cwt. per acre).

Combination		Yield	(R)
(1)	n_o, 7in	40.9	R_1
a	n_1, 7in	47.8	R_2
b	n_o, 11in	42.4	R_3
ab	n_1, 11in	50.2.	R_4

The results might be summarized as follows. Considering the simple effect of nitrogen, from Equations 2.9 and 2.10, we might report that the application of nitrogen increased the yields by

$R_2 - R_1 = 49.8 - 40.9 = 6.9$ cwt

with shallow ploughing and by

$R_4 - R_3 = 50.2 - 42.4 = 7.8$ cwt

with deep ploughing. These figures are called the simple effects of nitrogen. They represent the type of information that could be wanted for instance in giving advice to a farmer who always used shallow ploughing but was doubtful whether to apply nitrogen. For the simple effect of depth of ploughing from Equation 2.12 and 2.13 we might report that 11 in ploughing was superior to 7in by

$R_3 - R_1 = 42.4 - 40.9 = 1.5$ cwt

in the absence of nitrogen and by

$R_4 - R_2 = 50.2 - 47.8 = 2.4$ cwt

when nitrogen was applied.

This is another way of looking at the results. It sometimes happens that the effects of the factors are independent. By this we mean that the response to nitrogen is the same whether ploughing is shallow or deep, and that the difference between the effects of deep and shallow ploughing is the same whether nitrogen is present or not.

In this event, the two simple effects of nitrogen, 6.9 cwt and 7.8 cwt are estimates of the same quantity and differ only by experimental errors. On this supposition, we would naturally average the two figures in order to estimate the response to nitrogen. The average, 7.4 cwt, is called the average effect of nitrogen. It can be derived alternatively from Equation 2.11 thus:

$\frac{1}{2}(R_4 + R_2 - R_3 - R_2) = \frac{1}{2}(5o.2 + 47.8 - 42.4 - 40.9) = 7.4$ cwt.

Similarly, the average effect of depth of ploughing (11 in minus 7 in) is the average of 1.5 cwt and 2.4 cwt, or 1.9 cwt. Note that the average effect, is an average of the simple effects.

Consequently, if we are sure that the factors operate independently, the summary that was given above in terms of the simple effects may be replaced by another that is both more concise and more accurate. This might read as follows. "The application of nitrogen increased the yield of sugar by 7.4 cwt, while 11 in. ploughing increased the yield by 1.9 cwt as compared with 7in ploughing." It is worth repeating that when the factors are independent the figure 7.4 cwt is the best estimate not only of the average response to nitrogen, but also of the response on plots ploughed to 7in and of that on plots ploughed to 11in. In other words, the whole of the information in the experiment is contained in the average effects.

21

It is important to consider when the factors are independent. This is determined by knowledge of the processes by which the factors produce their effects. In the present example an agronomist might reason that deep ploughing should enable the plant to develop a more vigorous root system. With this the plant should be able to utilize more effectively any added nutrients such as nitrogen. Thus he might predict the response to nitrogen would be greater with deep than with shallow ploughing, though he would not expect it to be much greater. In short, he would predict that two factors would not be quite independent in their effects (Cochran and Cox, 1957).

In addition to the information that may be available from such reasoning, a factional experiment itself provides a test of the assumption of independence. If the depth of ploughing does affect the response to nitrogen, the difference between 7.8 cwt. (the response to the nitrogen with deep ploughing) and 6.9 cwt. (the response with shallow ploughing) is an estimate of this effect. The difference, 0.9 cwt., can be tested in the usual way by a T-test. If the test proves significant, the assumption of independence is rejected by the data. The difference (sometimes divided by a numerical factor) is called the interaction between nitrogen and depth of ploughing. T-test is a special case of F-test (Linton,1965).

Interchanging the roles of the two factors, allows us to consider whether the superiority of deep over shallow ploughing is affected by the presence of nitrogen. To measure the interaction in this case, we subtract 1.5 cwt. (superiority of deep ploughing when no nitrogen is added) from 2.4 cwt. (superiority when nitrogen is added). The difference is again 0.9 cwt. It is easy to see that this equality always holds with a 2x2 experiment (Cochran and Cox, 1957).

22

Example 2.1

To illustrate the average effect of a 2^2 factorial experiment, we use an inventory model with two experimental factors, A and B. The "low" and "high" values chosen for A and B are given in the "coding chart" in Table 2.5. The design matrix and corresponding response variables are given in Table 2.6. (The sign for A x B interaction are also included).

Table 2.5:   **Coding for A and B in the inventory model.**

Factors	-	+
A	20	60
B	70	120

Table 2.6:   **Design matrix and empirical results in a 2^2 factorial design on A and B for the inventory model.**

Factor combination	A	B	A x B	Response
1	-	-	+	118.280
2	+	-	-	141.060
3	-	+	-	136.807
4	+	+	+	152.789

From Equation 2.11, the average effects are:

$A = 1\backslash 2(R_4 + R_2 - R_3 - R_1)$ rearranging gives

$A = 1\backslash 2(-R_1 + R_2 - R_3 + R_4)$

$\quad = 1\backslash 2(-118.280 + 141.060 - 136.807 + 152.789)$

$\quad = \quad 19.381$

and

$B = \frac{1}{2}(-118.280 - 141.060 + 136.807 + 152.789)$

$\quad = 15.128$

and for A x B interaction,

$AB = 1\backslash 2(118.280 - 141.060 - 136.807 + 152.789)$

$\quad = -3.399$

Thus, the average effect of raising A from 20 to 60 was to raise the monthly cost by 19.381, and raising B from 70 to 120 increase the monthly cost by an average of 15.128. Therefore it appears that the smaller values of A and B would be preferable, since lower monthly costs are desired. Since the A x B interaction effect is negative there is indication that lower cost are observed by setting both A and B at either their − or + levels rather than one at the − level and one at the + level. The magnitude (absolute value) of the average interaction effect is much smaller than the magnitudes of the average main effects, which often happens in factorial experiments.

Calculation of average main effects and interaction of a 2^k factorial experiments is actually equivalent to estimating the parameters in a particular statistical regression model of how the response depends on the factors (Averill and

24

Kelton,1996). For the model above this regression model is:

$$R(A, B) = \eta + \beta X_A + \varnothing X_B + \varphi X_A X_B + \epsilon \ldots\ldots\ldots\ldots 2.17$$

Where R(A, B) = response as a function of A and B.

In order to transform A and B to coded factors we use,

$$X_A \quad = \quad \frac{(A - 40)}{20} \quad\quad \ldots\ldots\ldots\ldots\ldots 2.18$$

$$X_B \quad = \quad \frac{(B - 95)}{25} \quad\quad \ldots\ldots\ldots\ldots 2.19$$

and ϵ is random variable with mean zero. Given this regression model, we could estimate η, β, $\varnothing$ and φ by ordinary least-square regression and use the fitted model to forecast the value of E[R(A, B)] at combinations of A and B where no simulation has been done. Moreover the average main effect A is twice β, the least-square estimator of β. Similarly, $\quad B = 2 \varnothing, \quad\quad \ldots\ldots\ldots\ldots 2.20$

$$AB = 2\varphi \ldots\ldots 2.21 \quad \text{and} \quad \eta = \frac{1}{4} \sum_{i=1}^{4} R_i \ldots\ldots\ldots 2.22$$

Therefore the regression coefficient could be calculated directly from:

$$\eta = \frac{1}{2^k} \sum_{i=1}^{4} (R_i) \ldots\ldots\ldots\ldots\ldots\ldots 2.23$$

$$\beta = \frac{1}{2^k} \sum_{i=1}^{4} (X_A Ri) \ldots\ldots\ldots\ldots\ldots\ldots 2.24$$

$$\phi = \frac{1}{2^k} \sum_{i=1}^{4} (X_B Ri) \ldots\ldots\ldots\ldots\ldots 2.25$$

$$\varphi = \frac{1}{2^k} \sum_{i=1}^{4} (X_A X_B Ri) \ldots\ldots\ldots\ldots 2.26$$

where k = 2

25

2.3.5 ADVANTAGE OF FACTORIAL EXPERIMENT WHEN FACTORS ARE INDEPENDENT

The advantages of factorial experimentation naturally depend on the purpose of the experiment. We suppose for the present that the purpose is to investigate the effects of each factor over some pre-assigned range that is covered by the levels of that factor, which are used in the experiment. In other words the object is to obtain a broad picture of the effects of the factors rather than to find, say the combination of the levels of the factors that give maximum response. One procedure is to conduct separate experiment each of which deals only with a single factor. Another is to include all factors simultaneously by means of a factorial experiment.

If all factors are independent in their effects, the factorial approach will result in a considerable saving of time and maternal devoted to the experiments. The saving results from two facts. First, as we have seen, when factors are independent, all simple effects of a factor are equal to average effect, so that average effects are the only quantities needed to describe fully the consequences of variations in the factor. Secondly, in a factorial experiment, each average effect is estimated with the same precision as if the whole experiment had been devoted to that factor alone. Thus, in the preceding example, half the plot receives nitrogen and half do not. Consequently, the average effect of nitrogen is estimated just as precisely as it would be in a simple experiment of the same size devoted to nitrogen alone. The some result holds for the effect of depth of ploughing. The two single-factor experiments would require twice the total number of plots in order to equal the precision obtained by the factorial experiment. If there are n factors, all at two levels and independent, the single-factor approach would necessitate n times as much experimental material as a factorial arrangement of equal precision.

26

2.3.6   FACTORIAL EXPERIMENTATION WHEN FACTORS ARE

DEPENDENT

We assume that the purpose is still to investigate each factor over the range represented by its levels. When factors are not independent the simple effects of a factor vary according to the particular combination of the other factors with which these are produced. In this case, the single-factor approach is likely to provide only a number of disconnected pieces of information that cannot easily be put together. In order to conduct an experiment on a single factor A, some decision must be made about the levels of other factors B,C,D say, that are not used in the experiment (e.g, whether all plots should be ploughed 7in, 9in, or 11in deep in an experiment on nitrogen). The experiment reveals the effects of A for this particular combination of B,C, and D, but no information is provided for predicting the effects of A, with any other combination of B, C, and D. With a factional approach on the other hand, the effects of A are examined for every combination of B, C, and D, that is included in the experiment. Thus a great deal of information is accumulated both about the effect of the factor and about their interrelationships.

In this connection, Fisher (1966) has pointed out that it is sometime advisable to introduce into an experiment an extra factor that is not itself of interest, in order that the experiment may form the basis for sounder recommendations about the other factors. In agricultural experimentation in Britain, farmyard manure has served as a subsidiary factor of this kind (Cochran and Cox, 1957).

2.3.7 A 2^3 DESIGN

If three factors A, B, and C each at two levels, are under study, the design is called a 2^3 factorial, and eight treatment combinations can now be displayed graphically as a cube, as shown in Figure 2.2. The treatment combination in standard order are (1), a, b, ab, c, ac, bc and abc (Montgomery,1976).
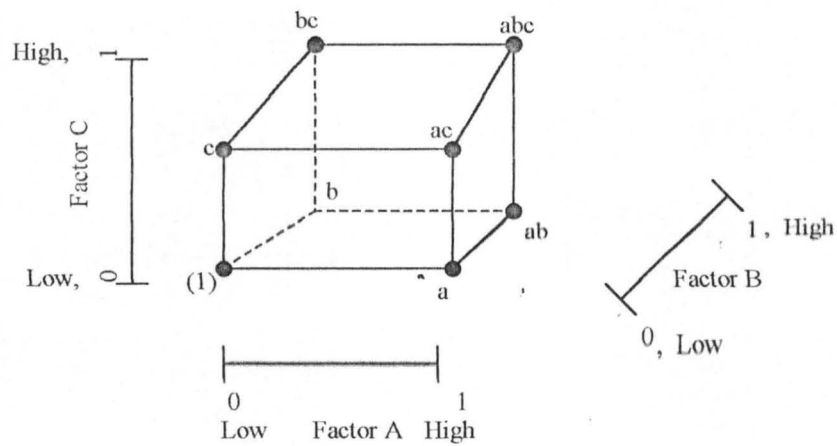
Figure 2.2:   **Treatment combination in 2^3 design.**

If the factors A, B, and C are changed to 1, 2, and 3 the form of the experiment can be represented in tabular form, as exemplified in Table 2.7.

Table 2.7: **Design matrix table for a 2^3 factorial design**

Factor comb.	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	Resp. (R_i)
1 = (1)	+	-	-	-	+	+	+	-	R_1
2 = a	+	+	-	-	-	-	+	+	R_2
3 = bc	+	-	+	-	-	+	-	+	R_3
4 = ab	+	+	+	-	+	-	-	-	R_4
5 = c	+	-	-	+	+	-	-	+	R_5
6 = ac	+	+	-	+	_	+	-	-	R_6
7 = ac	+	-	+	+	-	-	+	-	R_7
8 = abc	+	+	+	+	+	-+	+	+	R_8

The variables R_i for i = 1, 2........8 are the values of the response when running the simulation with ith combination of factor levels. For example, R_6 is the response resulting from running the simulation with factor 1 at its + level, factor 2 at its – level, and factor 3 at its + level. We shall see later that writing down this array of + and – signs, called the design matrix, facilitates calculation of the factors effects and interactions.

The average main effect of factor j is the average change in the response due to moving factor j from its – level to its + level while holding all other factors fixed. For the 2^3 factorial the average main effect of design of factor 1 is thus:

$$e_1 = \frac{(R_2 - R_1) + (R_4 - R_3) + (R_6 - R_5) + (R_8 - R_7)}{4} \qquad \dots\dots\dots\dots\dots\dots2.27$$

Note that in combinations 1 and 2, factors 2 and 3 remain fixed, as they do in combinations 3 and 4, 5 and 6, and 7 and 8. The average main effect of factors 2 is

$$e_2 = \frac{(R_3 - R_1) + (R_4 - R_2) + (R_7 - R_5) + (R_8 - R_6)}{4} \qquad \dots\dots\dots\dots2.28$$

29

and that of factor 3 isS

$$e_3 = \frac{(R_5 - R_1) + (R_6 - R_2) + (R_7 - R_3) + (R_8 - R_4)}{4} \quad \dots\dots\dots\dots 2.29$$

Examination of Table 2.7 and the above expression for average main effects e_j's lead to an alternative way of defining average main effects, as well as a simpler way of computing them. That is, e_j is the difference between the average response when factor j is at its $+$ level and the average response when factor j is at its $-$ level. Thus, to compute e_j we simply apply the signs in "factor j" column to the corresponding R_i's, add them up and divide by 2^{k-1}. For example, in the 2^3 factorial design of Table 2.7,

$$e_2 = \frac{-R_1 - R_2 + R_3 + R_4 - R_5 - R_6 + R_7 + R_8}{4} \quad \dots\dots\dots\dots 2.30$$

which is equivalent to earlier equation 2.28.

The average main effects measures the average change in the response due to a change in an individual factor. However, it could be, that the effect of factor j_1 depends on the level of some other factor, j_2, in which case factors j_1 and j_2 are said to interact. We measure the degree of this interaction by the two factor interaction effect, $e_{j_1 j_2}$, between factors j_1 and j_2. It is defined to be half the difference between the average effect of factor j_1, when factor j_2 is at its $+$ level (all factors other than j_1 and j_2 are held constant) and average effect of j_1 when j_2 is at its $-$ level. ($e_{j_1 j_2}$ is also called $j_1 \times j_2$ interaction.) For example in the design of Table 2.7 we have:

$$e_{12} = \frac{1}{2}\left[\frac{(R_4 - R_3) + (R_8 - R_7)}{2} - \frac{(R_2 - R_1) + (R_6 - R_5)}{2}\right] \quad \dots\dots\dots 2.31$$

$$e_{13} = \frac{1}{2}\left[\frac{(R_6 - R_5) + (R_8 - R_7)}{2} - \frac{(R_2 - R_1) + (R_4 - R_3)}{2}\right] \quad \dots\dots\dots 2.32$$

$$e_{23} = \frac{1}{2}\left[\frac{(R_7 - R_5) + (R_8 - R_6)}{2} - \frac{(R_3 - R_1) + (R_4 - R_2)}{2}\right] \ldots\ldots\ldots 2.33$$

As with main effects, there is a much simpler way to compute average interaction effects, based on examination of the design matrix. If we rearrange the above expression for e_{13}, for example, so that the R_i's appear in increasing

order of the i's, we get:

$$e_{13} = \frac{R_1 - R_2 + R_3 - R_4 - R_5 + R_6 - R_7 + R_8}{4} \quad \ldots\ldots 2.34$$

Now if we create column, labeled $x_1 \times x_3$, of 8 signs by "multiplying" the ith sign in the "factor 1" column by the ith sign in "factor 3" column, the product of like signs is a "+", and the product of opposite signs is a "-". we get a column of signs which gives us precisely the signs of the R_i's used to obtain e_{13}. As with main effects, the divisor is 2^{k-1}. The computation of e_{12} and e_{23} are thus given below:

$$e_{12} = \frac{R_1 - R_2 - R_3 + R_4 + R_5 - R_6 - R_7 + R_8}{4} \ldots\ldots\ldots 2.35$$

and

$$e_{23} = \frac{R_1 + R_2 - R_3 - R_4 - R_5 - R_6 + R_7 + R_8}{4} \ldots\ldots\ldots 2.36$$

Finally, we note that two factors interaction effects are completely symmetric; that is, $e_{12} = e_{21}, e_{23}, = e_{32}$ etc.

Although their interaction become more difficult, we can define (and compute) three and higher factor interaction effects, all the way up to k-factor interaction. For example, in the 2^3 factorial design of Table 2.6, the three-factor interaction between factor 1, 2 and 3 is half the different between the average two factor interaction effect between factor 1 and 2 when factor 3 is at it's + level and the average two-factor interaction effect between factors 1 and 2 when factor 3 is at its - level. That is,

31

$$e_{123} = \frac{1}{2}\left[\frac{(R_8 - R_7) - (R_6 - R_5)}{2} - \frac{(R_4 - R_3) - (R_2 - R_1)}{2}\right]$$

$$= \frac{-R_1 + R_2 + R_3 - R_4 + R_5 - R_6 - R_7 + R_8}{4} \quad\dots\dots\dots\dots\dots\dots\dots.2.37$$

The second expression for e_{123} is obtained by multiplying the signs from columns for factor 1,2, and 3 in Table 2.6 and applying them to Ri. The denominator is once again 2^{k-1}. Three factors and higher interaction effects are also symmetric, for example, $e_{123} = e_{132} = e_{213}$ etc (Averill and Kelton,1996).

If the regression coefficients are named b_o, b_2, b_{12}, etc using Equation 2.23-2.26 the regression coefficient for 2^3 factorial are:

$$b_0 = \frac{R_1 + R_2 + R_3 + R_4 + R_5 + R_6 + R_7 + R_8}{8} \quad\dots\dots\dots\dots\dots.2.38$$

$$b_1 = \frac{-R_1 + R_2 - R_3 + R_4 - R_5 + R_6 - R_7 + R_8}{8} \quad\dots\dots\dots\dots\dots\ 2.39$$

and

$$b_{123} = \frac{-R_1 + R_2 + R_3 - R_4 + R_5 - R_6 - R_7 + R_8}{8} \quad\dots\dots\dots\dots\ 2.40$$

Or generally,

$$b_0 = \frac{1}{2^k}\sum_{i=1}^{2^k} Ri \quad\dots\dots\dots\dots\dots.2.41$$

$$b_j = \frac{1}{2^k}\sum_{i=1}^{2^k} (SiRi) \quad\dots\dots\dots\dots\dots.2.42$$

Where b_j = the coefficient of variables x_1, x_2, x_{12} etc,

Si = the sign along the column of the variable in question from the design matrix table and Ri = the response corresponding to Si.

32

2.3.8   A GENERAL 2^k DESIGN

Before any 2^k factorial could be analyzed G-test is used to check if the output factors have the maximum accuracy of the replication. It ascertains the possibility of carrying out regression analysis. The condition of homogeneity is:

$$G\left[\alpha,(r-1),N\right] > Gcal \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 2.43$$

Here G-cal $= \dfrac{Su^2\,max}{\sum\limits^{2^k} Su^2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 2.44$

The value of Su^2 max and $\sum\limits^{2^k} Su^2$ are gotten from table of response and their replicate.

The method of analysis that we have presented thus far may be generalized to the case of a 2^k factorial design. If the coded factors are x_1, x_2, x_{12} etc. The regression coefficient for response yi, may be calculated using the general formula

$$bo = \frac{1}{2^K}\sum_{i=1}^{2^K} yi \qquad \ldots\ldots\ldots\ldots\ldots\ldots 2.45$$

$$bj = \frac{1}{2^K}\sum_{i=1}^{2^K}(Siyi) \qquad \ldots\ldots\ldots\ldots\ldots\ldots 2.46$$

The significance of coefficient of the regression model could be tested using the individual F- test. We use F-test by rejecting the null hypothesis,

$$Ho: bj = 0 \qquad \ldots\ldots\ldots\ldots\ldots\ldots 2.47$$

when Fcal $>$ F$[\alpha$, df$_R$, N(r-1)] $\ldots\ldots\ldots\ldots\ldots\ldots$ 2.48

a coefficient is significant.

$$F_{cal} = \frac{MS_R}{MS_E} = \frac{\frac{SS_R}{df_R}}{\frac{SS_E}{N(r-1)}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 2.49$$

The sum of squares for any contrast can be computed from Equation 2.16, thus:

33

$$SS_R = SS_{bj} = \frac{(r.contrast)^2}{r.N} \qquad \dots\dots\dots\dots\dots\dots 2.50$$

The total sum of squares is found in the usual way by,

$$SS_T = \sum_{i=1}^{r} (y_{ri})^2 - \frac{\left(\sum_{i=1}^{r} y_r\right)^2}{r.N} \qquad \dots\dots\dots\dots\dots\dots 2.51$$

and

$$SS_E = SS_T - \sum SS_R \qquad \dots\dots\dots\dots\dots\dots 2.52$$

The table below present analysis of variance for a general 2^k design (Montgomery, 1976).

Table 2.8: General analysis of variance

Source of variation	Sum of squares (SS)	Degree of freedom (df)	Mean square (MS)	F-cal (MS/MS_E)
k main effects				
A	SS_A	1	$SS_A/1$	SS_A/MS_E
B	SS_B	1	$SS_B/1$	SS_B/MS_E
.	.	.	.	.
.	.	.	.	.
K	SS_k	1	$SS_k/1$	SS_k/MS_E
Two-factor interactions				
AB	SS_{AB}	1	$SS_{AB}/1$	SS_{AB}/MS_E
AC	SS_{AC}	1	$SS_{AC}/1$	SS_{AC}/MS_E
.	.	.	.	.
.	.	.	.	.
Jk	SS_{jk}	1	$SS_{jk}/1$	SS_{jk}/MS_E
Three-factor interactions				
ABC	SS_{ABC}	1	$SS_{ABC}/1$	SS_{ABC}/MS_E
ABD	SS_{ABD}	1	$SS_{ABD}/1$	SS_{ABD}/MS_E
.	.	.	.	.
.	.	.	.	.
Ijk	SSijk	1	SSijk/1	SSijk/MS_E
k-factor interactions				
ABC....k	$SS_{ABC....k}$	1	$SS_{ABC....k}/1$	$SS_{ABC....k}/MS_E$
Error	SS_E	$2^K(r-1)$	$SS_E/2^K(r-1)$	
Total	SS_T	$r.2^k - 1$		

Alternatively it could also be tested using the T-test because, T-test is a special form of F-test (Box et al,1978). A coefficient is significant if and only if

$$\text{T-cal} > \text{T} [\propto, \text{N} (r\text{-}1)] \qquad \dots\dots\dots\dots\dots\dots\dots 2.53$$

$$\text{T-cal} = t_j \quad \frac{|bj|}{Sbj} \qquad \dots\dots\dots\dots\dots\dots\dots 2.54$$

Also

$$Sbj = \sqrt{\frac{Su}{N.r}} \qquad \dots\dots\dots\dots\dots\dots\dots 2.55$$

2.3.9   A 2^4 DESIGN

Using the coded variables (factors) x_1, x_2,- x_k and the average response of a design. The design matrix Table for a 2^4 full factorial design is given below in Table 2.9.

Table 2.9:   Design matrix table for a 2^4 experiment

Factorials effects

Factor Comb.	x_0	x_1	x_2	x_3	x_4	x_{12}	x_{13}	x_{23}	x_{123}	x_{14}	x_{24}	x_{34}	x_{124}	x_{134}	x_{234}	x_{1234}	Y
1	+	-	-	-	-	+	+	+	-	+	+	+	-	-	-	+	Y_1
2	+	+	-	-	-	-	-	+	+	-	+	+	+	+	-	-	Y_2
3	+	-	+	-	-	-	+	-	+	+	-	+	+	-	+	-	Y_3
4	+	+	+	-	-	+	-	-	-	-	-	+	-	+	+	+	Y_4
5	+	-	-	+	-	+	-	-	+	+	+	-	-	+	+	-	Y_5
6	+	+	-	+	-	-	+	-	-	-	+	-	+	-	+	+	Y_6
7	+	-	+	+	-	-	-	+	-	+	-	-	+	+	-	+	Y_7
8	+	+	+	+	-	+	+	+	+	-	-	-	-	-	-	-	Y_8
9	+	-	-	-	+	+	+	+	-	-	-	-	+	+	+	-	Y_9
10	+	+	-	-	+	-	-	+	+	+	-	-	-	-	+	+	Y_{10}
11	+	-	+	-	+	-	+	-	+	-	+	-	+	+	-	+	Y_{11}
12	+	+	+	-	+	+	-	-	-	+	+	-	-	-	-	-	Y_{12}
13	+	-	-	+	+	+	+	-	+	-	-	+	+	-	-	+	Y_{13}
14	+	+	-	+	+	-	+	-	-	+	-	+	-	+	-	-	Y_{14}
15	+	-	+	+	+	-	-	+	-	-	+	+	-	-	+	-	Y_{15}
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	Y_{16}

2.3.10: 2^{k-p} FRACTIONAL FACTORIAL DESIGN

For a model with k factors, the designs of the previous section require at least

one simulation run of the 2^k possible combinations of factors .It is not at all difficult

to imaged a simulation model with as many as k= 11 factors, for which a full 2^k

factorial design would require at least 2048 simulation run. If we wanted to estimate the variances of the effects of this 2^{11} model with says n = 5 replications, at each design point, we will need 10,240 total replications. If a single replication of the model took say 1 minute of C P U time (which is modest amount of time for many complex real-world simulations), we would need over a full week of round-the-clock computing to complete the experiment.

Fractional factorial designs provide a way to get estimates of (for example) the main effects or low-order interaction of interest at a fraction of the experimental effort required by a full 2^k factorial design. These kinds of design are especially useful as a first step in experimentation when many factors are present and we want to screen out those factors which appear to be relatively unimportant without having to perform an excessive amount of simulation. This will save a bulk of the computing budget for a more intensive study of the important factors.

Basically, a 2^{k-p} fractional factorial design, is constructed by choosing a certain subset (of size 2^{k-p}) of the 2^k possible factor combinations and then running the simulation only for these chosen combinations. Thus only $1\backslash 2^p$ of the 2^k factor combinations are actually run. The important question of which 2^{k-p} combination to choose is a whole subject into itself, discussed at length in most books on experimental design (Box et al, 1978). This choice should obviously be made carefully and might depend on which main effects and interactions are of great interest.

38

2.3.11 VARIETAL TRIAL EXPERIMENT

In varietal trials, treatments like (a) different varieties of crop, (b) several feeds for animals,(c) different doses of drugs, etc are under investigation. In fact, different level of only one factor usually form the treatment in varietal trials. The main purpose of such experiment is to compare the treatment in all possible pairs.

2.3.12 BIO-ASSAY EXPERIMENT

The third type of experiment is the bio-assay. In one category of these experiments usually two preparations of drugs are taken, each at several doses. These doses form the treatments. The main comparison required for the assays are (i) a comparison giving the difference between the two preparations, (ii) a comparison representing the slope of the effects of the doses of each preparation on the doses.

2.4   FACTORS AFFECTING RATE OF CHEMICAL REACTION

The production of oxalic acid from the bark of eucalyptus camaldulensis, having a chemical reaction could be influenced by some of the factors that affect the rate of chemical reaction.

On the basis of elementary collision theory, the rate of a reaction would depend on the frequency of effective collisions between reactant particles.   (Lambert and Holderness, 1980).   Some of the important factors, which influence the rate of chemical reaction, include:

1.   Nature of reactants

2.   Concentration of reactant.

3.   Surface area of reactants

4.   Temperature of reaction mixture

39

Effect of the nature of reactants

When a piece of iron is place in dilute hydrochloric acid, there is a slow evolution of hydrogen gas, with a piece of zinc, hydrogen is evolved rapidly, and with a piece of gold, there is no evidence of reaction. Thus the rate of a chemical reaction is determined by the chemical nature of the reactant as different substances have different energy contents.

Effect of concentration of reactants.

Reactant particles will collide more often if they are crowded in a small space, i.e. frequency of collision is depending upon concentration. An increase (or decrease) in the concentration of the reactants will results in corresponding increase (or decrease) in effective collisions of the reactant hence in the reaction rate (Osei, 1990).

Effect of surface area of contact (Particle size)

This is a very important factor especially when one of the reactants is a solid, because only the particles on the surface of the solid are in contact and hence able to react with the other reacting particles. To bring about greater contact between the reacting particles, the exposed surface area of the solid reactants must be increased by subdividing or breaking the solid into smaller pieces. The greater the surface area of the reactant, the higher the rate of reaction.

Effect of temperature

When temperature of a reaction is increased, heat is supplied to the particles involved, in the reaction, thereby increasing the kinetic energy of the particles. If the particles travel at a greater speed when the temperature is increased then they will collide with one another at more frequent intervals and we could expect the reaction to proceed at a faster rate (Osei, 1990).

With respect to the above explanation, an increase in temperature within a definite limit will increase the rate of chemical reaction.

2.5   APPLICATION OF COMPUTER IN FACTORIAL DESIGN

One way common use of computer involves the development of models that simulate real-world systems. These models are coded in computer language, which take into account many, but presumably not all of the features of the real system. If the most important features of the real systems can be identified and built into the model without making it too complex and unwieldy, useful information about the real system may be obtained by working with the model. Being simpler, a model is usually more convenient to work with. Experiments can be run for various initial conditions, or values of key parameters, often much more conveniently than if the real systems were used .In fact, experiments are often run that would be impossible to do with the real system, because of hazardous conditions or for other equal good reasons. From such experiments it is hoped that a better understanding of certain aspects of the real system may be achieved. Often enough, the results obtained from a preliminary model suggest direction for improvement of the model, to explore previously unrecognized aspects of the real system. Because they represent efforts to identify the essential features of a real system, models have sometimes been described as "abstractions" of reality (Francis, 1983).

41

In computer-aided design, computers are used to perform certain functions in design, production, and manufacturing. This technology has moved into the direction of greater integration of design and manufacturing, two activities which have traditionally been treated as a distinct and separate functions in a production firm (Oguntoyimbo, 1993). On a more general note, computer aided design is utility that enables speedy processing of design procedures (Onifade, 2000).

In spite of their impressive capabilities, computers still have to be told exactly what to do, in a step-by-step fashion. The process of satisfactorily achieving the required level of detail is called problem analysis and it is the user's responsibility. This can be divided into several more or less distinct parts (Francis, 1983).

1.   **The problem must be thoroughly understood**: Surely if the user does not understand his or her own problem, there is little hope that the computer will understand it better. A careful examination of the inputs provided or the questions asked or type of output expected, perhaps even manual processing of simplified cases, may be useful in figuring out just what has to be done. This sometimes exposes fuzzy spots, which need clearing up.

2.   **A solution method is chosen, or developed:** When the problem does appear to be in good focus, a solution method has to be found, a path leading from what has been given to what is required. Often a path will be obvious. There may even be several likely prospects. When paths are abundant, the choice between them is made using such criteria as computer time needed or vulnerability to error. For other types of problems, no satisfactory solution method may be known, and the luxury of choice gives way to the necessity of invention or concession. That is either a method is developed, or the problem has to be simplified to the point where a solution becomes feasible.

42

3.   **The solution process is described step by step:** When a solution

method has been selected, it must be reduced to the level of detail that

computers understand. This is usually done in stages, beginning with a relative

crude listing of the various parts of the process, and gradually adding more

and more detail until the result has the step-by-step character of a program.

When this point is reached, the transition to a program is almost painless. A

description of the solution method involving some respectable level of detail is

called solution algorithm. Although of ancient origin, the word "algorithm"

has found its first extensive use in computer science.

4.   **The algorithm is programmed, and the program tested:** Algorithms can be

written in ordinary language, or using formal procedures that lie somewhere

between ordinary and programming languages. If this has been done with

sufficient detail, it will be relatively easy to convert the algorithm into a

program. If details have been postponed, they will have to be supplied in the

programming step. It is also true that some programming languages are more

accommodating than others, taking care of certain details that other languages

leave to the programmer. Once written, a program must be tested, or verified.

Even computer scientist soon learns how embarrassingly easy it is to make

mistakes. Typing and key punching errors are common place but are usually

found quickly with the machine's help. Diagnostic output is provided for this

purpose. Other programming errors may be harder to pinpoint, but detection

procedures do exist. Assuring a program's correctness is called program

verification, and the popular term for error removal is debugging.

5. **The algorithm is validated:** Even when a program correctly implements the steps of an algorithm, there is the question of whether the algorithm itself correctly solves the problem. To answer this, the user can run the program under a broad variety of conditions (test cases), and evaluate the output as best as he or she can (Francis, 1983).

CHAPTER THREE

3.0 DEVELOPMENT OF THE MODEL

The 2^k full factorial design was used to develop the source code. The flowchart for the source code is shown in Fig 3.1. The hand calculation is on Appendix A, while the program list is on Appendix B.

The design code consists of four files, extract.dat, extract.out, extract.pas and extract.exe. The file extract.pas is the Pascal source code. This is the actual code that causes the actions produced by the program when executed. It is a menu-oriented and user-friendly program that is easy to use with instruction and information appearing at each stage of execution of the program. The syntax of the content of this file follows the Pascal language requirements. This file is compiled into executable form so that the program can be run independent of the Pascal compiler.

Extract.exe is the executable file, generated from extract.pas which can be easily executed by calling the file i.e. C:\> extract ↵. Extract.dat is the default data file where the input data is stored, however provision is made in the program for using other data filenames. While extract.out is the file where the output generated by the program is stored. This file can be copied to a diskette or even printed after execution. Provision is also made in the program for using other output filenames.

3.1 MODELS TO PREDICT PERCENTAGE EXTRACT

In the course of this work, three 2^3 full fractional experiments with different factors were used. The comprehensive analysis of the results is given in chapter four. The three model equations used have the following general form:

$$y = \eta + \beta x_1 + \varnothing x_2 + \gamma x_3 + \theta x_1 x_2 + \mu x_1 x_3 + \delta x_2 x_3 + \varphi x_1 x_2 x_3 \quad ....3.1$$
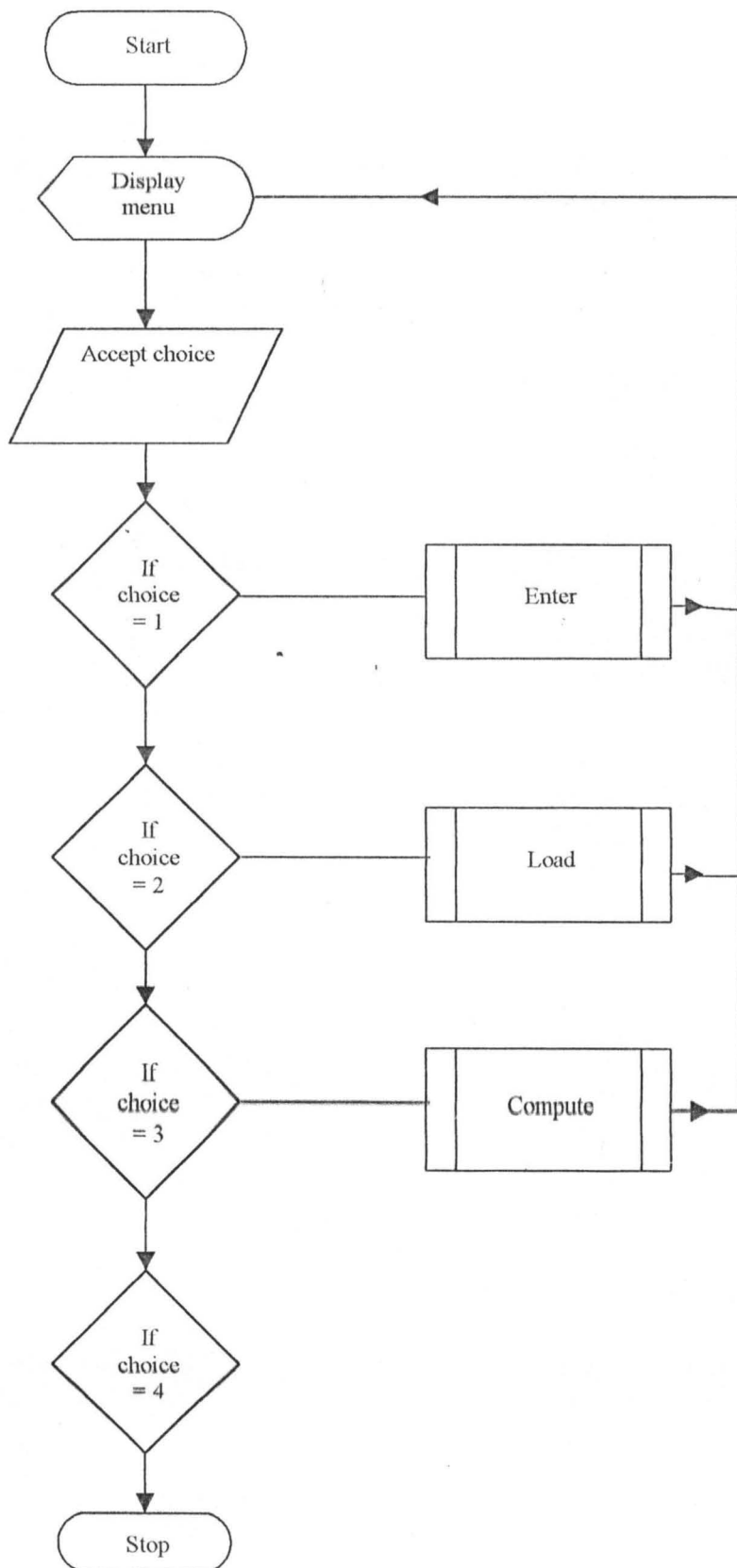
Fig. 3.1:   **Flow diagram for source code and running the Program**
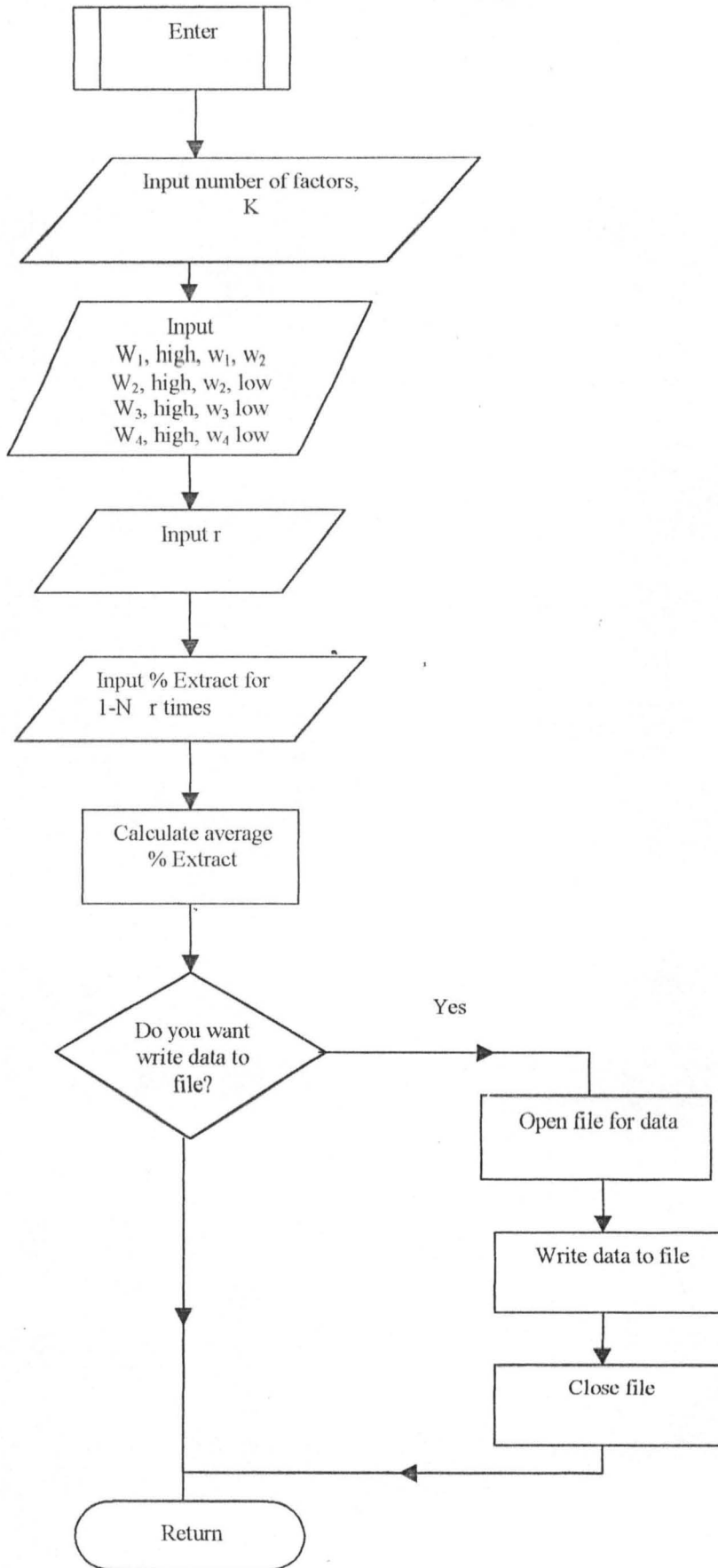
Fig 3.2:   **Flow diagram for sub-routine Enter**
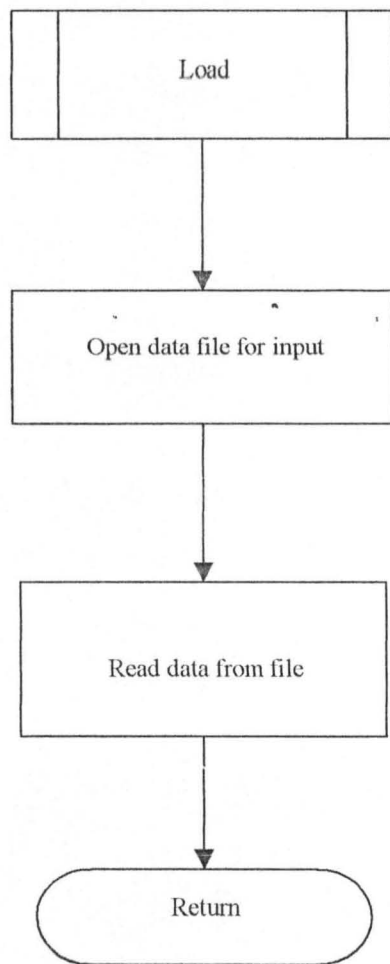
Fig. 3.3:   **Flow diagram for subroutine Load.**
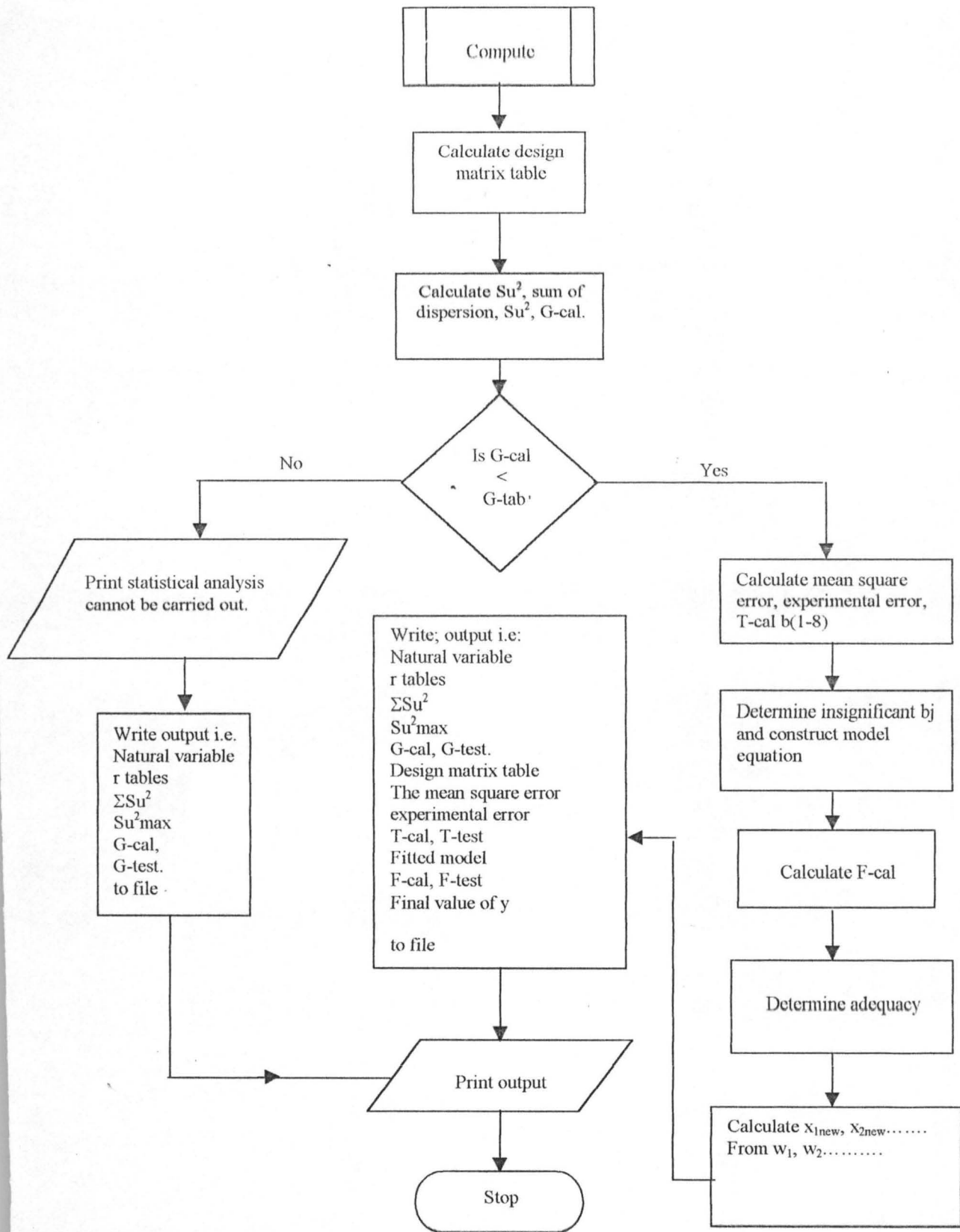
FIG. 3.4: **Flow diagram for sub-routine Compute.**

The flow diagram contains the following elements:

Compute

↓

Calculate design matrix table

↓

Calculate Su^2, sum of dispersion, Su^2, G-cal.

↓

Is G-cal < G-tab?

No →

Print statistical analysis cannot be carried out.

↓

Write output i.e.
Natural variable
r tables
ΣSu^2
$Su^2 max$
G-cal,
G-test.
to file

↓

Yes →

Calculate mean square error, experimental error, T-cal b(1-8)

↓

Determine insignificant bj and construct model equation

↓

Calculate F-cal

↓

Determine adequacy

↓

Calculate x_{1new}, x_{2new}·······
From w_1, w_2·········

Write; output i.e:
Natural variable
r tables
ΣSu^2
$Su^2 max$
G-cal, G-test.
Design matrix table
The mean square error
experimental error
T-cal, T-test
Fitted model
F-cal, F-test
Final value of y

to file

↓

Print output

↓

Stop

49

CHAPTER FOUR

4.0   SIMULATION OUTPUT

The prediction models are based on experimental results generated from the final year projects of Muhammad Adamu, Shafi'i Salihu and Bissallah Awwal on the production of oxalic acid from the bark of eucalyptus camaldulensis. The experimental procedure is given in Section 2.2.3 with Equation 2.8 summarising the reaction as follows:

$$CaC_2O_4 \; + \; H_2SO_4 \; \rightarrow \; (CO_2H)_2 \; + \; CaSO_4$$

Calcium	Sulphuric	Oxalic	Calcium
oxalate	acid	acid	sulphate
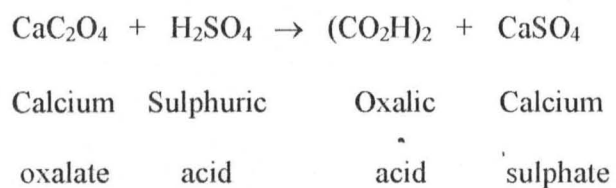
N.B.

Average size		Size range
0.3750mm	$\Rightarrow$	0.2500 - 0.5000mm
0.9250mm	$\Rightarrow$	0.8500 - 1.0000mm
1.2000mm	$\Rightarrow$	1.0000 - 1.4000mm

4.1   RESULTS FOR MODEL 1a

Constant (fixed) level parameters

Volume of H_2SO_4 = 250 mls

Concentration of H_2SO_4 = 4%

Reaction temperature = 70°C

Table 4.1:   Natural factors and their coded levels

Level of factors	Code	Mass	Size	Time	
High	+1	50.0000	1.2000	6.0000	0.0000
Low	-1	25.0000	0.9250	3.0000	0.0000

Table 4.2:   Experimental results

Factor combination $N=2^3$	Natural w_1 (mass) (g)	variables w_2 (size) (mm)	w_3 (time) (hrs.)	Y_i (% extract)
1	25	0.9250	3	3.2360
2	50	0.9250	3	2.4400
3	25	1.2000	3	3.1980
4	50	1.2000	3	0.9040
5	25	0.9250	6	1.8140
6	50	0.9250	6	2.2390
7	25	1.2000	6	1.5760
8	50	1.2000	6	0.7660

Table 4.3:   **Experimental results with replicates**

N	1	2	3	4	5	6	7	8
Yr1	3.3600	2.5480	3.1920	1.5260	1.7120	2.6140	1.5640	0.6640
Yr2	3.1120	2.3320	3.2040	0.2820	1.9160	1.8640	1.5880	0.8680
Y %	3.2360	2.4400	3.1980	0.9040	1.8140	2.2390	1.5760	0.7660
Yr1-Y	0.1240	0.1080	-0.0060	0.6220	-0.1020	0.3750	-0.0120	-0.1020
Yr2-Y	-0.1240	-0.1080	0.0060	-0.6220	0.1020	-0.3750	0.0120	0.1020
Sqr1	0.0154	0.0117	0.0000	0.3869	0.0104	0.1406	0.0001	0.0104
Sqr2	0.0154	0.0117	0.0000	0.3869	0.0104	0.1406	0.0001	0.0104
Su^2	0.0308	0.0233	0.0001	0.7738	0.0208	0.2813	0.0003	0.0208

The Sum of the dispersion   = 1.1511

The maximum Su^2 = 0.7738

G-Calculated   = 0.6722

G-Statistical Table $[\alpha,(r-1),N]$ = 0.6800

G-TEST:

It is possible to carry out regression

analysis, since G-stat > G-cal

Table 4.4:   Design matrix Table for a 2^ 3 full factorial design with the interactions

X0	X1	X2	X3	X1*X2	X1*X3	X2*X3	X1*X2*X3	Y
1	-1	-1	-1	1	1	1	-1	3.2360
1	1	-1	-1	-1	-1	1	1	2.4400
1	-1	1	-1	-1	1	-1	1	3.1980
1	1	1	-1	1	-1	-1	-1	0.9040
1	-1	-1	1	1	-1	-1	1	1.8140
1	1	-1	1	-1	1	-1	-1	2.2390
1	-1	1	1	-1	-1	1	-1	1.5760
1	1	1	1	1	1	1	1	0.7660

The mean square error   = 0.14388

The experimental error = 0.37932

T-TEST:

T-Statistical Table [α,N(r-1)] : 1.8600

The constant and the variable for the following

are insignificant since T-cal < T-table

b23 b123

The fitted model then becomes:

$Y = 2.022 \ -0.434x1 \ -0.411x2 \ -0.423x3 \ -0.342x1x2 + \ 0.338x1x3$

53

Table 4.5:   Table of calculated T-values

Regression coefficient	Estimated effect	Confidence interval	T values
b0	2.022	0.176	21.31837
b1	-0.434	0.176	4.58056
b2	-0.411	0.176	4.33011
b3	-0.423	0.176	4.45929
b12	-0.342	0.176	3.60249
b13	0.338	0.176	3.56558
b23	-0.017	0.176	0.18059
b123	0.033	0.176	0.34667

F-TEST :

F-Statistical Table $[\alpha, dFr, N(r-1)]$: 5.3200

The constant and the variable for the following

are insignificant since F-cal < F-table

b23 b123

Table 4.6:   Complete analysis of variance

Source of variation	Sum of squares (SS)	Degree of freedom (df)	Mean square	F-cal
b1	3.019	1	3.019	20.98149
b2	2.698	1	2.698	18.74984
b3	2.861	1	2.861	19.88524
b12	1.867	1	1.867	12.97795
b13	1.829	1	1.829	12.71339
b23	0.005	1	0.005	0.03261
b123	0.017	1	0.017	0.12018

Error (SSE)   =   1.15107

Total (SST) = 13.44752

N(r-1) = 8

Nr-1 = 15

SSE/(N(r-1) = 0.14388

Table 4.7: Experimental and calculated percentage extract

N	Y	Ycal	eu=Y-Ycal	eu^2=(Y-Ycal)^2
1	3.236	3.286	-0.050000	0.002500
2	2.440	2.424	0.015750	0.000248
3	3.198	3.148	0.050000	0.002500
4	0.904	0.920	-0.015750	0.000248
5	1.814	1.764	0.050000	0.002500
6	2.239	2.255	-0.015750	0.000248
7	1.576	1.626	-0.050000	0.002500
8	0.766	0.750	0.015750	0.000248

F-TEST (Fisher):

F-Calculated = 0.0254655

F-Statistical Table = 3.5800000

The fitted model is adequate

Since F-cal < F-table

Value of w1,x1 = 40.0000 0.2000

Value of w2,x2 = 1.2000 1.0000

Value of w3,x3 = 4.0000 -0.3333

Y = 2.022 -0.434x1 -0.411x2 -0.423x3 -0.342x1x2

+ 0.338x1x3

Final value of Y = 1.5742

55

Value of w1,x1   =   50.0000   1.0000

Value of w2,x2   =   0.9250   -1.0000

Value of w3,x3   =   3.0000   -1.0000

$Y = 2.022 \ -0.434x1 \ -0.411x2 \ -0.423x3 \ -0.342x1x2$

$+ \ 0.338x1x3$

Final value of Y =   2.4242

Value of w1,x1   =   25.0000   -1.0000

Value of w2,x2   =   0.9250   -1.0000

Value of w3,x3   =   6.0000   1.0000

$Y = \ \ 2.022 \ \ -0.434x1 \ \ -0.411x2 \ \ -0.423x3 \ \ -0.342x1x2$

$+ \ 0.338x1x3$

Final value of Y =   1.7640

4.2   RESULTS FOR MODEL 1b

Constant (fixed) level parameters

Volume of H_2SO_4 = 250 mls

Concentration of H_2SO_4 = 4%

Reaction temperature = 70°C

Table 4.8: Natural factors and their coded levels

Level of factors	Code	Mass	Size	Time	
High	+1	25.0000	0.9250	3.0000	0.0000
Low	−1	50.0000	1.2000	6.0000	0.0000

Table 4.9: Experimental results

Factor combination N=2^3	Natural w_1 (mass) (g)	variables w_2 (size) (mm)	w_3 (time) (hrs.)	Y_i (% extract)
1	50	1.2000	6	0.7660
2	25	1.2000	6	1.5760
3	50	0.9250	6	2.2390
4	25	0.9250	6	1.8140
5	50	1.2000	3	0.9040
6	25	1.2000	3	3.1980
7	50	0.9250	3	2.4400
8	25	0.9250	3	3.2360

Table 4.10:   Experimental results with replicates

N	1	2	3	4	5	6	7	8
Yr1	0.6640	1.5640	2.6140	1.7120	1.5260	3.1920	2.5480	3.3600
Yr2	0.8680	1.5880	1.8640	1.9160	0.2820	3.2040	2.3320	3.1120
Y %	0.7660	1.5760	2.2390	1.8140	0.9040	3.1980	2.4400	3.2360
Yr1-Y	-0.1020	-0.0120	0.3750	-0.1020	0.6220	-0.0060	0.1080	0.1240
Yr2-Y	0.1020	0.0120	-0.3750	0.1020	-0.6220	0.0060	-0.1080	-0.1240
Sqr1	0.0104	0.0001	0.1406	0.0104	0.3869	0.0000	0.0117	0.0154
Sqr2	0.0104	0.0001	0.1406	0.0104	0.3869	0.0000	0.0117	0.0154
Su^2	0.0208	0.0003	0.2813	0.0208	0.7738	0.0001	0.0233	0.0308

The Sum of the dispersion   = 1.1511

The maximum Su^2 = 0.7738

G-Calculated   = 0.6722

G-Statistical Table $[\alpha,(r-1),N]$ = 0.6800

G-TEST:

It is possible to carry out regression

analysis, since G-stat > G-cal

Table 4.11: Design matrix Table for a 2^ 3 full factorial design with the interactions

X0	X1	X2	X3	X1*X2	X1*X3	X2*X3	X1*X2*X3	Y
1	-1	-1	-1	1	1	1	-1	0.7760
1	1	-1	-1	-1	-1	1	1	1.5760
1	-1	1	-1	-1	1	-1	1	2.2390
1	1	1	-1	1	-1	-1	-1	1.1840
1	-1	-1	1	1	-1	-1	1	0.9040
1	1	-1	1	-1	1	-1	-1	3.1980
1	-1	1	1	-1	-1	1	-1	2.4400
1	1	1	1	1	1	1	1	3.2360

The mean square error   =   0.14388

The experimental error   =   0.37932

T-TEST:

T-Statistical Table $[\alpha, N(r-1)]$: 1.8600

The constant and the variable for the following

are insignificant since T-cal < T-table

b23 b123

The fitted model then becomes:

$Y = 2.022 + 0.434x1 + 0.411x2 + 0.423x3 - 0.342x1x2 + 0.338x1x3$

Table 4.12: **Table of calculated T-values**

Regression coefficient	Estimated effect	Confidence interval	T values
b0	2.022	0.176	21.31837
b1	0.434	0.176	4.58056
b2	0.411	0.176	4.33011
b3	0.423	0.176	4.45929
b12	-0.342	0.176	3.60249
b13	0.338	0.176	3.56558
b23	-0.017	0.176	0.18059
b123	-0.033	0.176	0.34667

F-TEST :

F-Statistical Table [α,dFr,N(r-1)] : 5.3200

The constant and the variable for the following

are insignificant since F-cal < F-table

b23 b123

Table 4.13: Complete analysis of variance

Source of variation	Sum of squares(SS)	Degree of freedom(df)	Mean square	F-cal
b1	3.019	1	3.019	20.98149
b2	2.698	1	2.698	18.74984
b3	2.861	1	2.861	19.88524
b12	1.867	1	1.867	12.97795
b13	1.829	1	1.829	12.71339
b23	0.005	1	0.005	0.03261
b123	0.017	1	0.017	0.12018

Error (SSE)   =   1.15107

Total (SST)   =   13.44752

$N(r-1)$ =   8

$Nr-1$ =   15

$SSE/(N(r-1)$ =   0.14388

Table 4.14: Experimental and calculated percentage extract

N	Y	Ycal	eu=Y-Ycal	eu^2=(Y-Ycal)^2
1	0.766	0.750	0.015750	0.000248
2	1.576	1.626	-0.050000	0.002500
3	2.239	2.255	-0.015750	0.000248
4	1.814	1.764	0.050000	0.002500
5	0.904	0.920	-0.015750	0.000248
6	3.198	3.148	0.050000	0.002500
7	2.440	2.424	0.015750	0.000248
8	3.236	3.286	-0.050000	0.002500

61

F-TEST (Fisher):

F-Calculated   =   0.0254655

F-Statistical Table   =   3.5800000

The fitted model is adequate

Since F-cal < F-table

Value of w1,x1   =   40.0000   -0.2000

Value of w2,x2   =   1.2000   -1.0000

Value of w3,x3   =   4.0000   0.3333

Y =   2.022   +   0.434x1 +   0.411x2 +   0.423x3   -0.342x1x2

+   0.338x1x3

Final value of Y   =   1.5742

Value of w1,x1   =   50.0000   -1.0000

Value of w2,x2   =   0.9250   1.0000

Value of w3,x3   =   3.0000   1.0000

Y =   2.022   +   0.434x1 +   0.411x2 +   0.423x3   -0.342x1x2

+   0.338x1x3

Final value of Y   =   2.4242

Value of w1,x1   =   25.0000   1.0000

Value of w2,x2   =   0.9250   1.0000

Value of w3,x3   =   6.0000   -1.0000

Y =   2.022   +   0.434x1 +   0.411x2 +   0.423x3   -0.342x1x2

+   0.338x1x3

Final value of Y   =   1.7640

4.3   RESULTS FOR MODEL 2

Constant (fixed) level parameters

Particle size = 0.375mm

Mass of bark = 50g

Reaction temperature = 70°C

Table 4.15:   Natural factors and their coded levels

Level of factors	Code	Conc.	Time	Vol.	
High	+1	4.0000	6.0000	750.0000	0.0000
Low	−1	2.0000	3.0000	250.0000	0.0000

Table 4.16:   Experimental results

Factor combination $N=2^3$	Natural w_1 (conc.) (%)	variables w_2 (time) (hrs.)	w_3 (vol.) (mls)	Y_i (% extract)
1	2	3	250	2.950
2	4	3	250	4.240
3	2	6	250	2.492
4	4	6	25	3.850
5	2	3	750	5.478
6	4	3	750	7.305
7	2	6	750	5.048
8	4	6	750	5.638

Table 4.17:   Experimental results with replicates

N	1	2	3	4	5	6	7	8
Yr1	3.0820	3.9220	1.9120	4.1700	5.2480	7.2020	5.1260	7.1820
Yr2	2.8200	4.1700	3.0700	3.5200	5.7060	7.4080	4.9700	4.0940
Y %	2.9510	4.0460	2.4910	3.8450	5.4770	7.3050	5.0480	5.6380
Yr1- Y	0.1310	-0.1240	-0.5790	0.3250	-0.2290	-0.1030	0.0780	1.5440
Yr2-Y	-0.1310	0.1240	0.5790	-0.3250	0.2290	0.1030	-0.0780	-1.5440
Sqr1	0.0172	0.0154	0.3352	0.1056	0.0524	0.0106	0.0061	2.3839
Sqr2	0.0172	0.0154	0.3352	0.1056	0.0524	0.0106	0.0061	2.3839
Su^2	0.0343	0.0308	0.6705	0.2112	0.1049	0.0212	0.0122	4.7679

The Sum of the dispersion   = 5.8529

The maximum Su^2   = 4.7679

G-Calculated   = 0.8146

G-Statistical Table [α,(r-1),N] = 0.6800

G-TEST:

It is not possible to carry out regression analysis, since G-stat < G-cal

4.4   RESULTS FOR MODEL 3

Constant (fixed) level parameters

Volume of H_2SO_4 = 250 mls

Mass of bark = 50g

Reaction temperature = 70°C

Table 4.18: Natural factors and their coded levels

Level of factors	Code	Conc.	Time	Size	
High	+1	4.0000	6.0000	1.2000	0.0000
Low	-1	2.0000	3.0000	0.9250	0.0000

Table 4.19: Experimental results

Factor combination $N=2^3$	Natural variable w_1 (conc.) (%)	w_2 (time) (hrs.)	w_3 (size) (mm)	Y_i (% extract)
1	2	3	0.9250	1.1800
2	4	3	0.9250	2.4400
3	2	6	0.9250	1.1970
4	4	6	0.9250	2.2460
5	2	3	1.2000	1.0930
6	4	3	1.2000	0.9000
7	2	6	1.2000	0.7930
8	4	6	1.2000	0.7700

Table 4.20: Experimental results with replicates

N	1	2	3	4	5	6	7	8
Yr1	1.1160	2.5480	1.2280	2.6140	0.9820	1.5280	1.1860	0.6640
Yr2	1.2540	2.3320	1.1660	1.8640	1.2020	0.2820	0.4000	0.8680
Y %	1.1850	2.4400	1.1970	2.2390	1.0920	0.9050	0.7930	0.7660
Yr1-Y	-0.0690	0.1080	0.0310	0.3750	-0.1100	0.6230	0.3930	-0.1020
Yr2-Y	0.0690	-0.1080	-0.0310	-0.3750	0.1100	-0.6230	-0.3930	0.1020
Sqr1	0.0048	0.0117	0.0010	0.1406	0.0121	0.3881	0.1544	0.0104
Sqr2	0.0048	0.0117	0.0010	0.1406	0.0121	0.3881	0.1544	0.0104
Su^2	0.0095	0.0233	0.0019	0.2813	0.0242	0.7763	0.3089	0.0208

The Sum of the dispersion   = 1.4462

The maximum Su^2   = 0.7763

G-Calculated   = 0.5368

G-Statistical Table [α,(r-1),N] = 0.6800

G-TEST:

It is possible to carry out regression

analysis, since G-stat > G-cal

Table 4.21:   **Design matrix Table for a 2^ 3 full factorial design with the**

   interactions

X0	X1	X2	X3	X1*X2	X1*X3	X2*X3	X1*X2*X3	Y
1	-1	-1	-1	1	1	1	-1	1.1850
1	1	-1	-1	-1	-1	1	1	2.4400
1	-1	1	-1	-1	1	-1	1	1.1970
1	1	1	-1	1	-1	-1	-1	2.2390
1	-1	-1	1	1	-1	-1	1	1.0920
1	1	-1	1	-1	1	-1	-1	0.9050
1	-1	1	1	-1	-1ʳ	1	-1	0.7930
1	1	1	1	1	1	1	1	0.7660

The mean square error   = 0.18077

The experimental error   = 0.42517

T-TEST:

T-Statistical Table $[\alpha, N(r-1)]$: 1.8600

The constant and the variable for the following

are insignificant since T-cal < T-table

b2 b12 b23 b123

The fitted model then becomes:

$Y = 1.327 + 0.260x1 - 0.438x3 - 0.314x1x3$

Table 4.22: Table of calculated T-values

Regression coefficient	Estimated effect	Confidence interval	T values
b0	1.327	0.198	12.48547
b1	0.260	0.198	2.44958
b2	-0.078	0.198	0.73734
b3	-0.438	0.198	4.12184
b12	-0.007	0.198	0.06233
b13	-0.314	0.198	2.95291
b23	-0.031	0.198	0.29282
b123	0.047	0.198	0.43864

F-TEST :

F-Statistical Table $[\alpha, dFr, N(r-1)]$: 5.3200

The constant and the variable for the following

are insignificant since F-cal < F-table

b2 b12 b23 b123

Table 4.23: Complete analysis of variance

Source of variation	Sum of squares(SS)	Degree of freedom(df)	Mean square	F-cal
b1	1.085	1	1.085	6.00046
b2	0.098	1	0.098	0.54368
b3	3.071	1	3.071	16.98955
b12	0.001	1	0.001	0.00388
b13	1.576	1	1.576	8.71965
b23	0.016	1	0.016	0.08574
b123	0.035	1	0.035	0.19241

Error (SSE)   =   1.44619

Total (SST)   =   7.32771

$N(r-1)$ =   8

Nr-1   =   15

SSE/(N(r-1)   =   0.18077

68

Table 4.24:   Experimental and calculated percentage extract

N	Y	Ycal	eu=Y-Ycal	eu^2=(Y-Ycal)^2
1	1.185	1.191	-0.006000	0.000036
2	2.440	2.340	0.100500	0.010100
3	1.197	1.191	0.006000	0.000036
4	2.239	2.340	-0.100500	0.010100
5	1.092	0.943	0.149500	0.022350
6	0.905	0.835	0.069500	0.004830
7	0.793	0.943	-0.149500	0.022350
8	0.766	0.835	-0.069500	0.004830

F-TEST (Fisher):

F-Calculated   = 0.2064285

F-Statistical Table   = 3.5800000

The fitted model is adequate

Since F-cal < F-table

Value of w1,x1   =   4.0000   1.0000

Value of w2,x2   =   3.0000   -1.0000

Value of w3,x3   =   0.9250   -1.0000

Y =   1.327 + 0.260x1 -0.438x3 -0.314x1x3

Final value of Y =   2.3395

Value of w1,x1   =   3.0000   0.0000

Value of w2,x2   =   3.0000   -1.0000

Value of w3,x3   =   1.2000   1.0000

Y =   1.327 + 0.260x1 -0.438x3 -0.314x1x3

Final value of Y =   0.8890

69

```
Value of w1,x1    =    2.0000   -1.0000
Value of w2,x2    =    4.0000   -0.3333
Value of w3,x3    =    0.9250   -1.0000
Y =   1.327  +  0.260x1  -0.438x3  -0.314x1x3
Final value of Y  =    1.1910
```

CHAPTER FIVE

5.0   DISCUSSION OF RESULTS

The prediction model is developed based on the hand calculation of 2^3 full factorial experiment of Appendix A. But the program was also written to carry out the analysis and prediction for 2^2, 2^3 and 2^4 full factorial experiments. However, since this work is based on the results for a 2^3 full factorial, the results presented are for 2^3 experiments.

Looking through the results for model 1b as presented in Tables 4.8 – 4.14 and comparing it with the results for the hand calculation of Appendix A, the model for both case is:

$$y = 2.022 + 0.434x_1 + 0.411x_2 + 0.423x_3 - 0.342x_1x_2 + 0.338x_1x_3$$

The value for other parameters e.g. G-cal, F-cal, e.t.c. are the same in both the manual and the program output are the same, except for small errors resulting from round offs (approximation).

Averill and Kelton (1996), reported that there is no general prescription for the choice of level. Model 1 was therefore analyzed using two versions. One version differs from the other in the choice of level of the factors. This is done so that we can investigate the effect of choice of levels. In Table 4.1 which is for the levels of model 1a, the high numerical value is taken for high level while the low numerical value is considered as low level. On the other hand, Table 4.8 shows the choice of level for model 1b where low numerical value is chosen as high level and high numerical value as low level. From Table 4.2, the average percentage extract for the eight combination of model 1a are: 3.2360, 2.4400, 3.1980, 0.9040, 1.8140, 2.2390, 1.5760 and 0.7660. And in Table 4.9, due to change in choice of level, the percentage extract

71

for model 1b are arranged in a reverse order thus: 0.7660, 1.5760, 2.2390, 1.8140, 0.9040, 3.1980, 2.440 and 3.2360. This is so because all the books on experimental design especially Montgomery (1976), said that the responses must be arranged in accordance with the design matrix table (standard order).

After the necessary test and calculations, the model equation for model 1a is:

$$y = 2.022 - 0.434x_1 - 0.411x_2 - 0.423x_3 - 0.342x_1x_2 + 0.338x_1x_3 \ldots\ldots\ldots 5.1$$

It can be noticed that the variables x_2x_3, $x_1x_2x_3$ and their coefficients did not appear in the model equation because they failed the T-test and F-test. The effect of raising a natural variable from its low level to high level is given by the coefficient of the variable (Averill and Kelton, 1996). *Negative (-) sign implies a decrease while positive (+) sign implies an increase. In the model for yield of sugar in Section 2.3.4, Cox and Cochran (1957) concluded thus: "the application of nitrogen increased the yield of sugar by 7.4cwt, while 11in ploughing increased the yield by 1.9cwt as compared with 7in ploughing". From Equation 5.1, raising the mass of eucalyptus camaldulensis from 25g to 50g decreases the percentage extract by 0.4340. Raising the average size from 0.9250 to 1.2000mm decreases the percentage extract by 0.4110, while raising the time from 3 hours to 6 hours decrease the percentage extract by 0.4230. In the inventory model of Example 2.1, Averill and Kelton (1996) also concluded that the smaller value of A and B would be preferable, since lower monthly costs are desired. The same reasoning is applicable here, since the higher values of percentage extract is desired, it appears that smaller values of the mass, size and time are preferable.

72

The Equation resulting from model 1b is given below:

$$y = 2.022 + 0.434x_1 + 0.411x_2 + 0.423x_3 - 0.342x_1x_2 + 0.338x_1x_3 \ldots\ldots\ldots 5.2$$

Here raising the mass from 50 (low level) to 25 (high level) increase the percentage extract by 0.4340 while raising the size and time from their low level to high level increases the percentage extract by 0.4110 and 0.4230 respectively. Again since higher of the percentage extract are desired and an increase is obtained by moving in the direction of the low numerical values, the lower mass, size and time are desired.

Muhammed (1998), Bisallah (1998), and Shafii (1998) concluded from their calculation of average percentage extract that the percentage extract of oxalic acid decreases with increase in size and time. Though their basis for this conclusion is not sound, the conclusion was correct. It agrees with our finding base on factorial experiment analysis which is a well-established method of analysis. They could not make any deduction for the mass because an individual was given a fraction of factorial experiment to handle in that regard. But with factorial analysis we concluded that the percentage extract decreases with increase in mass over the range of level used.

To justify our finding, we refer to our knowledge of chemistry and extraction. For the size, we know that the smaller the particle size, the greater the surface area of reactant and the higher the rate of reaction (Osei, 2000). The higher the rate of reaction, the more the product formed after a given time. As far as time is concerned, we know that any reaction has the optimum time for it to come to completion. Any moment after that time, there may be side reaction, which might results in decrease in the desired product. From the experimental procedure, we learnt that there is always unreacted H_2SO_4 in the mixture. When all the CaC_2O_4 are converted to $(CO_2H)_2$ after a given time, if the reaction is not stopped, since there is continuous heating the excess H_2SO_4 dehydrates some of the oxalic acid formed to equimolar mixture of carbon monoxide and dioxide (Dazeley, 1979). This is summarized by Equation 2.6 thus:

73

$$(CO_2H)_2 \xrightarrow[heat]{conc.H_2SO_4} CO_{(g)} + CO_{2(g)} + H_2O_{(g)}$$
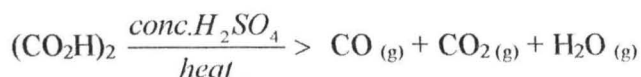
This implies that smaller time between 3 and 6hours is most suitable for this reaction. For the mass, since we are using a fixed volume (250mls) of H_2SO_4 for both 50g and 25g, the decrease recorded shows that smaller mass between 25g and 50g is required for the volume of acid used. That is to say only small amount of CaC_2O_4 in the bulk of the bark is required to react with 250mls of H_2SO_4, leaving the remaining oxalate unreacted. This is why a decrease is recorded in the percentage extract of oxalic acid obtained with increase in mass form 25g to 50g.

Although there is differences in signs of coefficient of x_1, x_2 and x_3 in model 1a and model 1b, it can be seen from the above discussion that they represent the same thing. In addition when the same values of variables are used in both models to predict percentage extract, the results were obtained. For instance, when mass of 40g, 1.2000mm average size and time of 4 hours was used to predict the percentage extract in both models, the predicted percentage extract is 1.5742.

The results for model 2 are given in Table 4.15 - 4.17 of Section 4.3. The G-calculated is 0.8146 while the G-Table is 0.6800. Since G-table < G-cal it is not possible to carryout regression analysis based on 5% level of significance ($\propto = 0.05$). Therefore the prediction model cannot be fitted.

Table 4.18 to 4.24 of Section 4.4 presents the results for model 3. The fitted model equation is:

$$y = 1.327 + 0.260x_1 - 0.438x_3 - 0.314x_1x_3 \ldots\ldots\ldots\ldots 5.3$$

Raising the concentration from 2 to 4% increases the percentage extract by 0.2600, and raising the average size from 0.9250 to 1.2000mm decreases the percentage extract by 0.4380. Here it will be noticed that the coded variable x_2 for time does not appear in the model equation. This is because the T-calculated for x_2, 0.7373 is less than 1.8600 which is the T-table. Since the T-cal < T-table, b_2 and x_2 are statistically insignificant. The variable also failed the F-test, all based on 5% level

74

of significance. This does not mean that time is useless in the experiment, because in Table 4.22 the estimated effect of time is 0.0780. Showing that raising the time from 3 to 6 hours decrease the percentage extract by 0.0780. Therefore, the higher value of concentration is desired for high percentage extract while the lower values of time and particle size are desired.

From our knowledge of factors affecting rate of chemical reaction, we know that the higher the concentration of reactant, the faster the rate of reaction (Lambert and Holderness, 1980). This implies that there would be more product and percentage extract for a specified reaction time. Since our percentage extract increases with increase in concentration, it agrees with this theory. As for the time ands size, the explanation is the same with that of model 1.

According to Averill and Kêlton (1996), given this regression model:
$$R(A, B) = \eta + \beta x_A + \varnothing x_B + \varphi x_A x_B \dots\dots\dots\dots\dots\dots\dots\dots 5.4$$ We could estimate η, β, $\varnothing$ and φ by ordinarily square regression and use the fitted model to forecast (predict) the value of $R(A,B)$ at combinations of A and B where no simulation has been done. The same way in the last page of Section 4.4 where concentration of 4%, time of 3hours and average size of 0.9250mm were used to predict the percentage extract of oxalic acid, the final value of percentage extract gotten is 2.3400. The coded variables x_1, x_2 and x_3 calculated are 1, -1 and -1 respectively. These correspond to the number 2-factor combination of Table 4.21. The calculated percentage extract for this factor combination in Table 4.24 is 2.3400, which is the same as the predicted percentage extract. This is one way of showing the level of accuracy of the prediction models. Another way is to look at the difference between the experimental percentage extract and the calculated percentage extract. In Table 4.7 and 4.14 for models 1a and b the highest difference is $\pm$ 0.0500. The difference is given in Table 4.24 for model 3 as $\pm$ 0.1495. For both models the difference is small as compared with their confidence limit of $\pm$ 0.1760 for model 1 and $\pm$ 0.1980 for model 3. Finally the experimental error for model 1 is 0.3793 while for model 3 it is 0.4251. Showing that model 1 is more accurate than model 3.

Generally, factorial experiment provide opportunity to study not only the individual effect of each factor (variable), but also their interaction (Das and Giri, 1979). Because when experiments are conducted factor by factor changing the level of one factor at a time and keeping other factors at constant level, the effect of interaction cannot be investigated.

CHAPTER SIX

6.0 **CONCLUSION AND RECOMMENDATION**

6.1 **CONCLUSION**

For model 1, the discussion shows that the choice of level of variables does not affect the model. Therefore, one can decide to choose the level any how, but it is advisable to choose high numerical value as high level and low for low to avoid confusion.

From the discussion above, we can see that for model 1a and 1b the smaller values of the mass, size and time gives higher the percentage extract. Therefore, the optimum percentage extract is gotten in the region of the smaller mass, size and time. However optimization technique will be needed to get the exact optimum values. In model 3, the higher the concentration, the higher the percentage extract of Oxalic acid. While the lower the time and size, the higher the percentage extract of Oxalic Acid obtained. It can also be concluded from the calculated experimental error that model 1 (0.3793) is more accurate than model 3 (0.4251).

6.2 **RECOMMENDATION**

The supervisor is the planner of experiments and it is because he planned and supervised the 2^3 full factorial experiment from the scratch that we are able to use it for further analysis. I suggest that he plans experiments for higher factorials, especially 2^4 and the fractional experiments i.e. 2^{k-p}.

I also recommend that further work be done on these results to see how possible it will be to use it for optimization. There is no problem about these because from Section 2.3.1 of the literature review we learnt that experimentation is usually an iterative process. With one experiment answering some questions and simultaneously posing others.

77

REFERENCES

Anderson V. L. and McLean R. A., 1974, "Design of Experiments: A Realistic Approach" , pp. 500, Marcel Dekkar Inc., New York.

Anonymous, 1982, "The New Encyclopedia Britannia", 15th ed. Vol. VII, pp. 645, Encyclopedia Britannica incorporated, Chicago.

Anonymous, 1982, "The New Encyclopedia Britannica" 15th ed. Vol. III pp. 986, Encyclopedia Britannica incorporated, Chicago.

Anonymous, 1982, "McGraw Hill Encyclopedia of Science and Technology", 5th ed. Vol. 9, pp. 710, McGraw Hill Book Company, New York.

Anonymous, 1994, "Encyclopedia Americana" International ed. Vol. pp. 648 – 649, Grolier Incorporated, Danbury. Reference:

Anonymous, 1994, "Encyclopedia Americana" International ed. Vol. 21 pp. 69, Grolier Incorporated, Danbury.

Arene E. O. and Kitwood T.M 1979, "An Introduction to Chemistry of Carbon Compounds", pp240-244, Longman, England.

Averill, M. L., and Kelton, W. D., 1976, "Simulation Modeling and Analysis", pp. 372-380, McGraw Hill Book Company, New York.

Bisallah Awwal, 1998. "Production of oxalic acid from the bark of Eucalyptus camaldulensis": (A final year B.Tech. project submitted to the Department of Chemical Engineering, Federal University of Technology, Minna, Nigeria), (Unpublished).

Box, G. E. P. Hunter, W. G. and Hunter, J. S., 1978, "Statistics for experimenters: An Introduction to Design Analysis and Model Building" pp. 319 – 322, Wiley and Sons Incorporated, New York.

Buckingham, J., 1982, "Dictionary of Organic Compounds" 5th ed. Vol. 4. pp. 4425, Chapman and Hall, New York.

Cochran, W. G. and Cox, G. M., 1957, "Experimental Design" 2nd ed. Pp. 148 – 152 John Wiley and Sons Incorporated, London.

Das, M. N. and Giri, N. C., 1979, "Design and Analysis of experiments" pp. 78 – 79, Wiley Eastern Limited, New Delhi.

Dazeley, G. H., 1979 "Organic Chemistry" pp. 288-289 Cambridge University Press, London.

Fisher R. A.; 1966, "The Design of experiments", 8th ed. pp. 527, Hafner Publishing Company, New York.

Francis, S., 1983, "Schaum's Outline of Theory and Problems of Computers and Programming" International ed. pp. 274, McGraw Hill Book Company, Singapore.

Guttman, L., Wilks, S. S. and Hunter, J. S., 1965, "Introductory Engineering Statistics." 3rd ed. John Wiley and Sons Incorporated, New York.

Hill, A. F. 1979, "Economic Botany", 2nd ed. pp. 79-81, Tata Mc. Graw Hill Publishing company Ltd. New Delhi.

Irwin, P. T., 1981, "Plant Physiology" pp. 310, Addison Wesley Publishing Company, London.

Keay, R. W. J., 1989, "Trees of Nigeria" pp. 78, Oxford Science Publication, New York.

Kempthrone, O. 1952, "The Design and Analysis of Experiments", pp. 432, Wiley, New York.

Lambert, J., and Holderness, A., 1980, A New Certificate Chemistry 5th ed. pp. 192-196, Heinman Educational Books, London.

Linton, C. F., 1965, "Elementary Applied Statistics", John Wiley and Sons Incorporated, New York.

Montgomery, C. D., 1976, "Design and Analysis of Experiments", pp. 181-191, John Wiley and Sons Incorporated, New York.

Muhammed, A, 1998, "Production of oxalic acid from the bark of Eucalyptus camaldulensis": (A final year B.Tech. project submitted to the Department of Chemical Engineering, Federal University of Technology, Minna, Nigeria), (Unpublished).

Oguntoyimbo, S., 1993, "Computer Aided Design in Nitel R and D", Journal, Vol. No. 2. pp 41.

Onifade, K. R. 2000 "Computer Aided Design Module for Multicomponet Distillation Column". A. U. Journal of Technology, 4(1) 26 – 38.

Osei Y. A. 1990, New School Chemistry, Senior Secondary Science. New ed. pp. 212 – 223. Africana – F.E.P. Publishers Limited, Nigeria.

Perry, R. H., and Green, B. W., 1984, "Chemical Engineering Handbook", 6th ed. McGraw Hill, New York.

Shafi'i, S., 1998. "Production of oxalic acid from the bark of Eucalyptus camaldulensis": (A final year B.Tech. project submitted to the Department of Chemical Engineering, Federal University of Technology, Minna, Nigeria), (Unpublished).

HAND CALCULATION

The required model equation fore predicting the yield is given as

$$y = b_o + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3$$

Table 1:   Natural factors and their coded levels

Levels of factors	code	w_1	w_2	w_3
High	+1	25	0.925	3
Low	-1	50	1.200	6

Table 2:   Experimental results

N	w_1 (Mass)	w_2 (Size)	w_3 (Time)	Y (% Extract)
1	50	1.200	6	0.766
2	25	1.200	6	1.576
3	50	0.925	6	2.239
4	25	0.925	6	1.814
5	50	1.200	3	0.904
6	25	1.200	3	3.198
7	50	0.925	3	2.440
8	25	0.925	3	3.236

Table 3: Design matrix table for a 2^3 full factorial with the interaction of coded factors.

N	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	$X_1X_2X_3$	Y
1	+	-	-	-	+	+	+	-	0.766
2	+	+	-	-	-	-	+	+	1.576
3	+	-	+	-	-	+	-	+	2.239
4	+	+	+	-	+	-	-	-	1.814
5	+	-	-	+	+	-	-	+	0.904
6	+	+	-	+	-	+	-	-	3.198
7	+	-	+	+	-	-	+	-	2.440
8	+	+	+	+	+	+	+	+	3.236

Table 4: Experimental results with their replicates.

Replicates (r)

N	Y_{r1}	Y_{r2}	Y_i	$Y_{r1}-Y_i$	$Y_{r2}-Y_i$	$(Y_{r1}-Y_i)^2$	$(Y_{r2}-Y_i)^2$	$S\mu^2$
1	0.664	0.868	0.766	-0.102	0.102	0.0104	0.0104	0.0208
2	1.564	1.588	1.576	-0.012	0.012	0.0001	0.0001	0.0002
3	2.614	1.864	2.239	0.375	-0.375	0.1406	0.1406	0.2812
4	1.712	1.916	1.814	-0.102	0.102	0.0104	0.0104	0.0208
5	1.526	0.282	0.904	0.622	-0.622	0.3869	0.3869	0.7738
6	3.192	3.204	3.198	-0.006	0.006	0.0000	0.0000	0.0000
7	2.548	2.332	2.440	0.108	-0.108	0.0117	0.0117	0.0234
8	3.360	3.112	3.236	0.124	-0.124	0.0154	0.0154	0.0308

i.   Mean $Y = \dfrac{1}{r}\sum yr$ 2

ii.   The dispersion of replicated observations is given by

$$Su^2 = \frac{1}{r-1}\sum (y_r - y_i)^2 \quad ....3$$

iii.   The sum of the dispersion $\sum Su^2 = 1.1511$

iv. Su^2, max $= 0.7738$

v.   The homogeity of the dispersion was determined using Cochran criterion of

Calculated G-value

i.e.   G-cal $= \dfrac{Su^2 \, max}{\sum Su^2}$ 4

G-cal $= \dfrac{0.7738}{1.1511} = 0.6722$

vi.   G-Test was used to check if the output factors of the replication have maximum

accuracy of replication.

G-Test.

The condition of homogeneity is:

G [α, (r-1), N] > G-cal 5

from statistical table

G(0.05, 1, 8) = 0.0680

Since 0.680 > 0.6722

It is possible to carry out regression analysis.

83

Analysis:

1. The mean square error was determined by $Su^2 = \dfrac{1}{N}\sum Su^2$ 6

$$Su^2 = \dfrac{1}{8}(1.1511)$$

$$= 0.14389$$

2. The experimental error was given as:

$$Su = \sqrt{Su^2} = \sqrt{0.14389} = 0.37933$$

3. The regression coefficients are given by :

$$b_0 = \dfrac{1}{N}\sum_{i=1}(yi)\dots\dots\dots\dots\dots.7$$

$$bj = \dfrac{1}{N}\sum_{i=1}^{N}(siyi)\dots\dots\dots\dots.8$$

Therefore,

$$b_0 = \dfrac{1}{8}(0.766 + 1.576 + 2.239 + 1.814 + 0.904 + 3.198 + 2.440 + 3.236)$$

$$= \dfrac{16.173}{8}$$

$$= 2.022$$

$$b_1 = \dfrac{1}{8}(-0.766 + 1.576 - 2.239 + 1.814 - 0.904 + 3.198 - 2.440 + 3.236)$$

$$= \dfrac{3.475}{8}$$

$$= 0.434$$

$$b_2 = \dfrac{1}{8}(-0.766 - 1.576 + 2.239 + 1.814 - 0.904 - 3.198 + 2.440 + 3.236)$$

$$= \frac{3.285}{8}$$

$$= 0.411$$

$$b_3 = \frac{1}{8}(-0.766 - 1.576 - 2.239 - 1.814 + 0.904 + 3.198 + 2.440 + 3.236)$$

$$= \frac{3.383}{8}$$

$$= 0.423$$

$$b_{12} = \frac{1}{8}(0.766 - 1.576 - 2.239 + 1.814 + 0.904 - 3.198 - 2.440 + 3.236)$$

$$= \frac{-2.733}{8}$$

$$= -0.342$$

$$b_{13} = \frac{1}{8}(0.766 - 1.576 + 2.239 - 1.814 - 0.904 + 3.198 - 2.440 + 3.236)$$

$$= \frac{2.705}{8}$$

$$= 0.338$$

$$b_{23} = \frac{1}{8}(0.766 + 1.576 - 2.239 - 1.814 - 0.904 - 3.198 + 2.440 + 3.236)$$

$$= \frac{-0.137}{8}$$

$$= -0.017$$

$$b_{123} = \frac{1}{8}(-0.766 + 1.576 + 2.239 - 1.814 + 0.904 - 3.198 - 2.440 + 3.236)$$

$$= \frac{-0.263}{8}$$

$$= -0.033$$

Testing significance of the regression coefficients.

T-test.

A coefficient is considered significant if and only if

$$T\text{-cal} > T\ [\alpha,\ N\ (r-1)] \dots\dots\dots\dots\dots 9$$

$$T\text{-table} = T[\alpha,\ N(r-1)] \dots\dots\dots\dots 10$$

from statistical table

$$T(0.05,\ 8) \quad = \quad 1.860$$

$$\text{i.e. T-table} \quad = \quad 1.860$$

$$Sb = \frac{Su}{\sqrt{N.r}} \dots\dots\dots\dots\dots 11$$

$$= \quad \frac{0.37933}{\sqrt{16}} = \frac{0.37933}{4}$$

$$= \quad 0.09483$$

$$Tj \quad = \quad \frac{|bj|}{sb} \dots\dots\dots\dots\dots 12$$

$$T_0 \quad = \quad \frac{bo}{Sb} = \frac{2.022}{0.09483} = 21.322$$

$$T_1 \quad = \quad \frac{b_1}{Sb} = \frac{0.434}{0.09483} = 4.577$$

$$T_2 \quad = \quad \frac{b_2}{Sb} = \frac{0.411}{0.09483} = 4.334$$

$$T_3 \quad = \quad \frac{b_3}{Sb} = \frac{0.423}{0.09483} = 4.461$$

$$T_{12} \quad = \quad \frac{b_2}{Sb} = \frac{0.349}{0.09483} = 3.606$$

$$T_{13} = \frac{b_{13}}{Sb} = \frac{0.339}{0.09483} = 3.564$$

$$T_{23} = \frac{0.017}{0.09483} = 0.179$$

$$T_{123} = \frac{0.033}{0.09483} = 0.348$$

b_0, b_1, b_2, b_3, b_{12}, and b_{13} are significant since T-cal > T-table. But b_{23} and b_{123} are not significant since T-cal < T-table.

∴ The filted model become:

$$y = 2.022 + 0.434x_1 + 0.411x_2 + 0.423x_3 - 0.342x_1x_2 + 0.338x_1x_3 \ ....13$$

Confidence intervals for the regression coefficients with confidence, α are of the general terms,

$$b_j \pm T\ [\alpha, N(r-1)]Sb.....................14$$

Confidence interval Δb_j = T-table x Sb.........15

$$= 1.860 \times 0.0948$$

$$= 0.176$$

Table 5: **The estimated effects, confidence interval and calculated T-values.**

Regression estimated confidence T-values

Regression Coefficients	Estimated Effect	Confidence interval.	T-values
b_0	2.022	0.176	21.322
b_1	0.434	0.176	4.577
b_2	0.411	0.176	4.334
b_3	0.423	0.176	4.461
b_{12}	0.342	0.176	3.606
b_{13}	0338	0.176	3.564
b_{23}	0.017	0.176	0.179
b_{123}	0.033	0.176	0.348

87

$$y = 2.022 + 0.434x_1 + 0.411x_2 + 0.423x_3 - 0.342x_1x_2 + 0.338x_1x_3$$

viii. The adequacy of the model was evaluated by the null hypothesis,
$$H_0: \ b_j = 0 \ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \ 16$$

on the individual regression coefficients. The analysis of variance is very useful in confirming the significance of the coefficients. In the 2^k factorial design with r replicates, the regression sum of squares for any effect is:

$$SS_R = \frac{(r.contrast)^2}{r.N} \ \dots\dots\dots\dots\dots\dots\dots.17$$

and has the degree of freedom ($df_R = 1$).

$$SS_T = \sum (y_{ri})^2 - \frac{(\sum y_{ri})^2}{r.N} \ \dots\dots\dots\dots\dots.18$$

The error sum of squares was given by

$$SS_E = SS_T - \sum SS_R \ \dots\dots\dots\dots\dots\dots\dots.19$$

$$SS_E = SS_T - \sum (SS_{bj}) \ \dots\dots\dots\dots\dots\dots.20$$

Testing the significance of each coefficient was carried out by F-test

$$F_{cal} = \frac{MS_R}{MS_E} = \frac{\frac{SS_R}{df_R}}{\frac{SS_E}{N(r-1)}} \ \dots\dots\dots\dots.21$$

The calculated F-values were compared with the appropriate critical table value. The null hypothesis is rejected using

$$F_{cal} > F[\alpha, df_R, N(r-1)]\dots\dots\dots.22$$

$$contrast = \sum (x_j y_i) \dots\dots\dots.23$$

Calculations:

$$SS_{b1} = \frac{1}{2x8}[2(-0.766 + 1.576 - 2.239 + 1.814 - 0.904 + 3.198 - 2.440 + 3.236)]^2 = \frac{(6.95)^2}{16} = 3.0189$$

$$SS_{b2} = \frac{1}{2x8}\left[2(-0.766 - 1.576 + 2.239 + 1.814 - 0.904 - 3.198 + 2.440 + 3.236)\right]^2 = \frac{(6.57)^2}{16} = 2.6978$$

$$SS_{b3} = \frac{1}{2x8}\left[2(-0.766 - 1.576 - 2.239 - 1.814 + 0.904 + 3.198 + 2.440 + 3.236)\right]^2 = \frac{(6.766)^2}{16} = 2.8612$$

$$SS_{b12} = \frac{1}{2x8}\left[2(0.766 - 1.576 - 2.239 + 1.814 + 0.904 - 3.198 - 2.440 + 3.236)\right]^2 = \frac{(-5.466)^2}{16} = 1.8673$$

$$SS_{b13} = \frac{1}{2x8}\left[2(0.766 - 1.576 + 2.239 - 1.814 - 0.904 + 3.198 - 2.440 + 3.236)\right]^2 = \frac{(5.41)^2}{16} = 1.8293$$

$$SS_{b23} = \frac{1}{2x8}\left[2(0.766 + 1.576 - 2.239 - 1.814 - 0.904 - 3.198 + 2.440 + 3.236)\right]^2 = \frac{(0.286)^2}{16} = 0.00511$$

$$SS_{b123} = \frac{1}{2x8}\left[2(-0.766 + 1.576 + 2.239 - 1.814 + 0.904 - 3.198 - 2.440 + 3.236)\right]^2 = \frac{(0.526)^2}{16} = 0.0173$$

$$\sum SS_R = 3.0189 + 2.6978 + 2.8612 + 1.8673 + 1.8293 + 0.00511 + 0.0173 = 12.2969$$

ix.   Sum of square error,

$$SS_E = SS_T - \sum SS_R$$

but,

$$SS_T = \sum \left(y_{ri}\right)^2 - \frac{\left(\sum y_{ri}\right)^2}{r.N}$$

$$\sum \left(y_{ri}\right)^2 = 0.664^2 + 0.862^2 + 1.564^2 + 1.588^2 + 2.614^2 + 1.864^2 + 1.712^2 + 1.916^2 + 1.526^2$$
$$+ 0.282^2 + 3.192^2 + 3.3204^2 + 2.548^2 + 2.332^2 + 3.360^2 + 3.112^2$$
$$= 78.8288$$

$$\left(\sum y_r\right)^2 = (0.664 + 0.862 + 1.564 + 1.588 + 2.614 + 1.864 + 1.712 + 1.916 + 1.526 + 0.282$$
$$+ 3.192 + 3.3204 + 2.548 + 2.332 + 3.360 + 3.112)^2$$
$$= 1046.1343$$

Hence $SS_T = 78.8288 - \dfrac{1046.1343}{8(2)} = 78.8288 - 65.3834 = 13.4454$

Therefore $SS_E = 13.4458 - 12.2969 = 1.1485$

$$F_{cal} = \frac{MS_R}{MS_E} = \frac{\frac{SS_R}{df_R}}{\frac{SS_E}{N(r-1)}}$$

But $df_R = 1$ and $\dfrac{SS_E}{N(r-1)} = \dfrac{1.1485}{8} = 0.1435$

For b_1, $F_{cal} = \dfrac{3.0189}{0.1435} = 21.0376$

For b_2, $F_{cal} = \dfrac{2.6978}{0.1435} = 18.8$

For b_3, $F_{cal} = \dfrac{2.86129}{0.1435} = 19.9387$

For b_{12}, $F_{cal} = \dfrac{1.8673}{0.1435} = 13.0125$

For b_{13}, $F_{cal} = \dfrac{1.8293}{0.1435} = 12.7038$

For b_{23}, $F_{cal} = \dfrac{0.00511}{0.1435} = 0.0356$

For b_{123}, $F_{cal} = \dfrac{0.0173}{0.1435} = 0.1206$

Table 6:   **Complete analysis of variance**

Source of variation	Sum of squares (SS)	Degree of freedom (df)	Mean square (MS)	F-cal
B_1	3.0189	1	3.0189	21.0376
B_2	2.6978	1	2.6978	18.8000
B_3	2.8612	1	2.8612	19.9387
B_{12}	1.8612	1	1.8612	13.0125
B_{13}	1.8293	1	1.8293	12.7038
B_{23}	0.00511	1	0.0051	0.0356
B_{123}	0.0173	1	0.0173	0.1206
Error	1.1485	$N(r-1) = 8$	$SS_E/N(r-1)$	
Total	13.4454	$Nr-1$		

F-Test

For any coefficient if

F-cal > F[(α, df$_R$, N(r-1)]

We reject the null hypothesis, H$_0$: bj = 0 hence that coefficient is significant.

From statistical table ,

F $(0.05,1,8) = 5.32$

Therefore b$_1$, b$_2$, b$_3$, b$_{12}$ and b$_{13}$ are significant since their F-cal > F-table. The fitted model is still,

y = 2.022 + 0.432x$_1$ + 0.411x$_2$ + 0.423x$_3$ + 0.342x$_1$x$_2$ + 0.338x$_1$x$_3$

From Table 2,

Y$_1$ = 2.022 + 0.434(-1) + 0.411(-1) + 0.423(-1) − 0.342(+1) + 0.338(+1) = 0.750

Y$_2$ = 2.022 + 0.434(+1) + 0.411(-1) + 0.423(-1) − 0.342(-1) + 0.338(-1) = 1.626

Y$_3$ = 2.022 + 0.434(-1) + 0.411(+1) + 0.423(-1) − 0.342(-1) + 0.338(+1) = 2.255

Y$_4$ = 2.022 + 0.434(+1) + 0.411(+1) + 0.423(-1) − 0.342(+1) + 0.338(-1) = 1.764

Y$_5$ = 2.022 + 0.434(-1) + 0.411(-1) + 0.423(+1) − 0.342(+1) + 0.338(-1) = 0.920

Y$_6$ = 2.022 + 0.434(+1) + 0.411(-1) + 0.423(+1) − 0.342(-1) + 0.338(+1) = 3.148

Y$_7$ = 2.022 + 0.434(-1) + 0.411(+1) + 0.423(+1) − 0.342(-1) + 0.338(-1) = 2.424

Y$_8$ = 2.022 + 0.434(+1) + 0.411(+1) + 0.423(+1) − 0.342(+1) + 0.338(+1) = 3.286

Table 7:   **Table of experimental and calculated percentage extract.**

N	Y	ycal	$e = y-ycal$	$e^2 = (y - ycal)^2$
1	0.766	0.750	0.016	0.000256
2	1.576	1.626	-0.050	0.002500
3	2.239	2.255	-0.016	0.000256
4	1.814	1.764	0.050	0.002500
5	0.904	0.924	-0.020	0.000400
6	3.198	3.148	0.050	0.002500
7	2.440	2.424	0.016	0.000256
8	3.236	3.286	-0.050	0.002500 ·

The dispersion of adequacy for the replicate experiment is

$$S_a^2 d = \frac{r}{N - \lambda} \sum_{}^{N} (y - ycal)^2 \ldots\ldots\ldots\ldots\ldots 24$$

λ = Insignificant coefficient

$\lambda = 2$

$$\cdot S_a^2 d = \frac{2}{8-2}(0.000256 + 0.002500 + 0.000256 + 0.002500 + 0.000400 +$$

$$0.002500 + 0.000256 + 0.002500)$$

$$= \frac{2}{6}x0.01168$$

$$= 0.003723$$

F-test

Applying Fisher's criterion

$$\text{F-table} = F[\alpha, N - r, N(r-1)]$$

92

The fitted model is adequate if and only if $F[\alpha, N - r, N(r-1)] > $ f-cal.................25

$$F\text{-cal} = \frac{Sa^2d}{Su^2} \quad ....26$$

$$= \frac{0.003723}{0.14389}$$

$$= 0.02587$$

$F[\alpha, 8 - 2, 8(2 - 1)] = $ F-table

$F[0.05, 6, 8] = 3.58$

Since F-table > F-cal, the fitted model is adequate.

i.e. $y = 2.022 + 0.434x_1 + 0.411x_2 + 0.423x_3 - 0.342x_1x_2 + 0.338x_1x_3$ is

adequate.

To transfer natural variables to their coded levels to be use in calculations.

$$x_{1_{new}} = \frac{(2w_1 - w_1 high - w_1 low)}{(w_1 high - w_1 low)} \quad ....27$$

$$x_{2_{new}} = \frac{(2w_2 - w_2 high - w_2 low)}{(w_2 high - w_2 low)} \quad ....28$$

$$x_{3_{new}} = \frac{(2w_3 - w_3 high - w_3 low)}{(w_3 high - w_3 low)} \quad ....29$$

APPENDIX B

PROGRAM LIST

Program Percentage_Extract(Input,Output,Filedat,Fileout);

{$M 8192,0,0}   { Leave memory for child process }

```
Uses crt,dos;
Const
 xes4 : array[0..15] of string[8] =
 (' ','x1','x2','x3','x4','x1x2','x1x3','x1x4','x2x3','x2x4',
 'x3x4','x1x2x3','x1x2x4','x1x3x4','x2x3x4','x1x2x3x4');
 xes : array[0..7] of string[6] =
 (' ','x1','x2','x3','x1x2','x1x3','x2x3','x1x2x3');
 xes2 : array[0..3] of string[6] =
 (' ','x1','x2','x1x2');
 bes4 : array[0..15] of string[5] =
 ('b0','b1','b2','b3','b4','b12','b13','b14','b23','b24','b34',
 'b123','b124','b134','b234','b1234');
 bes : array[0..7] of string[4] =
 ('b0','b1','b2','b3','b12','b13','b23','b123');
 bes2 : array[0..3] of string[4] =
 ('b0','b1','b2','b12');


Var
 n0,n1,n2,n,r,i,j,k,lam:Integer;
 flag,choice:Integer;
 x: array[1..16,0..15] of integer;
 b,t,xnew,Fcal,SSr:array[0..15] of real;
 btest:array[0..15] of boolean;
 newy,yid,su,su2,diff:array[1..16] of real;
 eu,eu2:array[1..16] of real;
 yidr, yid_r, yid_2r: array[1..3,1..16] of real;
 w1high,w1low,w0high,w0low,w2high,w2low,w3high,w3low:real;
 coninv,tvalue,gvalue,gcal,sumsu,maxsu,eerror,merror,sb:real;
 fvalue1,fvalue2,fcal1,fcal2,sumeu,sad,x1,x2,x3,x4,finy:real;
 w1value,w2value,w3value,w4value,w4low,w4high:real;
 w1name,w2name,w3name,w4name:String[6];
 SSmean,Sumry2,Sumry,SumSSr,SSt,SSe:real;
 equa : string;
 ans:char;
 Filedat,Fileout:Text;
 Filenam : String[12];

function Int2Str(L : Real) : string;
var
 S : string[6];
begin
```

```pascal
  Str(L:6:3, S);
  Int2Str := S;
end;

Function power2(valu:real):real;
Begin
  power2 := valu * valu;
end;

function FileExists(FileName: string) : Boolean;
var
  f: text;
begin
  {$I-}
  Assign(f, FileName);
  Reset(f);
  Close(f);
  {$I+}
  FileExists := (IOResult = 0) and
  (FileName <> '');
end;  { FileExists }

Procedure Clears;
Begin
  Clrscr;

gotoxy(15,1);Writeln('================================================
=====');
  gotoxy(15,2);Writeln('Prediction Models for the % Extract of Oxalic acid');
  gotoxy(15,3);Writeln('   from the bark of Eucalyptus Camaldulensis');

gotoxy(15,4);Writeln('================================================
=====');
  Writeln; Writeln;
End;

Procedure Initialize;
Begin
  for i := 1 to 8 do
    begin
      x[i,0] := 1;
      if i mod 2 = 0 then
        x[i,1] := 1
      else
        x[i,1] := -1;
      if i in [1,2,5,6] then
        x[i,2] := -1
      else
        x[i,2] := 1;
      if i <= 4 then
```

95

```
          x[i,3] := -1
        else
          x[i,3] := 1;
      x[i,4] := x[i,1] * x[i,2];
      x[i,5] := x[i,1] * x[i,3];
      x[i,6] := x[i,2] * x[i,3];
      x[i,7] := x[i,1] * x[i,2] * x[i,3];
    end;
    for i := 0 to 7 do
      btest[i] := TRUE;
end;

Procedure Initial4;
Begin
  for i := 1 to 16 do
    begin
      x[i,0] := 1;
      if i mod 2 = 0 then
        x[i,1] := 1
      else
        x[i,1] := -1;
      if i in [1,2,5,6,9,10,13,14] then
        x[i,2] := -1
      else
        x[i,2] := 1;
      x[i,3] := 1;
      x[i,4] := 1;
      if i <= 4 then x[i,3] := -1;
      if (i >= 9) and (i <= 12) then x[i,3] := -1;
      if i <= 8 then x[i,4] := -1;
      x[i,5] := x[i,1] * x[i,2];
      x[i,6] := x[i,1] * x[i,3];
      x[i,7] := x[i,1] * x[i,4];
      x[i,8] := x[i,2] * x[i,3];
      x[i,9] := x[i,2] * x[i,4];
      x[i,10] := x[i,3] * x[i,4];
      x[i,11] := x[i,1] * x[i,2] * x[i,3];
      x[i,12] := x[i,1] * x[i,2] * x[i,4];
      x[i,13] := x[i,1] * x[i,3] * x[i,4];
      x[i,14] := x[i,2] * x[i,3] * x[i,4];
      x[i,15] := x[i,1] * x[i,2] * x[i,3] * x[i,4];
    end;
    for i := 0 to 15 do
      btest[i] := TRUE;
end;

Procedure Initial2;
Begin
  for i := 1 to 8 do
    begin
```

```
      x[i,0] := 1;
      if i mod 2 = 0 then
        x[i,1] := 1
      else
        x[i,1] := -1;
      if i in [1,2] then
        x[i,2] := -1
      else
        x[i,2] := 1;
      x[i,3] := x[i,1] * x[i,2];
    end;
  for i := 0 to 7 do
    btest[i] := TRUE;
end;


Procedure EnterData;
Begin
  flag := 1;
  Clears;
  repeat
    Write('Enter Number of Factors (2,3 or 4) ');Readln(n1);
  until (n1 >= 2) and (n1 <= 4);
  n0 := n1;
  if n1 = 2 then
    n := 4;
  if n1 = 3 then
    n := 8;
  if n1 = 4 then
    n := 16;
  w3high := 0;
  w3low := 0;
  w4high := 0;
  w4low := 0;
  Write('Enter Name of W1 (Max 6 Chars) : ');Readln(W1name);
  Write('Enter High Value for W1 : ');Readln(W1high);
  Write('Enter Low Value for W1 : ');Readln(W1low);
  Write('Enter Name of W2 (Max 6 Chars) : ');Readln(W2name);
  Write('Enter High Value for W2 : ');Readln(W2high);
  Write('Enter Low Value for W2 : ');Readln(W2low);
if n >= 8 then
begin
  Write('Enter Name of W3 (Max 6 Chars) : ');Readln(W3name);
  Write('Enter High Value for W3 : ');Readln(W3high);
  Write('Enter Low Value for W3 : ');Readln(W3low);

end;

if n = 16 then
begin
```

```
    Write('Enter Name of W4 (Max 6 Chars) : ');Readln(W4name);
    Write('Enter High Value for W4 : ');Readln(W4high);
    Write('Enter Low Value for W4 : ');Readln(W4low);
end;
 W0high := 1;
 W0low := -1;
 repeat
   Write('Enter Number of Replicates (2,3 or 4) ');Readln(r);
 until (r >= 2) or (r <= 4);
 Clears;
 for i := 1 to n do
 begin
  yid[i] := 0;
  for j := 1 to r do
    Begin
      Write('Enter % Extract ',i:2,' for replicate ',j:2,': ');readln(yidr[j,i]);
      yid[i] := yid[i] + yidr[j,i];
    end;
    yid[i] := yid[i] / r;
 end;
 Clears;
 Write('High Value for W1 : ');Writeln(W1high:7:4);
 Write('Low Value for W1  : ');Writeln(W1low:7:4);
 Write('High Value for W2 : ');Writeln(W2high:7:4);
 Write('Low Value for W2  : ');Writeln(W2low:7:4);
if n >= 8 then
begin
 Write('High Value for W3 : ');Writeln(W3high:7:4);
 Write('Low Value for W3  : ');Writeln(W3low:7:4);
end;
if n = 16 then
begin
 Write('High Value for W4 : ');Writeln(W4high:7:4);
 Write('Low Value for W4  : ');Writeln(W4low:7:4);
end;
 Write('Number of Replicates : ');Writeln(r);
 Writeln;
  for j := 1 to r do
    Begin
      Write('   Yr',j:1,'      ');
    end;
    Writeln;
 for i := 1 to n do
 begin
  for j := 1 to r do
    Begin
      Write(yidr[j,i]:10:4,'   ');
    end;
    Writeln;
 end;
```

```pascal
Repeat
   Write('Do you wish to write data to file (Y/N) ');readln(ans);
until (Upcase(ans) = 'Y') or (Upcase(ans) = 'N');
if upcase(ans) = 'Y' then
   begin
   Writeln('Enter Data File name (Max 12 chars - AAAAAAAA.AAA) : ');
   Readln(Filenam);
      Assign(Filedat,Filenam);
      Rewrite(Filedat);
      Writeln(Filedat,W0high);
      Writeln(Filedat,W1high);
      Writeln(Filedat,W2high);
      Writeln(Filedat,W3high);
      Writeln(Filedat,W4high);
      Writeln(Filedat,W0low);
      Writeln(Filedat,W1low);
      Writeln(Filedat,W2low);
      Writeln(Filedat,W3low);
      Writeln(Filedat,W4low);
      Writeln(Filedat,W1name);
      Writeln(Filedat,W2name);
      Writeln(Filedat,W3name);
      Writeln(Filedat,W4name);
      Writeln(Filedat,r);
      Writeln(Filedat,n);
      for i := 1 to n do
       begin
        for j := 1 to r do
          Writeln(Filedat,yidr[j,i]);
          Writeln(Filedat,yid[i]);
       end;
      close(Filedat);
   end;
End;

Procedure LoadData;
Begin
clears;
Writeln('Enter Data File name (Max 12 chars - AAAAAAAA.AAA) : ');
Readln(Filenam);

 if not FileExists(Filenam) then
  begin
   clears;
   WriteLn('File not found');
   Writeln('Press Enter Key to Continue...');
   Readln;
  end
 else
```

```
begin
assign(Filedat,Filenam);
flag := 2;
Reset(Filedat);
while not eof(Filedat) do
  begin
     Readln(Filedat,W0high);
     Readln(Filedat,W1high);
     Readln(Filedat,W2high);
     Readln(Filedat,W3high);
     Readln(Filedat,W4high);
     Readln(Filedat,W0low);
     Readln(Filedat,W1low);
     Readln(Filedat,W2low);
     Readln(Filedat,W3low);
     Readln(Filedat,W4low);
     Readln(Filedat,W1name);
     Readln(Filedat,W2name);
     Readln(Filedat,W3name);
     Readln(Filedat,W4name);
     Readln(Filedat,r);
     Readln(Filedat,n);
     for i := 1 to n do
      begin
       for j := 1 to r do
         Readln(Filedat,yidr[j,i]);
         Readln(Filedat,yid[i]);
      end;
  end;
  close(Filedat);
Clears;
Write('High Value for W1 : ');Writeln(W1high:7:4);
Write('Low Value for W1  : ');Writeln(W1low:7:4);
Write('High Value for W2 : ');Writeln(W2high:7:4);
Write('Low Value for W2` : ');Writeln(W2low:7:4);
Write('High Value for W3 : ');Writeln(W3high:7:4);
Write('Low Value for W3  : ');Writeln(W3low:7:4);
Write('High Value for W4 : ');Writeln(W4high:7:4);
Write('Low Value for W4  : ');Writeln(W4low:7:4);
Write('Number of Replicates : ');Writeln(r);
Writeln;
. for j := 1 to r do
  Begin
    Write('   Yr',j:1,'      ');
  end;
  Writeln;
for i := 1 to n do
begin
  for j := 1 to r do
  Begin
```

100

```pascal
        Write(yidr[j,i]:10:4,'    ');
      end;
      Writeln;
    end;
  Writeln;
  Writeln('Data has been Loaded into memory !');
  Writeln;
  Write('Press Enter to Continue ');readln;
end;
End;

Procedure Compute;
Begin
if n = 4 then n0 := 2;
if n = 8 then n0 := 3;
if n = 16 then n0 := 4;
if n0 = 4 then
 initial4;
if n0 = 3 then
 initialize;
if n0 = 2 then
 initial2;
 if flag = 0 then
   Begin
     Clears;
     gotoxy(20,10);Write('Empty Data !, Select options 1 or 2 before 3');
     Delay(3500);
   end
 else
   Begin
     n2 := n - 1;
     Clears;
     Writeln('Enter Output File name (Max 12 chars - AAAAAAAA.AAA) : ');
     Readln(Filenam);
     assign(Fileout,Filenam);
     Rewrite(Fileout);
     Write('Enter G-Value Tabulated - G[à,(r-1),N] : ');Readln(gvalue);
     sumsu := 0;
        for i := 1 to n do
         begin
          su2[i] := 0;
          for j := 1 to r do
            begin
              diff[i] := yidr[j,i]-yid[i];
              su2[i] := su2[i] + sqr(diff[i])
            end;
            su2[i] := su2[i] / (r-1);
            sumsu := sumsu + su2[i];
         end;
        maxsu := su2[1];
```

101

```pascal
        for i := 1 to n do
          if su2[i] > maxsu then maxsu := su2[i];
        gcal := maxsu / sumsu;

Writeln(Fileout,'==================================================================
==');
        Writeln(Fileout,'Prediction Models for the % Extract of Oxalic acid');
        Writeln(Fileout,'    from the bark of Eucalyptus Camaldulensis');

Writeln(Fileout,'==================================================================
==');
        Writeln(Fileout);
        Writeln(Fileout,'Natural Factors and their Coded levels');
        for i := 1 to 70 do Write(Fileout,'=');Writeln(Fileout);
        if n = 16 then
          Writeln(Fileout,'Level of Factors  Code   ',w1name:6,' ',w2name:6,'
',w3name:6,' ',w4name:6);
        if n = 8 then
          Writeln(Fileout,'Level of Factors  Code   ',w1name:6,' ',w2name:6,'
',w3name:6);
        if n = 4 then
          Writeln(Fileout,'Level of Factors  Code   ',w1name:6,' ',w2name:6);
        for i := 1 to 70 do Write(Fileout,'=');Writeln(Fileout);
          Writeln(Fileout,'High',' ':15,'+1 ',w1high:9:4,
          w2high:9:4,w3high:10:4,w4high:10:4);
          Writeln(Fileout,'Low ',' ':15,'-1 ',w1low:9:4,
          w2low:9:4,w3low:10:4,w4low:10:4);
          Writeln(Fileout); Writeln(Fileout);
        Writeln(Fileout,'Experimental Results with Replicates');
        for i := 1 to n*10+6 do Write(Fileout,'=');Writeln(Fileout);
        Write(Fileout,' N  |');
        for i := 1 to n  do
          Write(Fileout,' ',i:3,'  |');
          Writeln(Fileout);
        for i := 1 to n*10+6 do Write(Fileout,'=');Writeln(Fileout);
        for j := 1 to r do
        begin
        Write(Fileout,' Yr',j:1,' |');
        for i := 1 to n  do
          Write(Fileout,' ',yidr[j,i]:7:4,' |');
          Writeln(Fileout);
        end;
        Write(Fileout,' Y % |');
        for i := 1 to n  do
          Write(Fileout,' ',yid[i]:7:4,' |');
          Writeln(Fileout);
        for j := 1 to r do
        begin
        Write(Fileout,'Yr',j:1,'-Y|');
        for i := 1 to n  do
```

102

```pascal
     Write(Fileout,' ',yidr[j,i]-yid[i]:7:4,' |');
     Writeln(Fileout);
end;
for j := 1 to r do
begin
Write(Fileout,'Sqr',j:1,' |');
for i := 1 to n  do
   Write(Fileout,' ',sqr(yidr[j,i]-yid[i]):7:4,' |');
   Writeln(Fileout);
end;
Write(Fileout,'Su^2 |');
for i := 1 to n  do
   Write(Fileout,' ',su2[i]:7:4,' |');
   Writeln(Fileout);
   Writeln(Fileout);
Writeln(Fileout,'The Sum of the dispersion     = ',sumsu:7:4);
Writeln(Fileout,'The maximum Su^2             = ',maxsu:7:4);
Writeln(Fileout,'G-Calculated                = ',gcal:7:4);
Writeln(Fileout,'G-Statistical Table [à,(r-1),N] = ',gvalue:7:4);
Writeln(Fileout);
Writeln(Fileout,'G-TEST:');
if gvalue < gcal then
 begin
  Clears;
  Writeln(Fileout,'It is not possible to carry out regression');
  Writeln(Fileout,'analysis, since G-stat < G-cal');
  Writeln('It is not possible to carry out regression');
  Writeln('analysis, since G-stat < G-cal'); Writeln;
  Writeln('Press AnyKey to Continue');
  Readln;

 end
else
 begin
  Writeln(Fileout,'It is possible to carry out regression');
  Writeln(Fileout,'analysis, since G-stat > G-cal');
  Writeln(Fileout);

  Clears;
  Write('Enter T-Value Tabulated - T[à,N(r-1)] : ');Readln(tvalue);
  for i := 1 to 78 do Write(Fileout,'-');Writeln(Fileout);
   Writeln(Fileout);
   merror := sumsu / n;
   eerror := sqrt(merror);
Write(Fileout,'Design Matrix Table for a 2^',n0:2,' full factorial');
Writeln(Fileout,'Design with the interactions');
for i := 1 to n*10 do Write(Fileout,'=');Writeln(Fileout);
if n = 16 then
begin
  Write(Fileout,' X0    X1    X2    X3    X4    X1*X2');
```

103

```pascal
  Write(Fileout,'  X1*X3  X1*X4  X2*X3  X2*X4  X3*X4  X1X2X3');
  Writeln(Fileout,'  X1X2X4  X1X3X4  X2X3X4  X1X2X3X4    Y');
end;
if n = 8 then
begin
  Write(Fileout,'  X0    X1    X2    X3   X1*X2');
  Writeln(Fileout,'  X1*X3  X2*X3  X1*X2*X3   Y');
end;
if n = 4 then
begin
  Writeln(Fileout,'  X0    X1    X2   X1*X2   Y');
end;
for i := 1 to n*10 do Write(Fileout,'=');Writeln(Fileout);
for i := 1 to n  do
begin
 for j := 0 to n2 do
  Write(Fileout,' ',x[i,j]:3,' |');
  Write(Fileout,' ',yid[i]:7:4);
  Writeln(Fileout);
end;
  Writeln(Fileout); Writeln(Fileout);

Writeln(Fileout,'The mean square error    = ',merror:9:5);
Writeln(Fileout,'The experimental error   = ',eerror:9:5);
Writeln(Fileout);
  for i := 0 to n2 do
    begin
      b[i] := 0;
      for j := 0 to n2 do
        b[i] := b[i] + (x[j+1,i]*yid[j+1]);
        b[i] := b[i] / n;
    end;
      sb := eerror / sqrt(n*r);
      for i := 0 to n2 do
      begin
       t[i] := abs(b[i]) / sb;
       if t[i] < tvalue then btest[i] := FALSE;
      end;
  equa := 'Y = ';
  lam := 0;
  for i := 0 to n2 do
   if btest[i] then
    begin
    if (i <> 0) and (b[i] > 0) then equa := equa + '+';
    if n = 16 then
       equa := equa + (' '+Int2str(b[i])+xes4[i]+' ');
    if n = 8 then
       equa := equa + (' '+Int2str(b[i])+xes[i]+' ');
    if n = 4 then
      equa := equa + (' '+Int2str(b[i])+xes2[i]+' ');
```

104

```pascal
        end
      else
        lam := lam + 1;
    Writeln(Fileout,'T-TEST:');
    Writeln(Fileout,'T-Statistical Table [à,N(r-1)] :',tvalue:7:4);
      Writeln(Fileout,'The constant and the variable for the following');
      Writeln(Fileout,'are insignificant since T-cal < T-table');
      for i := 0 to n2 do
        if not btest[i] then
          begin
            if n = 16 then
              Write(Fileout,bes4[i],' ');
            if n = 8 then
              Write(Fileout,bes[i],' ');
            if n = 4 then
              Write(Fileout,bes2[i],' ');
          end;
            Writeln(Fileout);
    coninv := tvalue * sb;
    Writeln(Fileout);                    •
    Writeln(Fileout,'The fitted model then becomes:');
    Writeln(Fileout,equa);
    Writeln(Fileout); Writeln(Fileout);
    writeln(fileout,'Table of Calculated T-values');
    Writeln(Fileout,'Regression  Estimated  Confidence   t');
    Writeln(Fileout,'Coefficient  Effect    Interval  Values');
    for i := 1 to 45 do Write(Fileout,'=');Writeln(Fileout);
    for i := 0 to n2  do
    begin
     if n = 16 then
       Writeln(Fileout,bes4[i]:6,'      ',b[i]:7:3,'   ',
       coninv:7:3,'  ',t[i]:9:5);
     if n = 8 then
       Writeln(Fileout,bes[i]:6,'       ',b[i]:7:3,'   ',
       coninv:7:3,'  ',t[i]:9:5);
     if n = 4 then
       Writeln(Fileout,bes2[i]:6,'       ',b[i]:7:3,'   ',
       coninv:7:3,'  ',t[i]:9:5);
    end;
    for i := 1 to 45 do Write(Fileout,'=');Writeln(Fileout);
    for i := 1 to n do
     begin
      newy[i] := 0;
      for j := 0 to n2 do
       if btest[j] then
         newy[i] := newy[i] + (b[j]*x[i,j]);
    end;

     Clears;
     Write('Enter F-Value Tabulated - F[à,dFr,N(r-1)] : ');
```

105

```pascal
    Readln(fvalue1);
    for i := 1 to 78 do Write(Fileout,'-');Writeln(Fileout);
     Writeln(Fileout);
Writeln(Fileout);
Sumry2 := 0;
Sumry := 0;
SumSSr := 0;
    for i := 1 to n do
      for k := 1 to r do
       begin
        Sumry2 := sumry2 + power2(yidr[k,i]);
        Sumry := Sumry + (yidr[k,i]);
       end;

    for i := 1 to n2 do
     begin
      SSr[i] := power2((b[i] * r * n))/(r*n);
      SumSSr := SumSSr + SSr[i];
     end;
       SSt := Sumry2 - (power2(Sumry)/(r*n));
       SSe := SSt - SumSSr;
       SSmean := SSe / (n*(r-1));
       for i := 1 to n2 do
       begin
        Fcal[i] := SSr[i]/ SSmean;
        btest[i] := TRUE;
        if Fcal[i] < fvalue1 then btest[i] := FALSE;
       end;
Writeln(Fileout,'F-TEST :');
Writeln(Fileout,'F-Statistical Table [à,dFr,N(r-1)] :',fvalue1:7:4);
  Writeln(Fileout,'The constant and the variable for the following');
  Writeln(Fileout,'are insignificant since F-cal < F-table');
  for i := 1 to n2 do
   if not btest[i] then
    begin
     if n = 16 then
       Write(Fileout,bes4[i],' ');
     if n = 8 then
       Write(Fileout,bes[i],' ');
     if n = 4 then
       Write(Fileout,bes2[i],' ');
    end;
Writeln(Fileout); Writeln(Fileout);
writeln(Fileout,'Complete Analysis of Variance');
Writeln(Fileout,'Source of    Sum of    Degree of   Mean     F-cal');
Writeln(Fileout,'Variation  Squares(SS) Freedom(df) Square        ');
for i := 1 to 55 do Write(Fileout,'=');Writeln(Fileout);
for i := 1 to n2  do
begin
 if n = 16 then
```

```pascal
    Writeln(Fileout,bes4[i]:6,'        ',SSr[i]:7:3,'        1',
    '        ',SSr[i]:7:3,'  ',Fcal[i]:9:5);
    if n = 8 then
      Writeln(Fileout,bes[i]:6,'        ',SSr[i]:7:3,'        1',
      '        ',SSr[i]:7:3,'  ',Fcal[i]:9:5);
    if n = 4 then
      Writeln(Fileout,bes2[i]:6,'        ',SSr[i]:7:3,'        1',
      '        ',SSr[i]:7:3,'  ',Fcal[i]:9:5);
  end;
  for i := 1 to 55 do Write(Fileout,'=');Writeln(Fileout);
  Writeln(Fileout,'Error (SSE)    = ',SSe:9:5);
  Writeln(Fileout,'Total (SST)    = ',SSt:9:5);
  Writeln(Fileout,'N(r-1)        = ',n:5);
  Writeln(Fileout,'Nr-1          = ',n*r - 1:5);
  Writeln(Fileout,'SSE/(N(r-1)    = ',SSmean:9:5);
  Writeln(Fileout);

  Clears;
  Write('Enter F-Value Tabulated - F[à,N-r,N(r-1)] : ');Readln(fvalue2);
  Writeln(Fileout);
  writeln(Fileout,'Experimental and Calculated',
  ' percentage extract');
  Writeln(Fileout,' N        Y      Ycal    eu=Y-Ycal',
  ' eu^2=(Y-Ycal)^2');
  for i := 1 to 55 do Write(Fileout,'=');Writeln(Fileout);
  sumeu := 0;
  for i := 1 to n  do
   begin
    eu[i] := yid[i] - newy[i];
    eu2[i] := sqr(eu[i]);
    Writeln(Fileout,i:4,'    ',yid[i]:7:3,' ',newy[i]:7:3,
    '  ',eu[i]:10:6,'    ',eu2[i]:10:6);
    sumeu := sumeu + eu2[i];
   end;
  sad := r/(n-lam)*sumeu;
  fcal2 := sad / merror;
  Writeln(Fileout);
  for i := 1 to 55 do Write(Fileout,'=');Writeln(Fileout);
  Writeln(Fileout);
  Writeln(Fileout,'F-TEST (Fisher):');
  Writeln(Fileout,'F-Calculated          = ',fcal2:11:7);
  Writeln(Fileout,'F-Statistical Table    = ',fvalue2:11:7);
  Writeln(Fileout);
  if fcal2 > fvalue2 then
   begin
    Writeln(Fileout,'The fitted model is inadequate');
    Writeln(Fileout,'Since F-cal > F-table (Fisher) ');
    Writeln(Fileout);
   end
  else
```

```
          begin
            Writeln(Fileout,'The fitted model is adequate');
            Writeln(Fileout,'Since F-cal < F-table');
            Writeln(Fileout);
            for i := 1 to 78 do Write(Fileout,'-');Writeln(Fileout);
          end;
Repeat
Clears;
Repeat
  Write('Do you wish to continue by entering values for Computation (Y/N) ');
  readln(ans);
until (Upcase(ans) = 'Y') or (Upcase(ans) = 'N');
if upcase(ans) = 'Y' then
  begin
      repeat
      Write('Enter Value of W1 (must be between w1high & w1low): ');
      Readln(w1value);
      until (w1value >= w1high) and (w1value <= w1low)
          or (w1value <= w1high) and (w1value >= w1low);
      repeat
      Write('Enter Value of W2 (must be between w2high & w2low): ');
      Readln(w2value);
      until (w2value >= w2high) and (w2value <= w2low)
      or (w2value <= w2high) and (w2value >= w2low);
      x1 := ((2*w1value)-w1low-w1high)/(w1high-w1low);
      x2 := ((2*w2value)-w2low-w2high)/(w2high-w2low);
    if n >= 8 then
    begin
      repeat
      Write('Enter Value of W3 (must be between w3high & w3low): ');
      Readln(w3value);
      until (w3value >= w3high) and (w3value <= w3low)
          or (w3value <= w3high) and (w3value >= w3low);
      x3 := ((2*w3value)-w3low-w3high)/(w3high-w3low);
      xnew[0] := 1;
      xnew[1] := x1;
      xnew[2] := x2;
      xnew[3] := x3;
      xnew[4] := x1*x2;
      xnew[5] := x1*x3;
      xnew[6] := x2*x3;
      xnew[7] := x1*x2*x3;
    end;
    if n = 16 then
    begin
      repeat
      Write('Enter Value of W4 (must be between w4high & w4low): ');
      Readln(w4value);
      until (w4value >= w4high) and (w4value <= w4low)
          or (w4value <= w4high) and (w4value >= w4low);
```

```pascal
      x4 := ((2*w4value)-w4low-w4high)/(w4high-w4low);
      xnew[0] := 1;
      xnew[1] := x1;
      xnew[2] := x2;
      xnew[3] := x3;
      xnew[4] := x1*x2;
      xnew[5] := x1*x3;
      xnew[6] := x2*x3;
      xnew[7] := x1*x2*x3;
    end;
    if n = 4 then
    begin
      xnew[0] := 1;       .
      xnew[1] := x1;
      xnew[2] := x2;
      xnew[3] := x1*x2;
    end;
       finy := 0;
       for j := 0 to n2 do
        if btest[j] then              .
          finy := finy + (b[j]*xnew[j]);
       Writeln(Fileout);
    for i := 1 to 75 do Write(Fileout,'*');Writeln(Fileout);
       Writeln(Fileout,'Value of w1,x1    = ',w1value:9:4,x1:9:4);
       Writeln(Fileout,'Value of w2,x2    = ',w2value:9:4,x2:9:4);
    if n >= 8 then
       Writeln(Fileout,'Value of w3,x3    = ',w3value:9:4,x3:9:4);
    if n = 16 then
       Writeln(Fileout,'Value of w4,x4    = ',w4value:9:4,x4:9:4);
       Writeln(Fileout);
       Writeln(Fileout,equa);
       Writeln(Fileout);
       Writeln(Fileout,'Final value of Y  = ',finy:9:4);
       Writeln(Fileout);
    end;
  until (Upcase(ans) = 'N');
    for i := 1 to 75 do Write(Fileout,'*');Writeln(Fileout);
    close(Fileout);
    Clears;
    SwapVectors;
    Exec(GetEnv('COMSPEC'), '/C ' + 'edit '+Filenam);
    SwapVectors;
   end;
End;
end;

Begin
 flag := 0;
 assign(Filedat,'extract.dat');
 assign(Fileout,'extract.out');
```

```pascal
repeat
 repeat
  Clears;
  gotoxy(25,7);Writeln('M A I N   M E N U');
  gotoxy(25,8);Writeln('*****************');
  gotoxy(23,10) ;Writeln('1. Enter Fresh Data ');
  gotoxy(23,12);Writeln('2. Load Data from File');
  gotoxy(23,14);Writeln('3. Computation    ');
  gotoxy(23,16);Writeln('4. Quit Program      ');
  gotoxy(25,19);Write('Enter Choice (1-4) ');readln(choice);
 until (choice > 0) and (choice <= 4);
 Case Choice of
   1:EnterData;
   2:LoadData;
   3:Compute;
 end;
until choice = 4;
end.
```