OPTIMAL AND EFFICIENT PRODUCTION OF ROSE COCO BEANS THROUGH THE TWENTY FOUR POINTS SECOND ORDER ROTATABLE DESIGN

BY
ISAAC KIPKOSGEI TUM
(Msc. Biostatistics)

A THESIS SUBMITTED FOR EXAMINATION IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN BIOSTATISTICS IN THE DEPARTMENT OF STATISTICS AND COMPUTER SCIENCE, SCHOOL OF BIOLOGICAL AND PHYSICAL SCIENCES OF MOI UNIVERSITY

2018
DECLARATION

Declaration by the Candidate

This thesis is my original work and has not been presented for a degree in any other university. No part of this thesis may be reproduced without prior written permission of the author and/or Moi University.

Isaac Kipkosgei Tum

Signature: ___________________________ Date: ________________

PHD/BS/003/12

Declaration by the Supervisors

This thesis has been submitted with our approval as the university supervisors

Prof. J. K. Arap Koske

Signature: ___________________________ Date: ________________

Department of Statistics and Computer Science,
Moi University,
P.O. Box 3900,
Eldoret, Kenya.

Prof. J.M. Mutiso

Signature: ___________________________ Date: ________________

Department of Statistics and Computer Science,
Moi University,
P.O. Box 3900,
Eldoret, Kenya.
DEDICATION

I dedicate this work to my lovely wife Viola, son Ryan, daughter Jayleen, my parents; Samuel Tum and Rosemary Tum, grandma and my siblings. Thank you all for special love, support, understanding, encouragement and prayers during the write-up of this thesis.
ACKNOWLEDGEMENT

Firstly, I thank Almighty God for giving me life, the courage and the determination, as well as guidance in furthering my studies, despite all odds. Secondly, I thank my supervisors Prof. J.K. Koske and Prof. J.M. Mutiso, they have totally transformed me. Without their patience and encouragement, this thesis would not have been possible. I sincerely appreciate the invaluable academic support I received from them throughout the writing of this thesis, their kind, detailed and constructive comments were vital to the development of this dissertation. Thirdly, my acknowledgement goes to Moi University staff members in the school of biological and physical sciences (in particular Dean, Prof Kiprop who occasionally sent us letters to show the ‘Ph.D progress’), and I am also indebted to my workmates at the department of statistics and computer science. I wish to extend my gratitude to the classmates (2012 class) for their wonderful cooperation, encouragement and moral support, for being together right from the Masters class in the University of Eldoret after which we enrolled together for Ph.D class in the Moi University. Finally, my further gratitude goes to my family members from whom my success is derived from (especially Jebet who assisted much in the greenhouse work); I am truly blessed to have you all as my family. Indeed, your efforts and contributions in my work cannot be exhausted, though God will reward you abundantly. Thank you.
ABSTRACT

Response Surface Methodology (RSM) is a collection of mathematical and statistical techniques useful for the modeling and analysis of problems in which a response of interest is influenced by several variables and the objective is to optimize the response. The yield results of the twenty four points RSM design permitted a response surface to be fitted easily and provided spherical information contours besides the realizations of an optimum combination of the fertilizers in *rose coco* beans, which resulted in economic use of scarce resources for optimal production of *rose coco* beans. In this study an existing A-optimum and D-efficient second order rotatable design in three dimensions was used in the production of *rose coco* beans optimally and efficiently. The general objective of the study was to produce *rose coco* beans (*Phaseolus vulgaris*) optimally and efficiently using an existing A-optimum and D-efficient twenty four points second order rotatable design in three dimensions in a greenhouse setting using three inorganic fertilizers, namely, nitrogen, phosphorus and potassium. Thus the study was accomplished using the calculus optimum value of the free/letter parameter $f=1.1072569$. The specific objectives were to estimate the linear parameters, thereby making available the yield response of *rose coco* beans at calculus optimum value design for the first time. The generalized variance of the estimated linear parameters was also obtained, fitted and tested the three models adequacies via lack of fit test, and then found the settings of the experimental factors that produces optimal response using contour plots to assist visualizes the response surfaces. This study demonstrated the importance of statistical methods in the optimal and efficient production of *rose coco* beans. The results showed that the three factors: nitrogen, phosphorus, and potassium contributed significantly on the yield of *rose coco* beans ($p<0.05$). The regression coefficients were determined by employing least squares techniques to predict quadratic polynomial models for group 1 greenhouse (GP1G), group 2 greenhouse (GP2G) and group 3 greenhouse (GP3G) for the three fertilizer combinations. In GP1G second order model was inadequate with a $p$ value of 0.3178, in GP2G and GP3G, the second order model was adequate at 1% level of significance with $p$ values of 0.0065 and 0.0034 respectively. The analysis of variance (ANOVA) of response surface for *rose coco* yield showed that this design was adequate due to satisfactory levels of coefficient of determination, $R^2$, 0.6810 (GP1G), 0.6704 (GP2G), and 0.8066 (GP3G) and coefficient variation, $CV$ was 13.48, 14.47 and 10.30 for GP1G, GP2G, GP3G respectively. The canonical analysis showed that there were saddle points for the three groups, meaning there was no unique optimum; therefore ridge analysis was used to overcome the saddle problem. The results from ridge analysis provided the maximum yield of 58.78 grams, 48.36 grams and 70.25 grams in GP1G, GP2G and GP3G respectively for the various fertilizer combinations at radii of one. We therefore recommend the use of GP3G design since it gave above board the required coefficient of determination ($R^2=80.66$) and the maximum yield (70.25 grams) was achieved.
TABLE OF CONTENTS

DECLARATION...................................................................................................................II
DEDICATION .................................................................................................................... III
ACKNOWLEDGEMENT................................................................................................... IV
ABSTRACT ......................................................................................................................... V
LIST OF TABLES ............................................................................................................. XII
LIST OF FIGURES .......................................................................................................... XIV
ABBREVIATIONS AND ACRONYMS......................................................................... XVI
CHAPTER 1: INTRODUCTION.......................................................................................... 1
  1.1 BACKGROUND INFORMATION ............................................................................. 1
  1.2 STATEMENT OF THE PROBLEM ......................................................................... 4
  1.3 OBJECTIVES OF THE STUDY ............................................................................. 6
    1.3.1 General objective.............................................................................................. 6
    1.3.2 Specific objectives............................................................................................ 6
  1.4 SIGNIFICANCE OF THE STUDY ......................................................................... 7
  1.5 LIMITATION OF THE STUDY .............................................................................. 7
CHAPTER 2: LITERATURE REVIEW............................................................................... 8
  2.1 INTRODUCTION .................................................................................................... 8
  2.2 THE ROSE COCO BEAN PRODUCTION IN KENYA ........................................... 8
  2.3 ORGANIC AND INORGANIC USE OF FERTILIZERS ........................................... 9
  2.4 RESPONSE SURFACE METHODOLOGY .......................................................... 10
    2.4.1 Optimal Design Criteria .................................................................................. 12
    2.4.2 D-optimality ................................................................................................... 13
  2.5 MODEL ADEQUACY ............................................................................................. 14
CHAPTER 3: METHODOLOGY....................................................................................... 16
  3.1 INTRODUCTION .................................................................................................. 16
  3.2 ESTIMATION OF THE LINEAR PARAMETERS IN AN EXISTING A-OPTIMUM AND D-EFFICIENT CALCULUS OPTIMUM VALUE SECOND ORDER ROTATABLE DESIGN. ................. 16
    3.2.1 Twenty Four Points Calculus Optimum Value Experimental Design .......... 16
    3.2.2 Screening experiment .................................................................................... 17
3.2.3 Experiment Layout .............................................................................................. 18
  3.2.3.1 Layout of calculus optimum value design .................................................... 18
  3.2.3.2 Group 1 Greenhouse (GP1G) ....................................................................... 18
  3.2.3.3 Group 2 Greenhouse (GP2G) ....................................................................... 19
  3.2.3.4 Group 3 Greenhouse (GP3G) ....................................................................... 19
  3.2.4 Planting Rose coco Beans ................................................................................... 19
  3.2.5 Data collection ..................................................................................................... 20
  3.2.6 Method of estimating the parameters .................................................................. 21

3.3 OBTAINING THE GENERALIZED VARIANCE OF THE ESTIMATED LINEAR PARAMETERS.. 21
  3.3.1 Residual sum of squares and variance ................................................................. 21

3.4 FITTING AND TESTING THE THREE MODELS ADEQUACIES ............................................ 22
  3.4.1 Design of fitting the first-order model ................................................................. 22
  3.4.2 The second-order model ...................................................................................... 26
  3.4.3 Model adequacy checking ................................................................................... 28
  3.4.4 The residual analysis ........................................................................................... 29

3.5 THE SETTINGS OF THE EXPERIMENTAL FACTORS THAT PRODUCES THE OPTIMAL
  RESPONSE .......................................................................................................................... 30
  3.5.1 Coding the factor variables ................................................................................. 30
  3.5.2 The stationary point ............................................................................................. 31
  3.5.3 The canonical analysis ........................................................................................ 33
  3.5.4 The ridge analysis ................................................................................................ 34

CHAPTER 4: RESULTS AND DISCUSSIONS ................................................................ 36

4.1 INTRODUCTION ........................................................................................................... 36

4.2 ESTIMATION OF THE LINEAR PARAMETERS IN AN EXISTING A- OPTIMUM AND D-
  EFFICIENT CALCULUS OPTIMUM VALUE SECOND ORDER ROTATABLE DESIGN ............. 36
  4.2.2 Twenty four points rotatable design of coded levels and natural levels with the
  yields of rose coco beans-GP1G .................................................................................. 39
    4.2.2.1 The three replication yield of twenty four points second order rotatable
    design, Group 1-GP1G ............................................................................................. 41
  4.2.3 Twenty four points rotatable design of coded levels and natural levels with the
  yield of rose coco beans-GP2G .................................................................................. 43
4.2.3.1 The three replication yield of twenty four points second order rotatable
design, Group 2-GP2G ........................................................................................................44
4.2.4 Twenty four points rotatable design of coded levels and natural levels with the
yield of rose coco beans-GP3G ..........................................................................................46
  4.2.4.1 The three replication yield of twenty four point second order rotatable
design, Group 3-GP3G ......................................................................................................47
4.2.5 Parameter estimation for the first order model ......................................................50
  4.2.5.1 Linear parameters of first-order model for GP1G ........................................ 51
  4.2.5.2 Linear parameters of first-order model for GP2G ........................................ 51
  4.2.5.3 Linear parameters of first-order model for GP3G ........................................ 52
4.2.6 Parameter estimation for the second order model .............................................. 52
  4.2.6.1 The parameter estimation for GP1G ............................................................. 54
  4.2.6.2 The parameter estimation for GP2G ............................................................. 55
  4.2.6.3 The parameter estimation for GP3G ............................................................. 55
4.3 OBTAINING THE GENERALIZED VARIANCE OF THE ESTIMATED LINEAR PARAMETERS.. 56
  4.3.1 Residual variance ................................................................................................ 56
    4.3.1.1 The residual variance for GP1G ................................................................... 56
    4.3.1.2 The residual variance for GP2G ................................................................... 56
    4.3.1.3 The residual variance for GP3G ................................................................... 56
  4.3.2 First order model variance-covariance ............................................................ 57
    4.3.2.1 The variance-covariance for GP1G .............................................................. 57
    4.3.2.2 The variance-covariance for GP2G .............................................................. 58
    4.3.2.3 The variance-covariance for GP3G .............................................................. 58
  4.3.3 The second order model variance-covariance .................................................. 59
    4.3.3.1 The standard error, variance-covariance for GP1G ...................................... 59
    4.3.3.2 The standard error, variance-covariance for GP2G ...................................... 60
    4.3.3.3 The standard error, variance-covariance for GP3G ...................................... 61
4.4 FITTING AND TESTING THE THREE MODELS ADEQUACIES ............................................ 61
  4.4.1 The first-Order Model ......................................................................................... 61
    4.4.1.1 The first-order model fitting for GP1G ........................................................ 62
    4.4.1.2 The first-order model fitting for GP2G ........................................................ 63
4.4.1.3 The first-order model fitting for GP3G ......................................................... 63
4.4.2 The Analysis of variance of GP1G ................................................................... 63
  4.4.2.1 The test of significance in regression-GP1G ............................................... 64
  4.4.2.2 Normal probability plot of the residuals GP1G ........................................ 65
  4.4.2.3 The test of significance of the model ...................................................... 66
  4.4.2.4 The test of significance of parameter estimates in first order model ........ 66
  4.4.2.5 Test for $\beta_1$ ....................................................................................... 66
  4.4.2.6 Test for $\beta_2$ ....................................................................................... 67
  4.4.2.7 Test for $\beta_3$ ....................................................................................... 68
4.4.3 The analysis of variance of GP2G................................................................. 69
  4.4.3.1 The test of significance in regression-GP2G ............................................ 70
  4.4.3.2 Normal probability plot of the residuals GP2G ........................................ 70
  4.4.3.3 The test of significance of the model ...................................................... 71
4.4.4 The analysis of variance of GP3G................................................................. 72
  4.4.4.1 The test for significance in regression-GP3G ............................................ 73
  4.4.4.2 Normal probability plot of the residuals GP3G ........................................ 73
  4.4.4.3 The test of significance of the model ...................................................... 73
4.4.5 Design for fitting the second-order model-GP1G ......................................... 74
  4.4.5.1 Fitting of the second-order model ......................................................... 74
  4.4.5.2 The test of significance of parameter estimates in second order model .... 75
  4.4.5.3 Test for $\beta_i$ ....................................................................................... 75
  4.4.5.4 Test for $\beta_{ii}$ ..................................................................................... 76
  4.4.5.5 Test for $\beta_{ij}$ ..................................................................................... 77
  4.4.5.6 The lack-of-fit test for the three replicates GP1G .................................... 81
4.4.6 Design for fitting the second-order model-GP2G ......................................... 81
  4.4.6.1 The test of significance of parameter estimates in second order model .... 81
  4.4.6.2 Test for $\beta_i$ ....................................................................................... 82
  4.4.6.3 Test for $\beta_{ii}$ ..................................................................................... 82
  4.4.6.4 Test for $\beta_{ij}$ ..................................................................................... 83
  4.4.6.5 The lack of fit for the three replicates in GP2G ....................................... 86
4.4.7 Design for fitting the second-order model GP3G ......................................... 87
4.4.7.1 The test of significance of parameter estimates in second order model...... 87
4.4.7.2 Test for $\beta_i$........................................................................................................... 87
4.4.7.3 Test for $\beta_{ii}$ ........................................................................................................ 88
4.4.7.4 Test for $\beta_{ij}$ ....................................................................................................... 89
4.4.7.5 The lack of Fit for the three replicates in GP3G ................................................. 92

4.5 FINDING THE SETTINGS OF THE EXPERIMENTAL FACTORS THAT PRODUCES THE OPTIMAL RESPONSE.................................................................................................................. 93
4.5.1 Three dimensional surfaces and contour plots .................................................. 93
4.5.1.1 Three dimensional surfaces and contour plots of GP1G ............................ 93
4.5.1.2 Three dimensional surfaces and contour plots of GP2G ............................ 100
4.5.1.3 Three dimensional surfaces and contour plots of GP3G ......................... 106
4.5.2 The canonical analysis ................................................................................... 112
4.5.2.1 The canonical analysis of GP1G ................................................................. 112
4.5.2.2 The canonical analysis of GP2G ................................................................. 113
4.5.2.3 The canonical analysis of GP3G ................................................................. 114
4.5.3 The stationary points ..................................................................................... 115
4.5.3.1 The stationary point-GP1G ................................................................. 115
4.5.3.2 The stationary point-GP2G ................................................................. 117
4.5.3.3 The stationary point-GP3G ................................................................. 118
4.5.4 The ridge analysis .......................................................................................... 120
4.5.4.1 The ridge analysis-GP1G ........................................................................ 120
4.5.4.2 The ridge analysis-GP2G ........................................................................ 121
4.5.4.3 The ridge analysis-GP3G ........................................................................ 122

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS ............................. 124
5.1 INTRODUCTION ......................................................................................................... 124
5.2 ESTIMATION OF THE LINEAR PARAMETERS IN AN EXISTING A-OPTIMUM AND D-EFFICIENT CALCULUS OPTIMUM VALUE SECOND ORDER ROTATABLE DESIGN..................... 124
5.3 OBTAINING THE GENERALIZED VARIANCE OF THE ESTIMATED LINEAR PARAMETERS 126
5.4 FITTING AND TESTING THE THREE MODEL ADEQUACIES .............................. 127
5.5 FINDING THE SETTINGS OF THE EXPERIMENTAL FACTORS THAT PRODUCE THE OPTIMAL RESPONSE.................................................................................................................. 128
5.6 RECOMMENDATION FOR FURTHER RESEARCH .......................................................... 133
REFERENCES .................................................................................................................. 134
APPENDIX 1 .................................................................................................................... 138

1.1 THE CODES FOR GENERATING THE RESPONSE SURFACES USING SAS SOFTWARE...... 138
1.2 THE CODES FOR GENERATING THE CONTOUR PLOTS USING R SOFTWARE .............. 141
1.3 THE CODES FOR GENERATING THE RIDGE ANALYSIS USING SAS SOFTWARE............ 142
1.4 GREENHOUSE PHOTOS ............................................................................................... 143
LIST OF TABLES

Table 3.1: Twenty four point second order rotatable design in three dimensions. .......... 17
Table 4.1: The yield of 30 rose coco beans with straight fertilizers: nitrogen, phosphorus and potassium. ....................................................................................................................................................... 36
Table 4.2: Descriptive Analysis of straight fertilizers of N, P and K on rose coco beans mean yield. ............................................................................................................................................................... 37
Table 4.3: Analysis of variance for the yield of rose coco beans on straight fertilizers of N, P and K of table 4.1. ........................................................................................................................................................................... 37
Table 4.4: The 24-point's rotatable design of coded levels and natural levels with the yield of rose coco beans- GP1G at the ratio 10:20:30 of N : P : K fertilizers. ................................ 40
Table 4.5: The yield of 24-points rose coco beans (from table 4.4) under fertilizers at the ratio of 10 grams of nitrogen, 20 grams of phosphorus and 30 grams of potassium ...... 41
Table 4.6: The 24-point's rotatable design of coded levels and natural levels with the yield of rose coco beans GP2G at the ratio of 20:30:40 N : P : K fertilizers. ......................... 43
Table 4.7: The yield of 24-points rose coco beans (from table 4.6) under fertilizers at the ratio of 20 grams of nitrogen, 30 grams of phosphorus and 40 grams of potassium ... 44
Table 4.8: The 24-point's rotatable design of coded levels and natural levels with the yield of rose coco beans GP3G at the ratio of 30:40:50 N : P : K fertilizers. ......................... 46
Table 4.9: The yield of 24-points rose coco bean (from table 4.8) under fertilizers at the ratio of 30 grams of Nitrogen, 40 grams of Phosphorus and 50 grams of potassium .. 48
Table 4.10: The Analysis of Variance table first order for GP1G. ..................................... 64
Table 4.11: The Analysis of Variance table first order for GP2G. ..................................... 70
Table 4.12: The Analysis of Variance table first order for GP3G. ..................................... 72
Table 4.13: The ANOVA results on the three replicate of rose coco beans-GP1G. .......... 77
Table 4.14: The ANOVA results in the fertilizer concentration on rose coco beans-GP1G ........................................................................................................................................................................... 78
Table 4.15: The estimated effects & coefficients of the empirical model for GP1G ...... 79
Table 4.16: The Analysis of Variance Table for N,P,K in 2nd order model- GP1G. ...... 80
Table 4.17: The lack of fit test for second order model-GP1G. ....................................... 81
Table 4.18: The ANOVA results on the three replicate of rose coco beans-GP2G. ...... 84
Table 4.19: The ANOVA results in the fertilizer concentration on rose coco beans-GP2G. ...................................................................................................................................... 84

Table 4.20: The estimated effects & coefficients of the empirical model for GP2G. ........ 85

Table 4.21: The lack of fit test for GP2G. ........................................................................ 86

Table 4.22: The Analysis of Variance Table for N,P,K in 2nd order model- GP2G. ....... 86

Table 4.23: The ANOVA results on the three replicate of rose coco beans-GP3G. ........ 89

Table 4.24: The ANOVA results in the fertilizer concentration on rose coco bean-GP3G. ...................................................................................................................................... 89

Table 4.25: The estimated effects & coefficients of the empirical model for GP3G. ........ 91

Table 4.26: The lack of fit test for GP3G. ........................................................................ 92

Table 4.27: The Analysis of Variance Table for N,P,K in 2nd order model- GP3G. ....... 92

Table 4.28: The Canonical Analysis of Response Surface, Eigenvectors and Eigenvalues-GP1G. ......................................................................................................................... 112

Table 4.29: The Canonical Analysis of Response Surface, Eigenvectors and Eigenvalues-GP2G. ......................................................................................................................... 113

Table 4.30: The Canonical Analysis of Response Surface, Eigenvectors and Eigenvalues-GP3G. ......................................................................................................................... 114

Table 4.31: The Estimated Ridge of Minimum Response for Variable YIELD-GP1G.. 121

Table 4.32: The Estimated Ridge of Maximum Response for Variable YIELD-GP1G. 121

Table 4.33: The Estimated Ridge of Minimum Response for Variable YIELD-GP2G.. 121

Table 4.34: The Estimated Ridge of Maximum Response for Variable YIELD-GP2G. 121

Table 4.35: The Estimated Ridge of Minimum Response for Variable YIELD-GP3G.. 122

Table 4.36: The Estimated Ridge of Maximum Response for Variable YIELD-GP3G. 122

Table 5.1: Regression coefficients of predicted second order polynomial models for the response of GP1G, GP2G and GP3G. ......................................................................................................................... 124

Table 5.2: The eigenvalues, predicted rose coco yield at the stationary points, and uncoded critical values of fertilizer levels. ...................................................................................................................................... 126

Table 5.3: The analysis of variance of the factors and the critical values of the rose coco yield response ...................................................................................................................................... 127

Table 5.4: Canonical analysis based on coded and actual values, eigenvalues and eigenvectors for the three groups. ...................................................................................................................................... 129
LIST OF FIGURES

Figure 4.1: The plot of mean of yield of rose coco beans under each of the three fertilizers; N, P, K. ......................................................................................................................... 38

Figure 4.2: The yield of rose coco beans under each of the three straight fertilizers; N, P, K of table 4.1 ........................................................................................................................................ 39

Figure 4.3: The average yield of rose coco bean for the three groups. ................................. 49

Figure 4.4: The Normal Probability Plot of GP1G. ................................................................... 65

Figure 4.5: The Normal Probability Plot of GP2G. ................................................................... 70

Figure 4.6: The Normal probability plot for GP3G. .................................................................. 73

Figure 4.7a: The response surface plots for the treatments of nitrogen and phosphorus fertilizer concentrations in GP1G. ......................................................................................... 94

Figure 4.7b: The response surface plots for the treatments of nitrogen and potassium fertilizer concentrations in GP1G. ......................................................................................... 96

Figure 4.7c: The response surface plots for the treatments of phosphorus and potassium fertilizer concentrations in GP1G. ......................................................................................... 98

Figure 4.7a 1: The contour plot for nitrogen and phosphorus fertilizers in GP1G of 4.7a. 95

Figure 4.7a 2: The contour plot of nitrogen and potassium fertilizers in GP1G of 4.7b. 97

Figure 4.7a 3: The contour plot of phosphorus and potassium fertilizers in GP1G of 4.7c. 99

Figure 4.8a: The response surface plots for the treatments of nitrogen and phosphorus fertilizer concentrations in GP2G. ......................................................................................... 100

Figure 4.8b: The response surface plots for the treatments of nitrogen and potassium fertilizer concentrations in GP2G. ......................................................................................... 102

Figure 4.8c: The response surface plots for the treatments of phosphorus and potassium fertilizer concentrations in GP2G. ......................................................................................... 104

Figure 4.8b 1: The contour plot of nitrogen and phosphorus fertilizers in GP2G of 4.8a. 101

Figure 4.8b 2: The contour plot of nitrogen and potassium fertilizers in GP2G of 4.8b. 103

Figure 4.8b 3: The contour plot of phosphorus and potassium fertilizers in GP2G of 4.8c. 105

Figure 4.9a: The response surface plots for the treatments of nitrogen and phosphorus fertilizer concentrations in GP3G. ......................................................................................... 106
Figure 4.9 b: The response surface plots for the treatments of nitrogen and potassium fertilizer concentrations in GP3G. ................................................................. 108

Figure 4.9 c: The response surface plots for the treatments of phosphorus and potassium fertilizer concentrations in GP3G. ................................................................. 110

Figure 4.9 c 1: The contour plot of nitrogen and phosphorus fertilizers in GP3G of 4.9a. ............................................................................................................................................ 107

Figure 4.9 c 2: The contour plot of nitrogen and potassium fertilizers in GP3G of 4.9b. 109

Figure 4.9 c 3: The contour plot of phosphorus and potassium fertilizers in GP3G 4.9c. 111
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Average Variance Criterion</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>C</td>
<td>Information Matrix</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>D</td>
<td>Determinant Criterion</td>
</tr>
<tr>
<td>D_1</td>
<td>The twenty four points second order rotatable design in three dimensions</td>
</tr>
<tr>
<td>E</td>
<td>Eigen-value Criterion</td>
</tr>
<tr>
<td>GP1G</td>
<td>Group 1 greenhouse with N:P:K fertilizer in grams at the ratio of 10:20:30</td>
</tr>
<tr>
<td>GP1GA</td>
<td>the first replicate in GP1G</td>
</tr>
<tr>
<td>GP1GB</td>
<td>the second replicate in GP1G</td>
</tr>
<tr>
<td>GP1GC</td>
<td>the third replicate in GP1G</td>
</tr>
<tr>
<td>GP2G</td>
<td>Group 2 greenhouse with N:P:K fertilizer in grams at the ratio of 20:30:40</td>
</tr>
<tr>
<td>GP2GA</td>
<td>the first replicate in GP2G</td>
</tr>
<tr>
<td>GP2GB</td>
<td>the second replicate in GP2G</td>
</tr>
<tr>
<td>GP2GC</td>
<td>the third replicate in GP2G</td>
</tr>
<tr>
<td>GP3G</td>
<td>Group 3 greenhouse with N:P:K fertilizer in grams at the ratio of 30:40:50</td>
</tr>
<tr>
<td>GP3GA</td>
<td>the first replicate in GP3G</td>
</tr>
<tr>
<td>GP3GB</td>
<td>the second replicate in GP3G</td>
</tr>
<tr>
<td>GP3GC</td>
<td>the third replicate in GP3G</td>
</tr>
<tr>
<td>K</td>
<td>Potassium</td>
</tr>
<tr>
<td>MATLAB</td>
<td>Matrix Laboratory</td>
</tr>
<tr>
<td>Mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>MT</td>
<td>Metric Tonnes</td>
</tr>
<tr>
<td>N</td>
<td>Nitrogen</td>
</tr>
<tr>
<td>P</td>
<td>Phosphorous</td>
</tr>
<tr>
<td>PRESS</td>
<td>Prediction error sum of squares</td>
</tr>
<tr>
<td>RSM</td>
<td>Response Surface Methodology</td>
</tr>
<tr>
<td>SAS</td>
<td>Statistical Analysis System</td>
</tr>
<tr>
<td>T</td>
<td>Trace criterion</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

1.1 Background Information

Generally, the objective for agricultural researchers is that, they are in constant search for new or improved technologies to increase productivity. The researcher is usually interested in finding maximum or optimal yield using minimum cost. For Kenya to meet the needs of its ever growing population, then new technologies is inevitable in agricultural fields. Since arable land in Kenya is a fraction, there’s need to produce maximally in such areas by utilizing design of an experiment like the one in this study. There are two types of supplies for agriculture, specifically fertilizer and pesticides. It can be said that the fertilizer is food and pesticides is medicine for plants in conventional agriculture (Arjumand, 2013). Soil fertility is diminishing gradually due to the erosion, loss of nutrients, accumulation of salts and other toxic elements and unbalanced nutrients compensation. Many efforts are being exercised to combat the adverse consequences of chemical farming (Faheed, 2008). During the last decades rose coco bean is becoming increasingly important legume for human nutrition and a major protein and calorie source in the world. The bean crop requires nitrogen in quite high amount in the first stage of development for the emergence of the nodules and builds up of the symbiotic nitrogen fixation. The amount of nitrogen which symbiotically bound depends on the kind of plant, the efficiency of the bacteria inoculated and soil properties (Bildirici and Yilmaz; 2005). Common bean (Phaseolus vulgaris), also referred to as dry beans, is an annual leguminous plant that belongs to the genus, Phaseolus, with pinnately compound trifoliate large leaves. It is largely a self-pollinated plant, though cross-pollination is possible if the stigma contacts with pollen coated bee when extended. Seeds are non-endospermic and vary greatly in size and colour from the small black wild type to the large white, brown, red, black or mottled seeds of cultivars, which
are 7-16 mm long (Cobley and Steele, 1976). The common *rose coco* beans (*Phaseolus vulgaris*) are the most commonly grown grain legumes that come second after maize as a subsistence crop (GoK, 2006). They serve as a source of protein which is relatively cheaper compared to animal proteins to the majority of the population in Kenya (Karanja et al., 2007). Consumption of common bean is high mostly because it is relatively inexpensive compared to meat (Pachico, 1993). Common bean plays a strategic role in alleviating malnutrition. Regular consumption of common bean is now promoted by health organizations because it reduces the risk of diseases such as cancer, diabetes or coronary heart diseases (Leterme and Munoz, 2002). This is because common bean is low in fat and is cholesterol free. It is also an appetite suppressant because it digests slowly and causes a low sustained increase in blood sugar. Bean provides a rich combination of carbohydrates (60-65%), proteins (21-25%), fats (less than 2%), vitamins and minerals (Ensminger *et al.*, 1994). In fact, with increasing health concerns, most people, especially the urban population are reducing consumption of animal proteins, and instead they are turning to pulses such as dry bean due to its low fat content. Hence the rationale for emphasis in more bean research is self-evident. The crop also provides farm households and traders with incomes and is therefore important from both the food security and income-generation. Hence there is a need for increased bean production to enhance exports as well as satisfies domestic market. It is an important staple food in the diet of people of all income categories. The beans are characterized as near perfect food because of their high protein content and generous amounts of iron, folic acid, complex carbohydrates and other diet essentials. The crop also matures within three months, enabling farmers to plant the crop almost three times annually. Therefore, we need to see this important crop doing well, so as to feed and supplement the right nutrients to its ever growing consumers.
The response surface methodology (RSM) emphasizes on finding a particular treatment combination, which causes the maximum or minimum responses. The use of analysis of the quadratic response function or RSM is necessary to obtain the optimum level of fertilizer requirements. In response surface analysis, the eigenvalues could be used to determine whether the solution gives a maximum, minimum or saddle point on the response curve. Moreover, the effects of treatment combinations, which have not been carried out in the experiment might still be estimated. The regression equation was fitted between the response variable, *rose coco* yield and the three fertilizer treatments, nitrogen (N), phosphorus (P) and potassium (K). The expected yield could be described as a continuous function of the application rate factor. When the fertilizer application rates are higher or lower than the optimum application rates, they might result in a reduction in yields. The purpose of implementing this RSM technique was to determine the optimum levels of fertilizer used in order to optimize *rose coco* yields.

In any treatment arrangement, we sought a treatment or treatment combination that could be used to either reverse existing methodologies, advance the current methods being used in farming in order to maximize the yield using the scarce resources available. In this research, the investigation geared towards searching the optimum combination of treatment combinations to maximize *rose coco* beans in a greenhouse setting using inorganic fertilizers was highlighted. It is seen how different components of fertilizers affect the output of the beans. The study therefore was used to determine; the best possible level of the identified fertilizer to maximize the yield in *rose coco* beans. In the technological context, this response surface methodology was applied to study the measured yield or output of a system as it varies in response to the changing levels of one or more physical input variables. The experimental design aspect deals
with the choice of suitable variables and their various levels, and regression analysis enables a mathematical form to be fitted to the observations to "model" the varying yields.

In order to conduct a first-order model, the operating conditions that maximize the yield of *rose coco* beans were studied. There are three independent variables which influence the process yield: *nitrogen*, *phosphorus* and *potassium* fertilizers. The low-order polynomial terms has been used to describe some part of the response surface. Once the estimated equation was obtained, we were able to use statistical techniques to check for the model adequacy. The objective was to determine the current levels or settings of the *nitrogen*, *phosphorus* and *potassium* fertilizers that resulted in a value of a response that was optimal or close to the optimum. When the first order model was inadequate, then second-order model was fitted. With the purpose of exploring the second-order model, the statistical modeling techniques to develop an appropriate approximating relationship between the yield and the process variables was used.

### 1.2 Statement of the problem

The beans are the second commonly grown grain legumes after maize as a subsistence crop. Most farmers plant the crop without fertilizers or they apply the available fertilizer without considering the plant optimal requirement for the optimal production, when the fertilizer application rates are higher or lower than the optimum application rates, reduction in yield is experienced. The fertilizers are wasted if the amount applied is more than the optimum rate and can be harmful to the soils, (Khamis, 2006). Kenya’s current annual bean production of approximately 215,000 MT, barely meets half the annual consumption of 450,000 MT. The average production per hectare is 500 kg or less, compared to 1800 to 2000 kg ha\(^{-1}\) potential (Africa Agriculture, 2008). The major limitation to bean production in many smallholder farms
is declining soil fertility as a result of continuous cropping with minimal inputs or rotation to replenish soil nutrients. Furthermore, there is minimal or no optimal use of the required inorganic fertilizers. The over use of N (nitrogen) relative to P\textsubscript{2}O\textsubscript{5} (phosphorus) and K\textsubscript{2}O (potassium) has raised concerns in environmental perspective. The phosphorus and potassium fertilizers have been in short supply and farmers have been more steadily adopting the use of nitrogenous fertilizers because of impressive response. There is evidence that soil P\textsubscript{2}O\textsubscript{5} and K\textsubscript{2}O level are declining. So determining the optimum balance of N, P\textsubscript{2}O\textsubscript{5} and K\textsubscript{2}O so as to produce high yield of *rose coco* has been of concern that the study tried to address. In the year 2014 and 2015, maize was majorly affected by a disease which swept the large acreage of farms, thus the need to embark on the research for the alternative crop that is *rose coco* bean. The study was geared towards estimating the linear parameters in one of the existing six specific second order rotatable designs in three dimensions in which Mutiso J.M.(1998) calculated the calculus optimal value to be 1.1072569 for the free/letter parameter (f), Koech, F.(2016) calculated the relative efficiencies for the six designs and their optimality criteria, Koech showed that the twenty four point second order rotatable design was the most D-efficient and A-optimal design. Therefore, out of these researches we dwelled on the twenty four points, second order rotatable design and proceeded to have a practical greenhouse experiment to realize the optimal and efficient production of *rose coco* bean using the three fertilizer components, namely, nitrogen, phosphorus and potassium. The study facilitated the estimation of the coefficient $\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$, $\beta_{11}$, $\beta_{22}$, $\beta_{33}$, $\beta_{12}$, $\beta_{13}$ and $\beta_{23}$ in an existing second order rotatable design of twenty four points in three dimensions. Achieving the production efficiency of *rose coco* has not been easy that this research tends to employ twenty four points second order rotatable design approach, focusing on the area of modeling *rose coco* yield production and highlighting the respective point of
intercept and gradient levels of fertilizer with which the variety of *rose coco* was capable of delivering the yield production efficiency. The study found out how the yields fitted the second order model and checked its model adequacy. The purpose of implementing the twenty four points’ rotatable design was to determine the optimum levels of fertilizer used in order to optimize *rose coco* yields.

Therefore the problem was to achieve an optimal and efficient production of *rose coco* beans through the twenty four point second order rotatable design in three dimensions for parameter estimation.

### 1.3 Objectives of the study

This section highlights the general objective and specific objectives.

#### 1.3.1 General objective

The general objective was to produce *rose coco* beans optimally and efficiently using an existing A-optimum and D-efficient twenty four point second order rotatable design in three dimensions.

#### 1.3.2 Specific objectives

1) To estimate the linear parameters in an existing A-optimum and D-efficient calculus optimum value second order rotatable design.

2) To obtain the generalized variance of the estimated linear parameters.

3) To fit and test the three models adequacies.

4) To find the settings of the experimental factors that produces the optimal response.
1.4 Significance of the study

The major nutrients, that is, nitrogen (N), phosphorus (P) and potassium (K) are essential for growth and production of bean, (Marschner, 2012). There is unbalanced nutrients compensation in the soil. Therefore, the need to determine the optimum balance of N, P and K so as to realize the maximum potential yield of *rose coco* beans. The study focused on the estimated linear parameters of a specific twenty four point second order rotatable design in three dimensions for the calculus optimal value for the first time, checked the adequacy of the models and obtained the generalized variance for the estimated linear parameters. Furthermore, the response surface for the optimal response yield of *rose coco* bean was found for the three groups. The study would provide the most reliable advice for policy makers on the right range application of fertilizers that are desirable to optimize/maximize the *rose coco* bean crop production.

1.5 Limitation of the study

The *rose coco* beans were planted simultaneously in a greenhouse (15mx10m) to guard soil contamination from the subsequent planting, we could have had three planting periods at different seasons say January-April, May-August and September-December, or instead built two or three greenhouses, this could have allowed the independent variables to be screened to get the amount of each component of the fertilizers. The choice of each fertilizer starting point in this study was done arbitrarily; there were so many infinite options available at our disposal. The top dressing and foliar spray was not applied, only one time initial fertilizer application was done during planting and left till harvest was done.
CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter presents some literature relevant to the study objectives. The subsections included are; the rose coco beans production in Kenya, organic and inorganic use of fertilizers, response surface methodology, optimal design criteria, D-optimality, Model adequacy.

2.2 The rose coco bean production in Kenya

Kenya’s current annual bean production of approximately 215,000 metric tonnes (MT), barely meets half the annual consumption of 450,000 MT. The average production per hectare is 500 kg or less, compared to 1800 to 2000 kg ha\(^{-1}\) potential (Africa Agriculture, 2008). The major limitation to bean production in many smallholder farms is declining soil fertility as a result of continuous cropping with minimal inputs or rotation to replenish soil nutrients. Furthermore, there is minimal or no use of the required inorganic fertilizers. Some of the options that are currently being pursued to address low soil fertility include integrated use of organic (e.g., crop residues, animal manures, agro forestry tree pruning) and inorganic (fertilizers) resources. Common beans are grown in pure stands by large scale farmers, but commonly intercropped with maize by smallholder farmers (Kiménjú, 2004). Like many legumes, beans thrive well in sufficiently aerated and well drained soils with a pH of 6.5 - 7.5 because they are very sensitive to soil acidity and an optimal amount of organic carbon above 2.4 percent (Baudoin et al., 2001). The optimum altitude should exceed 1000 m above sea level. Most common bean cultivars are short-season crops with a maturity period ranging from 65 to 110 days from emergence to physiological maturity, hence two bean harvests per year (Buruchara, 2007).
Common bean production in Kenya is mainly in the highlands and the midlands. About 75 percent of the annual cultivation occurs in four regions, namely: Rift Valley, Western, Nyanza, and Eastern Province (Wortmann, 1998). The Rift Valley contributes the biggest share, accounting for 33 percent of the national production followed by Nyanza and Western province accounting for 22 percent each (Katungi, 2010). Output from the eastern parts of the country and the coast is constrained by adverse climatic conditions. Although Kenya has two seasons for common bean, a significant number of farmers grow the crop once a year because of adverse climatic conditions. In Kenya, some of the Andean beans that are widely cultivated include the red mottled (occurring in different local names such as Rose coco or GLP 2 and Nyayo), red kidneys (such as Canadian Wonder or GLP 24), purple/gray speckled (locally known as Mwezi Moja). *Rose coco* type is the most widely grown followed by Canadian Wonder type. *Rose coco* and *Canadian Wonder* type are high yielding but require heavy rains and high soil fertility to yield well. Soil fertility is diminishing gradually due to the erosion, loss of nutrients, accumulation of salts and other toxic elements and unbalanced nutrients compensation. *Rose coco* bean (*Phaseolus vulgaris*), is an important legume for human nutrition and a major protein and calorie source. Bean production is declining in Kenya due to various factors; virus diseases and optimum combination of the needed component of fertilizers are some of the major yield reduction factors in bean production. The *rose coco* hardly meets the demand; therefore the need to search for improved methods.

### 2.3 Organic and inorganic use of fertilizers

The animal manures are valuable sources of nutrients and the yield-increasing effect of manure is well established. Apart from the nutrients in manure, its effects on the improvement of soil
organic matter, soil structure and the biological life of the soil are well recognized. There is also some evidence that it may contain other growth-promoting substances like natural hormones and B vitamins (Leonard, 1986). Plants can only use the nutrients that are in an inorganic form. Manure N and P are present in organic and inorganic forms, and are not totally available to plants. The organic forms must be mineralized or converted into inorganic forms over time before they could be used by plants. The organic manure could then be applied to plants as a source of nitrogen, phosphorus and potassium (N, P & K) which are the macro nutrients that limit crop growth (Kwabiah, 2003; Wasonga, 2008). The term “organic” describes production systems that optimize natural processes. Organic farming systems rely on ecologically-based practices such as cultural and biological pest management, and virtually exclude the use of synthetic chemicals in crop production and prohibit the use of antibiotics and hormones in livestock production. The chemical fertilizers are used in modern agriculture to correct known plant-nutrient deficiencies; to provide high levels of nutrition, which aid plants in withstanding stress conditions; to maintain optimum soil fertility conditions; and to improve crop quality. Adequate fertilization programs supply the amounts of plant nutrients needed to sustain maximum net returns (Leonard, 1986). In essence, fertilizers are used to make certain that soil fertility is not a limiting factor in crop production.

2.4 Response Surface Methodology

The response Surface Methodology (RSM) is a collection of mathematical and statistical techniques useful for the modeling and analysis of problems in which a response of interest is influenced by several variables and the objective is to optimize this response (Montgomery 2005). In most RSM problems, the form of the relationship between the response and the
independent variables is unknown. Thus the first step in RSM is to find a suitable approximation for the true functional relationship between the response and the set of independent variables which are subject to the control of the scientist or engineer. Usually, a low-order polynomial in some region of the independent variables is employed. If the response is well modeled by a linear function of the independent variables, then the approximating function is the first-order model and if there is curvature in the system, then a polynomial of higher degree such as the second order is used. The broad aims of the RSM are to investigate the nature of the response surface over a region of interest and to identify operating conditions associated with maximum or minimum responses. RSM is a sequential procedure and is generally conducted in three phases, as emphasized in (Myers and Montgomery, 2002). Phase 1 involves the screening of explanatory variables when the number of factors is large or when experimentation is expensive to identify those which have a significant effect or the most influential on the response(s) being investigated; phase 2 is concerned with the location of optimum operating conditions by conducting a sequence of suitable experiments; and phase 3 involves the fitting of an appropriate empirical model, usually a second-order polynomial model, in order to examine the nature of the response surface in the vicinity of the optimum. In order to get the most efficient result in the approximation of polynomials the proper experimental design must be used to collect data. Once the data are collected, the method of least squares was used to estimate the parameters in the polynomials. The response surface analysis was performed by using the fitted surface. The response surface designs are the types of designs for fitting response surface. The fundamental methods for quantitative variables involve fitting first-order (linear) or second-order (quadratic) functions of the predictors to one or more response variables, and then examining the characteristics of the fitted surface to decide
what action was appropriate. The $\beta$ could be estimated by using the least squares method as: $\beta = (x'x)^{-1} x'y$. The measure of accuracy of the column of estimators, $\beta$, is the variance-covariance matrix which is defined as; $V(\beta) = \sigma^2 (x'x)^{-1}$ (Myers, 1971), (Unal, 1996), (Craig, 1978), (Mitchell, 1974) where $\sigma^2$ is the variance of the error. The $V(\beta)$ matrix is a statistical measure of the goodness of fit. The $V(\beta)$ is a function of $(X'X)^{-1}$ and therefore, one would want to minimize $(X'X)^{-1}$ to improve the quality of the fit. Statisticians have shown that minimizing $(X'X)^{-1}$ is equivalent to maximizing the determinant of $X'X$ (Mitchell, 1974, Montgomery, 1991, Box and Draper, 1974). Therefore, generating a design matrix which enables the construction of a good least square approximation model translates to maximizing the determinant of the $X'X$ matrix and experimental designs that maximize $|X'X|$ are referred to as D optimal designs (Unal, et al 1996, Craig, 1978, Box and Draper, 1974) Here, “D” stands for the determinant of the $X'X$ matrix associated with the model. This analysis could easily be extended to the quadratic model with the same conclusions for D-optimality.

### 2.4.1 Optimal Design Criteria

The design optimality criteria are often called the alphabetical optimality criteria because they are named by some of the letters of the alphabet. Kiefer and Wolfowitz (1959) were among the first authors who developed these optimality criteria. There are many types of optimal criteria available to generate experimental designs which include the determinant criterion D-; the average-variance criterion A-; the eigenvalue criterion E-; and the trace criterion T-;.

Concerning optimality, the smallest value among the matrix means is usually taken thus the identification of optimality criteria. D-optimum design minimizes the content of the ellipsoidal confidence region for the parameter of the linear model. In terms of eigenvalues D-minimizes
the generalized variance of the parameter estimates. A-optimality minimizes the sum (or average) of the variance of the parameter estimates (Atkinson and Donev 1992). The smallest among the criteria determines the optimality for the generalized variance of the parameter estimates of the information matrix. These criteria measure the desirability of a design, E-reduces the variance of each individual parameter estimate, and T-optimum design is one that has not enjoyed much use because of the linearity aspect of the T-criterion. Evaluation of these criteria showed that the more homogeneous the design the higher optimal it becomes. The ultimate purpose of any optimality criterion is to measure ‘largeness’ of a non-negative definite information matrix $C_{sxs}$. At this point, the specific optimality values of the design by utilizing the methods of evaluation of the particular criteria as outlined by Pukelsheim (1993).

The determinant criterion, $D$, $\phi_0(C) = (\text{det } C)^{1/2}$.

The average variance criterion, $A$, $\phi_1(C) = \left(\frac{1}{s} \text{trace } C^{-1}\right)^{-1}$ (if C is positive definite)

The consideration is a three dimensional specific rotatable design of order two. The relation $C = (K'M^{-1}K)^{-1}$ is used for the second order model, where $M = 1/N(X'X)$ is the moment matrix. The information matrix is utilized in the determination of the optimality criteria and the exact values of the criteria.

\subsection*{2.4.2 D-optimality}

When considered historically, D-optimality (Kiefer, 1958) was the first alphabetical optimality criterion developed. It is also still among the most popular because of its simple computation, and the many available algorithms. In D-optimal designs, the goal is to minimize the generalized variance of the regression model parameter estimates. This is achieved by
maximizing the determinant of the information matrix, $|X'X|$. It turns out that the information matrix is inversely proportional to the variance-covariance matrix of the least squares estimates of the linear parameters of a model. Thus, by maximizing the information matrix, the generalized variance of the model parameters is minimized. The twenty four point second order rotatable design is $D$-optimal.

The focus of $D$-optimality is an estimation of model parameters through good attributes of the moment matrix, which is defined as $M = \frac{X'X}{N}$ where $X'X$ is the information matrix, and $N$, the total number of runs, is used as a penalty for larger designs. $D$-optimality requires one to maximize the determinant of the moment matrix, that is, a $D$-optimal design is the design, $D^*$, in the design space $W$ such that the determinant of $M$, $|M(D^*)| = \max_{D \in \Omega} |M(D)|$ Under the standard normality assumptions, $|X'X|$ is inversely proportional to the square of the volume of the confidence region for the regression coefficients. The larger the determinant of $X'X$ then better the estimation of the model parameters.

### 2.5 Model adequacy

The large value of $R^2$ does not necessarily imply that the regression model is good one, adding a variable to the model always increases $R^2$, regardless of whether the additional variable is statistically significant or not. Thus, it is possible for models that have large values of $R^2$ to yield poor predictions of new observations or estimates of the mean response. Because $R^2$ always increases as we add terms to the model, some regression model builders prefer to use an adjusted $R^2$ statistic defined as

$$R_{adj}^2 = 1 - \frac{SS_E/(n-p)}{SS_T/(n-1)} = 1 - \frac{n-1}{n-p}(1-R^2)$$

(2.1)
In general, the adjusted $R^2$ statistic is not always increasing as variables are added to the model. In fact, when unnecessary terms are added, the value of $R_{adj}^2$ often decreases. When $R^2$ and $R_{adj}^2$ differ dramatically, there is a good chance that non-significant terms have been included in the model. We are frequently interested in testing hypotheses on the individual unknown parameters. Such tests would be useful in determining the value of each of the regressor variables in the response surface model. For example, the model might be more effective with the inclusion of additional variables, or perhaps with the deletion of one or more of the variables already in the model. Adding a variable to the regression model always causes the sum of squares for regression to increase and the error sum of squares to decrease. We must decide whether the increase in the regression sum of squares is sufficient to warrant using the additional variable in the model. Furthermore, adding an unimportant variable to the model can actually increase the mean square error, thereby decreasing the usefulness of the model (Carley et al, 2004).

The significant terms in the model were found by analysis of variance (ANOVA) for each response. The significance was judged by determining the probability level that the F-statistic calculated from the data was less than 5%. The model adequacies were checked by $R^2$, adjusted-$R^2$. The analysis included examining normal probability plots. The maximization and minimization of the polynomials thus fitted was usually performed by the mapping of the fitted responses.
CHAPTER 3: METHODOLOGY

3.1 Introduction

This chapter gives explanations on how each one of the four specific objectives were achieved. The research was done with the combination of independent treatment factors of inorganic fertilizers to check how factors influence the optimal yield of rose coco beans. The experiment was carried out in a greenhouse of size 15m x 10m during the period of February and July 2016 on the twenty four points second order rotatable design in three dimensions.

3.2 Estimation of the linear parameters in an existing A-optimum and D-efficient calculus optimum value second order rotatable design.

This section gives the twenty four points’ calculus optimum value experimental design, screening experiment, experiment layout, planting rose coco beans and the method of estimating the parameters.

3.2.1 Twenty Four Points Calculus Optimum Value Experimental Design

The design of twenty four points second order rotatable design has the free parameter(s) \( f \) which has been determined using differential calculus by Mutiso J.M (1998). The study was geared towards estimation of the linear parameters in one of the existing six specific second order rotatable designs in three dimensions in which Mutiso J.M (1998) calculated the calculus optimal value to be 1.1072569 for the free/letter parameter, the other values of \( c_1 \) and \( c_2 \) were given as 0.7829487 and 1.2735263 respectively. We went further to obtain the yield for the calculus optimal value design in the greenhouse setting after which linear parameters were
estimated. We consider a set of twenty four point’s rotatable designs as highlighted by Mutiso (1998) and Koske et al (2008) is given as:

\[ D_1 = \left[ \frac{1}{2} G(f, f, 0) + \frac{1}{4} G(c_1, 0, 0) + \frac{1}{4} G(c_2, 0, 0) \right] \]  

(3.1)

Table 3.1: Twenty four point second order rotatable design in three dimensions.

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-f</td>
<td>-f</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>f</td>
<td>-f</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-f</td>
<td>f</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>f</td>
<td>f</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-f</td>
<td>0</td>
<td>-f</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-f</td>
<td>0</td>
<td>f</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>f</td>
<td>0</td>
<td>-f</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>f</td>
<td>0</td>
<td>f</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>-f</td>
<td>-f</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>-f</td>
<td>f</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>f</td>
<td>-f</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>f</td>
<td>f</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>c1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-c1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>c1</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>0</td>
<td>-c1</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>c1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>-c1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>c2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>-c2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>c2</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>0</td>
<td>-c2</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>c2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>-c2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

3.2.2 Screening experiment

The critical preliminary step was the screening stage at which we sought to identify the key list of factors that influence the bean production process. The goal of a screening experiment was to identify those factors which had an influence on the response, and then work with those significant factors in the study. The three straight fertilizers (nitrogen, phosphorus and potassium) of inorganic were screened, to determine the amount each could contribute to
optimal yield of *rose coco* beans. Each one of the straight inorganic fertilizer was applied at the following amounts; 10 grams, 20 grams, 30 grams, 40 grams, and 50 grams per hole to *rose coco* plant. We then checked what level of inorganic fertilizer gave maximum yield independently. Thereafter, the constructed rotatable design of twenty four points was used to find out the maximum optimal combination of the treatment factors.

3.2.3 Experiment Layout

This section gives the experimental layout of GP1G, GP2G and GP3G using the inorganic fertilizers.

3.2.3.1 Layout of calculus optimum value design

The *rose coco* beans were subjected to inorganic fertilizers N, P, K at different levels of a twenty four point second order rotatable design according to table 3.1. Therefore, each of the three groups (that is, GP1G, GP2G and GP3G) was given three replications each, for example in GP1G we had GP1GA, GP1GB and GP1GC.

3.2.3.2 Group 1 Greenhouse-(GP1G)

In GP1G a combination of 10 grams of nitrogen, 20 grams of phosphorus and 30 grams of potassium were the initial fertilizers applied to *rose coco* beans and acted as the center point. This group was given the straight N, P & K fertilizers, such that GP1GA, GP1GB and GP1GC were the three replications with each having twenty four *rose coco* plants.
3.2.3.3 Group 2 Greenhouse-(GP2G)

In GP2G a combination of 20 grams of nitrogen, 30 grams of phosphorus and 40 grams of potassium were the initial fertilizers applied to *rose coco* beans and acted as the center point. This group was given the straight N, P & K fertilizers, such that GP2GA, GP2GB and GP2GC were the three replications with each having twenty four *rose coco* plants.

3.2.3.4 Group 3 Greenhouse-(GP3G)

In GP3G a combination of 30 grams of nitrogen, 40 grams of phosphorus and 50 grams of potassium were the initial fertilizers applied to *rose coco* beans and acted as the center point. This group was given the straight N, P & K fertilizers, such that GP3GA, GP3GB and GP3GC were the three replications with each having twenty four *rose coco* plants.

In the study some of the cases (plants) were lost due to some environmental factors such as drying up or any other cause, so dead plants were replaced to allow the study to go on. The yields for the *rose coco* plant that died after one month from the day of planting were not obtained.

3.2.4 Planting *Rose coco* Beans

In the greenhouse, organic matter on the soil surface was cleared, and then prepared by *Jembe* ploughing followed by harrowing until fine tillage was obtained. The *rose coco* had three block groups with each group having three replications of twenty four design points. Certified, viable and uniform seeds of *rose coco* beans were planted in the plots in February 2016. Beans were planted as a pure stand in a greenhouse. Before planting, the bean seeds were dressed with Aldrin at the rate of 5g per kg of seeds, to control soil pests especially bean fly
(Melanargromyza phaseoli). Furadan (5% carbofuran) were applied at a rate of 2 g m\(^{-1}\) in the bean rows at sowing to control cutworms (Agrotis ipsilon). The bean seeds were surface sterilized using 3% sodium hypochlorite and pre-germinated on a nutrient free agar media before planting. Bean plant population density was 24 plants per line, meaning 72 plants per block of three replications (75 cm between row spacing and an intra-row spacing of 30 cm). All the plots were given a drip irrigation using drip line pipes. The plots received inorganic fertilizers at different combination. The inorganic fertilizers (nitrogen, phosphorus, potassium) combinations were applied in each experimental unit before planting two seeds (pre-germinated) of rose coco beans, and one week after germination, they were thinned to one plant per experimental unit. The twenty four rose coco bean plants per plot were retained after thinning. The first weeding was carried out on the greenhouse at two weeks after emergence. The second weeding was carried out four weeks later. The third weeding was done before flowering six weeks later, no other supplement were added either by top dressing or foliar spray.

3.2.5 Data collection

The harvesting of each rose coco plant was carried out after 85 days in stages to ensure pods mature well and it was done in a timely manner to avoid self-explosion, the harvest was placed in labeled white polythene bags. The yield of each harvest was weighed using sensitive weighing scale. Careful recording of the harvest was done to ensure correct yield for each rose coco plant crop was captured. The labeling was to assist for easy administration and management of the plants and the yields. The bean pods were shelled and grains dried to 12.5%
moisture content, after which weighing was done for each *rose coco* plant and yield recorded in grams for further analysis.

### 3.2.6 Method of estimating the parameters

In order to get the most efficient result in the approximation of polynomials the proper experimental design must be used to collect data. Once the data are collected, the method of least squares was used to estimate the parameters in the polynomials. The representation $y = \mathbf{x}\beta + \varepsilon$ was given where $y$ is a vector of observations, $\varepsilon$ is the vector of errors, $\mathbf{x}$ is the design matrix and $\beta$ is a vector of unknown model coefficients. The design matrix was a set of combinations of the values of the coded variables, which specifies the settings of the design parameters to be performed during experimentation. The $\beta$ could be estimated by using the least squares method as: $\beta = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'y$.

### 3.3 Obtaining the generalized variance of the estimated linear parameters

This section gives residual sum of squares and variance of parameter estimates.

#### 3.3.1 Residual sum of squares and variance

The method of least squares produces an unbiased estimator of the parameter $\beta$ in the multiple linear regression models. The important parameter was the sum of squares of the residuals

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2 = e'e$$

**(3.2)**

Because $\mathbf{X}'\mathbf{X}\beta = \mathbf{X}'y$, we could have a computational formula for $SSE$:

$$SSE = y'y - \mathbf{b}'\mathbf{X}'y$$

**(3.3)**

equation 3.3 is called the error or the residual sum of squares.
The unbiased estimator of $\sigma^2$ is

$$\sigma^2 = \frac{SSE}{n - p} \quad (3.4)$$

where $n$ is a number of observations and $p$ is a number of unknown parameters. Optimal experimental designs are a class of experimental designs that is optimal with respect to some statistical criterion of choice. The statistical criterion is in relation to any type of variance that is minimized, such as the generalized variance in the model parameter estimates or variance in predicted values of the response variable. Optimizing with respect to some statistical criterion allows the parameters to be estimated without bias with minimum variance or increase the precision of predicted values. The goal of optimal design was to eliminate the option of conducting a non-optimal experimental design, which would result in an increased number of experimental runs to estimate parameters with the same amount of precision compared to an optimal design (Wikipedia contributors). In general, each of the criteria deal with some aspect of either generalized variance of the model parameters or with the variance of predicted values.

### 3.4 Fitting and testing the three models adequacies

This section gives the design of fitting the first order model, the second order model, the model adequacy and residual analysis.

#### 3.4.1 Design of fitting the first-order model

In most RSM problems, the true response function is unknown; we therefore need to approximate the response function. In order to develop an appropriate approximation for response function, we model the data by starting with a first-order polynomial. When the response could be defined by a linear function of independent variables, that is, there is no
curvature then the approximating function is a first-order model. The linear regression model with $k$ independent variables takes the form:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i, \quad (i = 1, 2, 3, \ldots, n)$$

$$y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_{ij} + \varepsilon_i \quad (3.5)$$

The parameter $\beta_i$ measures the expected change in response $y$ per unit increase in $x_i$ when the other independent variables are held constant.

The eventual objective of RSM is to determine the optimum operating conditions for the system, or to determine a region of the factor space in which the operating specifications are satisfied. The general model relationship with $X_1, X_2, \ldots, X_k$ independent variables is given by

$$y = f(X_1, X_2, \ldots, X_k) + \varepsilon; \quad (3.6)$$

where the form of the true response function $f$ is unknown, and $\varepsilon$ is a term that represents other sources of variability not accounted for in $f$. Usually $\varepsilon$ includes effects such as measurement error in the response, background noise, the effect of other variables, and so on. Usually $\varepsilon$ is treated as a statistical error, often assuming it to have a normal distribution with mean zero and variance $\sigma^2$. Then

$$E(y) = \bar{y} = E[f(X_1, X_2, \ldots, X_k)] + E(\varepsilon) = f(X_1, X_2, \ldots, X_k); \quad (3.7)$$

The variables $X_1, X_2, \ldots, X_k$ in equation (3.7) are usually called the natural variables, because they are expressed in the natural units of measurement. In much RSM work it is convenient to transform the natural variables to coded variables $x_1, x_2, \ldots, x_k$, which are usually defined to be dimensionless with mean zero and the same standard deviation (Myers et al. 1989). In terms of the coded variables, the response function (3.7) is written as

$$\hat{y} = f(x_1, x_2, \ldots, x_k); \quad (3.8)$$
The successful use of RSM is critically dependent upon the experimenter’s ability to develop a suitable approximation for \( f \), because the form of the true response function \( f \) is unknown, we must approximate it. Usually a low-order polynomial in some relatively small region of the independent variable space is appropriate. In many cases, either a first-order or a second order model is used. The first-order model is likely to be appropriate when the experimenter is interested in approximating the true response surface over a relatively small region of the independent variable space in a location where there is little curvature in \( f \). In the case of three independent variables, the first-order model in terms of the coded variables is

\[
\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3; \tag{3.9}
\]

The form of the first-order model in equation (3.9) is sometimes called a main effects model, because it includes only the main effects of the three variables \( x_1 \), \( x_2 \) and \( x_3 \). If there is an interaction between these variables, it can be added to the model easily as follows:

\[
\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 \tag{3.10}
\]

This is the model with interaction. The addition of the interaction term introduces curvature into the response function. When the curvature in the true response surface is strong enough then the first-order model is inadequate. A second-order model is likely required in these situations.

The general expected first-order model is given by:

\[
\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k \tag{3.11}
\]

The low and high factor settings are coded as negative and positive, the midpoint coded as 0.

Finally, let’s note that there is a close connection between RSM and linear regression analysis.

For example, consider the model

\[
y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \epsilon \tag{3.12}
\]

The \( \beta \)'s are a set of unknown parameters.
The experimenter tries to quantify the relationship between a set of \( k \) predictor variables \( X' = (X_1, X_2, \ldots, X_k) \) and the response variable \( y \). Often the goal of the experiment has been to maximize or minimize \( E(y) \), the expected value of the response. In most cases, the \( X_i \) are transformed into coded \( x_i \) by \( (X_i - X_{i0})/(sc)_i \) \( i = 1, 2, \ldots, k \) where \( X_{i0} \) and \( (sc)_i > 0 \) are the centering and scaling constants, respectively.

The model in equation 3.5 could be represented in matrix form as;

\[ Y = X\beta + \varepsilon \]  

\[ Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \]

where,

\( Y \) is an \( (n \times 1) \) vector of observations,

\( X \) is an \( (n \times k) \) design matrix,

\( \beta \) is a \( (k \times 1) \) vector of unknown parameters, and

\( \varepsilon \) is a \( (n \times 1) \) vector of independently identically distributed random variables with mean zero and variance \( \sigma^2 \), (Montgomery 2005).

When the \( X'X \) is invertible, that is, there’s a determinant, then the linear system of equation (3.13) has a unique least squares solution given by;

\[ \hat{\beta} = (X'X)^{-1}X'y \]  

The total sum of squares is
$$SST = y'y - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} = \sum_{i=1}^{n} y_i^2 - \frac{\sum_{i=1}^{n} y_i}{n} \quad (3.16)$$

3.4.2 The second-order model

With the purpose of exploring the second-order model, we intend to use a statistical modeling to develop an appropriate approximating relationship between the yield and the process variables: the nitrogen, phosphorus and potassium of inorganic fertilizers. We fitted a first order model, and then we justify it with the second-order model. The second-order model is very flexible; consequently we could experiment with a wide variety of functional forms. The method of least squares was used to estimate the parameters. Once an appropriate approximating model was obtained, it was analyzed to determine the optimum conditions. When we found curvature in the response surface the first-order model was insufficient. The second-order model was useful in approximating a portion of the true response surface with curvature. The second-order model includes all the terms in the first-order model, and quadratic and cross product terms. It is usually represented as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \beta_{11} x_1^2 + \ldots + \beta_{kk} x_k^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \ldots + \beta_{k-1,k} x_{k-1} x_k + \varepsilon \quad \text{or}$$

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i<j}^{k} \beta_{ij} x_i x_j + \varepsilon \quad (3.17)$$

where $\beta_i$ is an unknown parameter and $\varepsilon$ is a random error.

The general expected second-order model is given by:

$$\overline{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=j}^{k} \beta_{ii} x_i^2 + \sum_{i<j}^{k} \beta_{ij} x_i x_j \quad (3.18)$$
where \( \hat{y} \) is the measured response, \( \beta_0 \) is the intercept term, \( \beta_i \) are the linear coefficients, \( \beta_{ij} \) is the logarithmic coefficient, \( \beta_{ii} \) are the quadratic coefficients and \( x_i \) are the coded independent variables.

In the case of three variables, the second-order model is given by:

\[
\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3
\]

Where, \( \hat{y} \) is the response variable (dependent variable), \( \beta_0 \) is intercept (constant), \( \beta_1, \beta_2 \) and \( \beta_3 \) are linear coefficients, \( \beta_{12}, \beta_{13}, \) and \( \beta_{23} \) are interaction coefficients, \( \beta_{11}, \beta_{22} \) and \( \beta_{33} \) are quadratic coefficients and \( x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1 x_2, x_1 x_3 \) and \( x_2 x_3 \) are the level of coded independent variables. This model would likely be useful as an approximation to the true response surface in a relatively small region. The second-order model is widely used in response surface methodology for several reasons:

1. The second-order model is very flexible. It can take on a wide variety of functional forms, so it often works well as an approximation to the true response surface.

2. It is easy to estimate the parameters (the \( \beta \)'s) in the second-order model. The method of least squares could be used for this purpose.

3. There is considerable practical experience indicating that second-order models work well in solving real response surface problems.

Often, a second order model fit to the experimental data, including all linear, quadratic and cross product terms for the \( x_i \). The second order model, usually fits by ordinary least squares is represented for a single response \( y_u, u=1,\ldots, N \). Experimental data were fitted to, a first and second-order polynomial model to obtain the regression coefficients. The yield of \textit{rose coco} beans was taken as the dependent response variable. The data obtained from \textit{rose coco} on response variables were subjected to the analysis of variance (ANOVA). The mean values of the
triplicate trials were fitted to a second-order polynomial according to twenty four point second order rotatable design.

### 3.4.3 Model adequacy checking

The goodness of the fit of the developed regression models was tested. It is always necessary to:

1. Examine the fitted model to ensure that it provides an adequate approximation to the true system;
2. Verify that none of the least squares regression assumptions are violated. We consider several techniques for checking model adequacy this includes: the lack-of-fit significance, coefficient of determination, $R^2$, coefficient of variation (CV), mean square error (MSE) and model significance were used to judge adequacy of model fit. The statistical significance of the model equation was determined by Fisher’s test value ($p$ value) and significance of each coefficient was determined using $t$-test, and the extent of variance that could be explained by the model was determined by the multiple coefficient of determination, $R$ squared ($R^2$) value, this assess the fitness of the polynomial model (Myers et al, 2008). When $R^2$ approaches unity, the better the empirical model fits the actual data. The smaller the value of $R^2$, the less relevant the dependent variables in the model have in explaining the behavior variation (Little and Hill, 1978). In general, the $R^2$ measures percentage of the variation of $y$ around $\bar{y}$ that is explained by the regression equation.

Then the coefficient of multiple determination, $R^2$ is defined as

$$R^2 = 1 - \frac{SS_E}{SS_T}$$  \hspace{1cm} (3.20)

From inspection of the analysis of variance identity (equation 3.20) we could see that $0 \leq R^2 \leq 1$. One can gauge the fit of the model using the $R^2$, which is usually defined as the proportion of variance of the response that could be explained by the independent variables. The higher values
of $R^2$ are generally taken to indicate a better model. In fact, though, when the $R^2$ is low there are two possible conclusions. The first is that the independent variables and the dependent variables have very little relation to each other; the second is that the model fit is poor. Further, there are a number of possible causes of this poor fit, some outliers; an incorrect form of the independent variables; may be even incorrect assumptions about the errors. The errors may not be normally distributed, or indeed the independent variables may not be fixed. To confound the issues further, the low $R^2$ may be any combination of the above aspects (Simpson P. et al, 2004). According to Myers et al (2008), for a good fit of a model, the correlation coefficient should be at least 0.80.

### 3.4.4 The residual analysis

The residuals from the least squares fit, defined by $e_i = y_i - \hat{y}_i$, $i = 1, 2, \ldots, n$, play an important role in judging model adequacy. The PRESS computes and displays the predicted residual sum of squares (PRESS) statistic for each dependent variable in the model. The large difference between the ordinary residual and the PRESS residual indicates a point where the model fits the data well, but a model built without that point predicts poorer. The significant terms in the model was found by analysis of variance (ANOVA) for each response. Significance was judged by determining the probability level that the F-statistic calculated from the data is less than 5%. The model adequacies were checked by $R^2$ and the prediction error sum of squares (PRESS). A good model has a large predicted $R^2$, and a low PRESS. This analysis included examining diagnostic plots such as normal probability plots. To analyze the multiple regression and variance, develop a regression equation between variables and response and numerically
optimize the procedure; the SAS statistical package (version 9.0, SAS Institute Inc., Cary, NC, USA) was employed.

3.5 The settings of the experimental factors that produces the optimal response

The experiment was run using the optimum values for the variables given by response optimization in order to validate the predicted optimum values of variable response of the *rose coco*. The *rose coco* beans were labeled in the greenhouse to facilitate the identification during the entire procedure. The average yield was computed for GP1G, GP2G and GP3G. The SAS software was used to generate the critical value of the responses and also to generate response surfaces, while holding one variable constant in the second-order polynomial model. The R-software was used to come up with contour plots. Canonical analysis (Khuri and Cornell, 1996) was carried out to determine the location and nature of the stationary point of the model. When the results showed a saddle point in response surfaces the ridge analysis of SAS RSREG procedure was employed to compute the estimated ridge of the maximum response for increasing radii from the center of original design. The fitted polynomial equations were generated to response surface and contour plots so as to visualize the relationship between the response and experimental levels of each factor. The regression analysis and contour graphs were obtained by using SAS software (Version 9.3, SAS Institute, Cary, NC) and R-software.

3.5.1 Coding the factor variables

For the results of the canonical and ridge analyses to be interpretable, the values of different factor variables should be comparable. This is because the canonical and ridge analyses of the response surface are not invariant with respect to differences in scale and location of the factor
variables. The analysis of variance is not affected by these changes. Although the actual predicted surface does not change, its parameterization does. The usual solution to this problem is to code each factor variable so that its minimum in the experiment is –1 and its maximum is 1 and to carry through the analysis with the coded values instead of the original ones. This practice has the added benefit of making 1 a reasonable boundary radius for the ridge analysis since 1 represents approximately the edge of the experimental region. By default, PROC RSREG computes the linear transformation to perform this coding as the data are initially read in, and the canonical and ridge analyses are performed on the model fit to the coded data. The actual form of the coding operation for each value of a variable is given by:

\[
\psi_{iu} = \frac{x_{iu} \psi_i}{s_i}, \quad \psi_{i*} = \frac{\sum_{u=1}^{N} \psi_{iu}}{N} \quad \text{and} \quad s_i = \left[ \frac{\sum_{u=1}^{N} (\psi_{iu} - \psi_{i*})^2}{N} \right]^{1/2}
\]

(3.21)

\[
\psi_{iu} = x_{iu} s_i \psi_{i*}.
\]

For \( \sum_{u=1}^{N} x_{iu}^2 = N \) and \( \sum_{u=1}^{N} x_{iu} = 0 \).

The second-order models illustrate quadratic surfaces such as minimum, maximum and saddle. The graphical visualization is very helpful in understanding the second-order response surface. Specifically, contour plots can help characterize the shape of the surface and locate the optimum response roughly.

### 3.5.2 The stationary point

The sign of the stationary point is determined from the signs of the eigenvalues of the matrix \( \mathbf{B} \).

The standard quadratic model could be written in matrix notation as:-
\[ y = \hat{\beta}_0 + x' b + x' B x + \varepsilon \]  

(3.22)

where \( x \) is a fixed combination of the levels of the \( k \) input variables, \( \hat{\beta}_0 \), \( b \) and \( B \) contains estimates of the intercept, linear and second order coefficients, respectively.

Where \( b = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \) and \( B = \begin{bmatrix} \hat{\beta}_{11} & \frac{1}{2} \hat{\beta}_{12} & \ldots & \frac{1}{2} \hat{\beta}_{1k} \\ \frac{1}{2} \hat{\beta}_{12} & \hat{\beta}_{22} & \ldots & \frac{1}{2} \hat{\beta}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2} \hat{\beta}_{1k} & \frac{1}{2} \hat{\beta}_{2k} & \ldots & \hat{\beta}_{kk} \end{bmatrix} \)  

(3.23)

\( x' = [x_1, x_2, \ldots, x_k] \) and \( B \) is the \( k \times k \) symmetric matrix and \( \varepsilon \) is the random error. The stationary point is the combination of design variables where the surface is at either a maximum or a minimum in all directions. If the stationary point is a maximum in some direction and minimum in another direction, then the stationary point is a saddle point. When the surface curves in one direction, but is fairly constant in another direction, then this type of surface is called ridge system (Oehlert, 2000). The stationary point could be found by using matrix algebra. The fitted second-order model in matrix form is as follows:

\[ \hat{y} = \hat{\beta}_0 + x' b + x' B x \]  

(3.24)

The derivative of \( \hat{y} \) with respect to the elements of the vector \( x \) is

\[ \frac{\partial \hat{y}}{\partial x} = b + 2Bx \]  

(3.25)

Therefore, setting the derivative vector to 0 yields the stationary point of the system:

\[ x_s = -\frac{1}{2} B^{-1} b \]  

(3.26)
This might be a maximum, minimum, or a saddle point of the fitted surface. The eigenvalues (call them $\lambda$s) and eigenvectors of $B$ are the key to characterizing the shape. The $x_s$ are a point of maximum if all $\lambda$'s are negative, the point of minimum if all $\lambda$'s are positive and saddle point if $\lambda$'s are of mixed sign.

Where $\mathbf{b} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} \hat{\beta}_{11} & \frac{1}{2} \hat{\beta}_{12} & \frac{1}{2} \hat{\beta}_{13} \\ \frac{1}{2} \hat{\beta}_{12} & \hat{\beta}_{22} & \frac{1}{2} \hat{\beta}_{23} \\ \frac{1}{2} \hat{\beta}_{13} & \frac{1}{2} \hat{\beta}_{23} & \hat{\beta}_{33} \end{bmatrix}$

(3.27)

$b$ is a (3x1) vector of the first-order regression coefficients and $B$ is a (3x3) symmetric matrix whose main diagonal elements are the quadratic coefficients and whose off-diagonal elements are one-half the mixed quadratic coefficients (Montgomery, 2005). The estimated response value at the stationary point is

$\hat{y}_s = \hat{\beta}_0 + \frac{1}{2} x_s'^{\mathbf{b}}$

(3.28)

3.5.3 The canonical analysis

Canonical analysis is a mathematical approach used to examine the overall shape of the response surface and to determine if the estimated response point is a maximum, minimum or a saddle point. If the stationary point is a maximum or minimum, a corresponding increase or decrease results in the response. In the case of a saddle point, the response may increase or decrease when we move away from the stationary point, depending on which direction is taken.

The eigenvalues and eigenvectors of the matrix of second-order characterize the shape of the response surface. The eigenvectors point in the directions of principal orientation of the surface, and the signs and magnitudes of the associated eigenvalues give the shape of the surface in
these directions. Positive eigenvalues indicate directions of upward curvature, and negative eigenvalues indicate directions of downward curvature. The larger an eigenvalue is in absolute value, the more pronounced is the curvature of the response surface in the associated direction. Often, all the coefficients of an eigenvector except for one are relatively small, indicating that the vector points roughly along the axis associated with the factor corresponding to the single large coefficient. In this case, the canonical analysis could be used to determine the relative sensitivity of the predicted response surface to variations in that factor (SAS user guide 2013).

### 3.5.4 The ridge analysis

Ridge analysis in RSM is a contour-based technique for investigating a quadratic response surface. It is a search for a new stationary point $S_R$ on a given radius $R$ such that the second order model has a minimum at this stationary point. Then, the maximum or minimum response value at different locations from the design center could be determined by comparing each ‘constrained’ stationary point (Myers and Montgomery 2002 and Khuri and Cornell 1996). This procedure of RIDGE was conducted by the SAS program (SAS, version 9.0). A RIDGE statement computes the ridge of optimum response. The ridge starts at a given point $x_0$, and the point on the ridge at radius $R$ from $x_0$ is the collection of factor settings that optimizes the predicted response at this radius. The ridge analysis could be used as a tool to help interpret an existing response surface or to indicate the direction in which further experimentation should be performed. The default starting point, $x_0$, has each coordinate equal to the point midway between the highest and lowest values of the factor in the design. The default radii at which the ridge is computed are 0, 0.1, ..., 0.9, 1. If the ridge analysis is based on the response surface fit to coded values for the factor variables, then these results in a ridge that starts at the point with a
coded zero value for each coordinate and extends toward, but not beyond, the edge of the range of experimentation. Alternatively, both the center point of the ridge and the radii at which it is to be computed could be specified. The starting point should be well inside the range of experimentation. The coded radii give the distances from the ridge starting point at which to compute the optimal (SAS user guide 2013).
CHAPTER 4: RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter presents the results and discussions of the study in line with the four specific objectives.

4.2 Estimation of the linear parameters in an existing A-optimum and D-efficient calculus optimum value second order rotatable design

This section gives the data collected on rose coco beans, twenty four points rotatable design of coded levels and natural levels with the yields of rose coco beans for GP1G, GP2G and GP3G.

Table 4.1: The yield of 30 rose coco beans with straight fertilizers: nitrogen, phosphorus and potassium.

<table>
<thead>
<tr>
<th>FERTILIZERS IN GRAMS</th>
<th>Potassium (K)</th>
<th>Phosphorus (P)</th>
<th>Nitrogen (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>38</td>
<td>52</td>
<td>25</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>65</td>
<td>25</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
<td>45</td>
<td>25</td>
</tr>
<tr>
<td>40</td>
<td>30</td>
<td>40</td>
<td>24</td>
</tr>
<tr>
<td>50</td>
<td>50</td>
<td>70</td>
<td>20</td>
</tr>
<tr>
<td>60</td>
<td>25</td>
<td>48</td>
<td>12</td>
</tr>
<tr>
<td>70</td>
<td>25</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>80</td>
<td>25</td>
<td>40</td>
<td>25</td>
</tr>
<tr>
<td>90</td>
<td>20</td>
<td>42</td>
<td>26</td>
</tr>
<tr>
<td>100</td>
<td>15</td>
<td>45</td>
<td>15</td>
</tr>
</tbody>
</table>
The table 4.1 shows the *rose coco* yield of the straight fertilizers, nitrogen, phosphorus and potassium. Each of the fertilizers was initially started with 10 grams then increased by 10 grams up to 100 grams for each *rose coco* plant.

Table 4.2: Descriptive Analysis of straight fertilizers of N, P and K on *rose coco* beans mean yield.

<table>
<thead>
<tr>
<th>Descriptives</th>
<th>Weight <em>rose coco</em></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>95% Confidence Interval for Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>28.800</td>
<td>9.71597</td>
<td>3.07246</td>
<td>21.8496 - 35.7504</td>
<td>15.00</td>
<td>50.00</td>
</tr>
<tr>
<td>Potassium (K)</td>
<td></td>
<td>10</td>
<td>49.700</td>
<td>10.25291</td>
<td>3.24226</td>
<td>42.3655 - 57.0345</td>
<td>40.00</td>
<td>70.00</td>
</tr>
<tr>
<td>Phosphorus (P)</td>
<td></td>
<td>10</td>
<td>21.700</td>
<td>4.85455</td>
<td>1.53514</td>
<td>18.2273 - 25.1727</td>
<td>12.00</td>
<td>26.00</td>
</tr>
<tr>
<td>Nitrogen (N)</td>
<td></td>
<td>10</td>
<td>21.700</td>
<td>4.85455</td>
<td>1.53514</td>
<td>18.2273 - 25.1727</td>
<td>12.00</td>
<td>26.00</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>30</td>
<td>33.400</td>
<td>14.67487</td>
<td>2.67925</td>
<td>27.9203 - 38.8797</td>
<td>12.00</td>
<td>70.00</td>
</tr>
</tbody>
</table>

Table 4.2 shows that phosphorus has more influence on the *rose coco* beans followed by potassium then nitrogen.

Table 4.3: Analysis of variance for the yield of *rose coco* beans on straight fertilizers of N, P and K of table 4.1.

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Weight <em>rose coco</em></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td></td>
<td>4237.400</td>
<td>2</td>
<td>2118.700</td>
<td>28.491</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td></td>
<td>2007.800</td>
<td>27</td>
<td>74.363</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>6245.200</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In table 4.3 the $p = 0.000$ is less than $\alpha = 0.05$, therefore, we reject the null hypothesis that states that the nitrogen, phosphorus and potassium has the same effects on rose coco beans and conclude that the treatments differ significantly, meaning that all the three fertilizers do not have the same effects on the rose coco beans.

**Means Plots**

![Means Plot Diagram](image)

Figure 4.1: The plot of mean of yield of *rose coco* beans under each of the three fertilizers; N, P, K.
Figure 4.2: The yield of *rose coco* beans under each of the three straight fertilizers; N, P, K of table 4.1.

Figure 4.2 shows the effect of individual fertilizers applied to *rose coco* beans. The fertilizers were applied once during the planting period and no other kind of fertilizers was applied after the initial application. Weeding and pest control was done. The phosphorus fertilizer had more influence on *rose coco* bean followed by potassium and then nitrogen.

4.2.2 Twenty four points rotatable design of coded levels and natural levels with the yields of *rose coco* beans-GPIG

The $x_{1u}$, $x_{2u}$ and $x_{3u}$ are coded values while $\Psi_{1u}$, $\Psi_{2u}$ and $\Psi_{3u}$ are natural values-GPIG. The natural values ($\Psi_{iu}$) of fertilizers N, P, K, were measured using a sensitive weighing scale and
applied to *rose coco* beans in a greenhouse which gave the observed yield \( y_i \). Predicted yield \( \hat{y} \) was calculated using a second order model of GP1G (equation 4.54).

Table 4. 4: The 24-point’s rotatable design of coded levels and natural levels with the yield of *rose coco* beans-GP1G at the ratio 10:20:30 of N: P: K fertilizers.

<table>
<thead>
<tr>
<th>( x_{1u} )</th>
<th>( x_{2u} )</th>
<th>( x_{3u} )</th>
<th>( \Psi_{1u} )</th>
<th>( \Psi_{2u} )</th>
<th>( \Psi_{3u} )</th>
<th>Observed Yield- ( y_i )</th>
<th>Predicted yield- ( \hat{y} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1072569</td>
<td>1.1072569</td>
<td>0</td>
<td>10.553628</td>
<td>20.332177</td>
<td>30</td>
<td>63</td>
<td>56.4342</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>1.1072569</td>
<td>0</td>
<td>9.446372</td>
<td>20.332177</td>
<td>30</td>
<td>71</td>
<td>61.1068</td>
</tr>
<tr>
<td>1.1072569</td>
<td>-1.1072569</td>
<td>0</td>
<td>10.553628</td>
<td>19.667823</td>
<td>30</td>
<td>35</td>
<td>35.8807</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>-1.1072569</td>
<td>0</td>
<td>9.446372</td>
<td>19.667823</td>
<td>30</td>
<td>56</td>
<td>53.5533</td>
</tr>
<tr>
<td>1.1072569</td>
<td>0</td>
<td>1.1072569</td>
<td>10.553628</td>
<td>20</td>
<td>31.107257</td>
<td>37</td>
<td>44.3546</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>0</td>
<td>1.1072569</td>
<td>9.446372</td>
<td>20</td>
<td>31.107257</td>
<td>62</td>
<td>69.5273</td>
</tr>
<tr>
<td>1.1072569</td>
<td>0</td>
<td>-1.1072569</td>
<td>10.553628</td>
<td>20</td>
<td>28.892743</td>
<td>57</td>
<td>55.7932</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>0</td>
<td>-1.1072569</td>
<td>9.446372</td>
<td>20</td>
<td>28.892743</td>
<td>54</td>
<td>52.9659</td>
</tr>
<tr>
<td>0</td>
<td>1.1072569</td>
<td>1.1072569</td>
<td>20</td>
<td>20.332177</td>
<td>31.107257</td>
<td>52</td>
<td>55.3883</td>
</tr>
<tr>
<td>0</td>
<td>-1.1072569</td>
<td>1.1072569</td>
<td>19.667823</td>
<td>31.107257</td>
<td>52</td>
<td>47.8349</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.1072569</td>
<td>-1.1072569</td>
<td>20.332177</td>
<td>28.892743</td>
<td>53</td>
<td>59.3270</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>-1.1072569</td>
<td>-1.1072569</td>
<td>19.667823</td>
<td>28.892743</td>
<td>40</td>
<td>38.7735</td>
<td></td>
</tr>
<tr>
<td>0.7829487</td>
<td>0</td>
<td>0</td>
<td>10.391474</td>
<td>20</td>
<td>30</td>
<td>43</td>
<td>45.8615</td>
</tr>
<tr>
<td>-0.7829487</td>
<td>0</td>
<td>0</td>
<td>9.608526</td>
<td>20</td>
<td>30</td>
<td>58</td>
<td>53.7617</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.7829487</td>
<td>20</td>
<td>30.782949</td>
<td>58</td>
<td>50.0108</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>-0.7829487</td>
<td>20</td>
<td>29.217051</td>
<td>42</td>
<td>48.1996</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.7829487</td>
<td>0</td>
<td>20</td>
<td>20.234885</td>
<td>30</td>
<td>48</td>
<td>52.1156</td>
</tr>
<tr>
<td>0</td>
<td>-0.7829487</td>
<td>0</td>
<td>20</td>
<td>19.765115</td>
<td>30</td>
<td>44</td>
<td>42.1783</td>
</tr>
</tbody>
</table>
4.2.2.1 The three replication yield of twenty four points second order rotatable design,

**Group 1-GP1G**

Table 4.5 below shows the three replication yields of twenty four points *rose coco* beans (from table 4.4) under fertilizers 10 grams of nitrogen, 20 grams of phosphorus and 30 grams of potassium.

Table 4.5: The yield of 24-points *rose coco* beans (from table 4.4) under fertilizers at the ratio of 10 grams of nitrogen, 20 grams of phosphorus and 30 grams of potassium

<table>
<thead>
<tr>
<th>N:P:K</th>
<th>10:20:30</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Bean No.</th>
<th>GP1GA Yield</th>
<th>GP1GB Yield</th>
<th>GP1GC Yield</th>
<th>AVERAGE YIELD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41</td>
<td>48</td>
<td>100</td>
<td>63</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>170</td>
<td>30</td>
<td>71</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>35</td>
<td>50</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>88</td>
<td>60</td>
<td>56</td>
</tr>
<tr>
<td>5</td>
<td>63</td>
<td>20</td>
<td>28</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>6</td>
<td>76</td>
<td>70</td>
<td>40</td>
<td>62</td>
</tr>
<tr>
<td>7</td>
<td>76</td>
<td>60</td>
<td>35</td>
<td>57</td>
</tr>
<tr>
<td>8</td>
<td>42</td>
<td>75</td>
<td>45</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>84</td>
<td>32</td>
<td>40</td>
<td>52</td>
</tr>
<tr>
<td>10</td>
<td>38</td>
<td>58</td>
<td>60</td>
<td>52</td>
</tr>
<tr>
<td>11</td>
<td>34</td>
<td>60</td>
<td>65</td>
<td>53</td>
</tr>
<tr>
<td>12</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>13</td>
<td>27</td>
<td>70</td>
<td>32</td>
<td>43</td>
</tr>
<tr>
<td>14</td>
<td>70</td>
<td>70</td>
<td>34</td>
<td>58</td>
</tr>
<tr>
<td>15</td>
<td>_</td>
<td>36</td>
<td>80</td>
<td>58</td>
</tr>
<tr>
<td>16</td>
<td>83</td>
<td>15</td>
<td>28</td>
<td>42</td>
</tr>
<tr>
<td>17</td>
<td>56</td>
<td>50</td>
<td>38</td>
<td>48</td>
</tr>
<tr>
<td>18</td>
<td>32</td>
<td>50</td>
<td>50</td>
<td>44</td>
</tr>
<tr>
<td>19</td>
<td>53</td>
<td>55</td>
<td>42</td>
<td>50</td>
</tr>
<tr>
<td>20</td>
<td>65</td>
<td>42</td>
<td>52</td>
<td>53</td>
</tr>
<tr>
<td>21</td>
<td>42</td>
<td>98</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>22</td>
<td>28</td>
<td>95</td>
<td>45</td>
<td>56</td>
</tr>
<tr>
<td>23</td>
<td>41</td>
<td>40</td>
<td>75</td>
<td>52</td>
</tr>
<tr>
<td>24</td>
<td>41</td>
<td>30</td>
<td>25</td>
<td>32</td>
</tr>
</tbody>
</table>

GP1G-10 grams nitrogen, 20 grams phosphorus, 30 grams potassium

Blank space indicated by a dash, the *rose coco* plant died so no yield was obtained.
4.2.3 Twenty four points rotatable design of coded levels and natural levels with the yield of rose coco beans-GP2G

The $x_{1u}$, $x_{2u}$ and $x_{3u}$ are coded values while $\Psi_{1u}$, $\Psi_{2u}$ and $\Psi_{3u}$ are natural values-GP2G. The natural values ($\Psi_{iu}$) of fertilizers N, P, K, were measured using a sensitive weighing scale and planted in a greenhouse which gave the observed yield-$y_i$. Predicted yield $\hat{y}$ was calculated using a second order model of GP2G (equation 4.62).

Table 4.6: The 24-point’s rotatable design of coded levels and natural levels with the yield of *rose coco* beans GP2G at the ratio of 20:30:40 N: P: K fertilizers.

<table>
<thead>
<tr>
<th>$(x_{1u})$</th>
<th>$x_{2u}$</th>
<th>$x_{3u}$</th>
<th>$\Psi_{1u}$</th>
<th>$\Psi_{2u}$</th>
<th>$\Psi_{3u}$</th>
<th>Observed Yield-$y_i$</th>
<th>Predicted yield-$\hat{y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1072569</td>
<td>1.1072569</td>
<td>0</td>
<td>20.553628</td>
<td>30.332177</td>
<td>40</td>
<td>49</td>
<td>53.3803</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>1.1072569</td>
<td>0</td>
<td>19.446372</td>
<td>30.332177</td>
<td>40</td>
<td>30</td>
<td>35.1918</td>
</tr>
<tr>
<td>1.1072569</td>
<td>-1.1072569</td>
<td>0</td>
<td>20.553628</td>
<td>29.667823</td>
<td>40</td>
<td>44</td>
<td>39.9260</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>-1.1072569</td>
<td>0</td>
<td>19.446372</td>
<td>29.667823</td>
<td>40</td>
<td>55</td>
<td>51.7374</td>
</tr>
<tr>
<td>1.1072569</td>
<td>0</td>
<td>1.1072569</td>
<td>20.553628</td>
<td>30</td>
<td>41.107257</td>
<td>59</td>
<td>55.9145</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>0</td>
<td>1.1072569</td>
<td>19.446372</td>
<td>30</td>
<td>41.107257</td>
<td>53</td>
<td>45.7260</td>
</tr>
<tr>
<td>1.1072569</td>
<td>0</td>
<td>-1.1072569</td>
<td>20.553628</td>
<td>30</td>
<td>38.892743</td>
<td>37</td>
<td>42.8240</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>0</td>
<td>-1.1072569</td>
<td>19.446372</td>
<td>30</td>
<td>38.892743</td>
<td>45</td>
<td>46.6354</td>
</tr>
<tr>
<td>0</td>
<td>1.1072569</td>
<td>1.1072569</td>
<td>20</td>
<td>30.332177</td>
<td>41.107257</td>
<td>64</td>
<td>58.5578</td>
</tr>
<tr>
<td>0</td>
<td>-1.1072569</td>
<td>1.1072569</td>
<td>20</td>
<td>29.667823</td>
<td>41.107257</td>
<td>56</td>
<td>57.6034</td>
</tr>
<tr>
<td>0</td>
<td>1.1072569</td>
<td>-1.1072569</td>
<td>20</td>
<td>30.332177</td>
<td>38.892743</td>
<td>57</td>
<td>49.9673</td>
</tr>
<tr>
<td>0</td>
<td>1.1072569</td>
<td>-1.1072569</td>
<td>20</td>
<td>29.667823</td>
<td>38.892743</td>
<td>54</td>
<td>54.0129</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>40</td>
<td>34</td>
<td>37.6742</td>
</tr>
<tr>
<td>-----------------</td>
<td>------</td>
<td>------</td>
<td>----------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>---------</td>
</tr>
<tr>
<td>0.7829487</td>
<td>0</td>
<td>0</td>
<td>20.391474</td>
<td>30</td>
<td>40</td>
<td>34</td>
<td>37.6742</td>
</tr>
<tr>
<td>-0.7829487</td>
<td>0</td>
<td>0</td>
<td>19.608526</td>
<td>30</td>
<td>40</td>
<td>39</td>
<td>35.4195</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.7829487</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>37</td>
<td>43.6884</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>-0.7829487</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>42</td>
<td>39.3817</td>
</tr>
<tr>
<td>0</td>
<td>0.7829487</td>
<td>0</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>47</td>
<td>39.6306</td>
</tr>
<tr>
<td>0</td>
<td>-0.7829487</td>
<td>0</td>
<td>20</td>
<td>29.765115</td>
<td>40</td>
<td>47</td>
<td>40.7235</td>
</tr>
<tr>
<td>1.2735263</td>
<td>0</td>
<td>0</td>
<td>20.636763</td>
<td>30</td>
<td>40</td>
<td>44</td>
<td>39.0672</td>
</tr>
<tr>
<td>-1.2735263</td>
<td>0</td>
<td>0</td>
<td>19.363237</td>
<td>30</td>
<td>40</td>
<td>30</td>
<td>35.3998</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1.2735263</td>
<td>20</td>
<td>30</td>
<td>41.273526</td>
<td>46</td>
<td>53.9337</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>-1.2735263</td>
<td>20</td>
<td>30</td>
<td>38.726474</td>
<td>46</td>
<td>46.9286</td>
</tr>
<tr>
<td>0</td>
<td>1.2735263</td>
<td>0</td>
<td>20</td>
<td>30.382058</td>
<td>40</td>
<td>41</td>
<td>45.9492</td>
</tr>
<tr>
<td>0</td>
<td>-1.2735263</td>
<td>0</td>
<td>20</td>
<td>29.617942</td>
<td>40</td>
<td>41</td>
<td>47.7269</td>
</tr>
</tbody>
</table>

### 4.2.3.1 The three replication yield of twenty four points second order rotatable design, Group 2-GP2G

Table 4.7 below shows the three replication yields of twenty four points rotatable design of *rose coco beans* (from table 4.6) under fertilizers at the ratio of 20 grams of nitrogen, 30 grams of phosphorus and 40 grams of potassium.

Table 4. 7: The yield of 24-points *rose coco* beans (from table 4.6) under fertilizers at the ratio of 20 grams of nitrogen, 30 grams of phosphorus and 40 grams of potassium

N:P:K

20:30:40
<table>
<thead>
<tr>
<th>Bean No.</th>
<th>GP2GA</th>
<th>GP2GB</th>
<th>GP2GC</th>
<th>AVERAGE YIELD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>42</td>
<td>60</td>
<td>45</td>
<td>49</td>
</tr>
<tr>
<td>2</td>
<td>47</td>
<td>18</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>62</td>
<td>30</td>
<td>44</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>35</td>
<td>70</td>
<td>55</td>
</tr>
<tr>
<td>5</td>
<td>36</td>
<td>82</td>
<td>_</td>
<td>59</td>
</tr>
<tr>
<td>6</td>
<td>51</td>
<td>50</td>
<td>58</td>
<td>53</td>
</tr>
<tr>
<td>7</td>
<td>69</td>
<td>22</td>
<td>20</td>
<td>37</td>
</tr>
<tr>
<td>8</td>
<td>73</td>
<td>20</td>
<td>42</td>
<td>45</td>
</tr>
<tr>
<td>9</td>
<td>75</td>
<td>72</td>
<td>45</td>
<td>64</td>
</tr>
<tr>
<td>10</td>
<td>58</td>
<td>58</td>
<td>52</td>
<td>56</td>
</tr>
<tr>
<td>11</td>
<td>51</td>
<td>65</td>
<td>55</td>
<td>57</td>
</tr>
<tr>
<td>12</td>
<td>72</td>
<td>50</td>
<td>40</td>
<td>54</td>
</tr>
<tr>
<td>13</td>
<td>22</td>
<td>32</td>
<td>48</td>
<td>34</td>
</tr>
<tr>
<td>14</td>
<td>50</td>
<td>32</td>
<td>35</td>
<td>39</td>
</tr>
<tr>
<td>15</td>
<td>43</td>
<td>18</td>
<td>50</td>
<td>37</td>
</tr>
<tr>
<td>16</td>
<td>36</td>
<td>35</td>
<td>55</td>
<td>42</td>
</tr>
<tr>
<td>17</td>
<td>46</td>
<td>55</td>
<td>40</td>
<td>47</td>
</tr>
<tr>
<td>18</td>
<td>56</td>
<td>35</td>
<td>50</td>
<td>47</td>
</tr>
<tr>
<td>19</td>
<td>32</td>
<td>58</td>
<td>42</td>
<td>44</td>
</tr>
<tr>
<td>20</td>
<td>30</td>
<td>40</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>21</td>
<td>77</td>
<td>21</td>
<td>40</td>
<td>46</td>
</tr>
</tbody>
</table>
Blank space indicated by a dash, the *rose coco* plant died, so no yield was obtained.

### 4.2.4 Twenty four points rotatable design of coded levels and natural levels with the yield of *rose coco* beans-GP3G

The $x_{1u}$, $x_{2u}$ and $x_{3u}$ are coded values while $\Psi_{1u}$, $\Psi_{2u}$ and $\Psi_{3u}$ are natural values-GP3G. The natural values ($\Psi_{iu}$) of fertilizers N, P, K, were measured using a sensitive weighing scale and planted in a greenhouse which gave the observed yield-$y_i$. Predicted yield $\hat{y}$ was calculated using a second order model of GP3G (equation 4.67).

Table 4. 8: The 24-point's rotatable design of coded levels and natural levels with the yield of *rose coco* beans-GP3G at the ratio of 30:40:50 N: P: K fertilizers.

<table>
<thead>
<tr>
<th>$(x_{1u})$</th>
<th>$x_{2u}$</th>
<th>$x_{3u}$</th>
<th>$\Psi_{1u}$</th>
<th>$\Psi_{2u}$</th>
<th>$\Psi_{3u}$</th>
<th>Observed Yield- $y_i$</th>
<th>Predicted yield- $\hat{y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1072569</td>
<td>1.1072569</td>
<td>0</td>
<td>30.553628</td>
<td>40.332177</td>
<td>50</td>
<td>47</td>
<td>53.3952</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>1.1072569</td>
<td>0</td>
<td>29.446372</td>
<td>40.332177</td>
<td>50</td>
<td>82</td>
<td>75.6594</td>
</tr>
<tr>
<td>1.1072569</td>
<td>-1.1072569</td>
<td>0</td>
<td>30.553628</td>
<td>39.667823</td>
<td>50</td>
<td>68</td>
<td>75.4937</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>-1.1072569</td>
<td>0</td>
<td>29.446372</td>
<td>39.667823</td>
<td>50</td>
<td>50</td>
<td>44.7579</td>
</tr>
<tr>
<td>1.1072569</td>
<td>0</td>
<td>1.1072569</td>
<td>30.553628</td>
<td>40</td>
<td>51.107257</td>
<td>69</td>
<td>63.6074</td>
</tr>
<tr>
<td>-1.1072569</td>
<td>0</td>
<td>1.1072569</td>
<td>29.446372</td>
<td>40</td>
<td>51.107257</td>
<td>53</td>
<td>59.8717</td>
</tr>
<tr>
<td>1.1072569</td>
<td>0</td>
<td>-1.1072569</td>
<td>30.553628</td>
<td>40</td>
<td>48.892743</td>
<td>87</td>
<td>78.5079</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>-1.1072569</td>
<td>0</td>
<td>-1.1072569</td>
<td>29.446372</td>
<td>40</td>
<td>48.892743</td>
<td>70</td>
<td>73.7722</td>
</tr>
<tr>
<td>0</td>
<td>1.1072569</td>
<td>1.1072569</td>
<td>30</td>
<td>40.332177</td>
<td>51.107257</td>
<td>67</td>
<td>60.8035</td>
</tr>
<tr>
<td>0</td>
<td>-1.1072569</td>
<td>1.1072569</td>
<td>30</td>
<td>39.667823</td>
<td>51.107257</td>
<td>73</td>
<td>67.9020</td>
</tr>
<tr>
<td>0</td>
<td>1.1072569</td>
<td>-1.1072569</td>
<td>30</td>
<td>40.332177</td>
<td>48.892743</td>
<td>84</td>
<td>86.7041</td>
</tr>
<tr>
<td>0</td>
<td>-1.1072569</td>
<td>-1.1072569</td>
<td>30</td>
<td>39.667823</td>
<td>48.892743</td>
<td>67</td>
<td>70.8025</td>
</tr>
<tr>
<td>0.7829487</td>
<td>0</td>
<td>0</td>
<td>30.391474</td>
<td>40</td>
<td>50</td>
<td>60</td>
<td>55.2395</td>
</tr>
<tr>
<td>-0.7829487</td>
<td>0</td>
<td>0</td>
<td>29.608526</td>
<td>40</td>
<td>50</td>
<td>51</td>
<td>52.2443</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.7829487</td>
<td>30</td>
<td>40</td>
<td>50.782949</td>
<td>51</td>
<td>53.2638</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>-0.7829487</td>
<td>30</td>
<td>40</td>
<td>49.217051</td>
<td>72</td>
<td>63.4465</td>
</tr>
<tr>
<td>0</td>
<td>0.7829487</td>
<td>0</td>
<td>30</td>
<td>40.234885</td>
<td>50</td>
<td>51</td>
<td>56.6047</td>
</tr>
<tr>
<td>0</td>
<td>-0.7829487</td>
<td>0</td>
<td>30</td>
<td>39.765115</td>
<td>50</td>
<td>54</td>
<td>53.4924</td>
</tr>
<tr>
<td>1.2735263</td>
<td>0</td>
<td>0</td>
<td>30.636763</td>
<td>40</td>
<td>50</td>
<td>57</td>
<td>59.4536</td>
</tr>
<tr>
<td>-1.2735263</td>
<td>0</td>
<td>0</td>
<td>29.363237</td>
<td>40</td>
<td>50</td>
<td>55</td>
<td>54.5818</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1.2735263</td>
<td>30</td>
<td>40</td>
<td>51.273526</td>
<td>55</td>
<td>60.9417</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>-1.2735263</td>
<td>30</td>
<td>40</td>
<td>48.726474</td>
<td>75</td>
<td>77.5047</td>
</tr>
<tr>
<td>0</td>
<td>1.2735263</td>
<td>0</td>
<td>30</td>
<td>40.382058</td>
<td>50</td>
<td>63</td>
<td>63.0059</td>
</tr>
<tr>
<td>0</td>
<td>-1.2735263</td>
<td>0</td>
<td>30</td>
<td>39.617942</td>
<td>50</td>
<td>58</td>
<td>57.9435</td>
</tr>
</tbody>
</table>

4.2.4.1 The three replication yield of twenty four point second order rotatable design, Group 3-GP3G

Table 4.9 shows the three replication yields of twenty four *rose coco* beans (from table 4.8) under fertilizers at the ratio of 30 grams of nitrogen, 40 grams of phosphorus and 50 grams of potassium.
Table 4.9: The yield of 24-points *rose coco* bean (from table 4.8) under fertilizers at the ratio of 30 grams of Nitrogen, 40 grams of Phosphorus and 50 grams of potassium

\[ \text{N:P:K} \]
\[ 30:40:50 \]

<table>
<thead>
<tr>
<th>Bean No.</th>
<th>GP3GA</th>
<th>GP3GB</th>
<th>GP3GC</th>
<th>AVERAGE YIELD GP3G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>61</td>
<td>40</td>
<td>40</td>
<td>47</td>
</tr>
<tr>
<td>2</td>
<td>136</td>
<td>40</td>
<td>70</td>
<td>82</td>
</tr>
<tr>
<td>3</td>
<td>84</td>
<td>80</td>
<td>40</td>
<td>68</td>
</tr>
<tr>
<td>4</td>
<td>72</td>
<td>38</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>77</td>
<td>80</td>
<td>50</td>
<td>69</td>
</tr>
<tr>
<td>6</td>
<td>67</td>
<td>48</td>
<td>44</td>
<td>53</td>
</tr>
<tr>
<td>7</td>
<td>93</td>
<td>80</td>
<td>88</td>
<td>87</td>
</tr>
<tr>
<td>8</td>
<td>82</td>
<td>60</td>
<td>68</td>
<td>70</td>
</tr>
<tr>
<td>9</td>
<td>21</td>
<td>100</td>
<td>80</td>
<td>67</td>
</tr>
<tr>
<td>10</td>
<td>80</td>
<td>58</td>
<td>81</td>
<td>73</td>
</tr>
<tr>
<td>11</td>
<td>100</td>
<td>80</td>
<td>72</td>
<td>84</td>
</tr>
<tr>
<td>12</td>
<td>_</td>
<td>72</td>
<td>62</td>
<td>67</td>
</tr>
<tr>
<td>13</td>
<td>78</td>
<td>70</td>
<td>32</td>
<td>60</td>
</tr>
<tr>
<td>14</td>
<td>58</td>
<td>65</td>
<td>30</td>
<td>51</td>
</tr>
<tr>
<td>15</td>
<td>71</td>
<td>50</td>
<td>32</td>
<td>51</td>
</tr>
<tr>
<td>16</td>
<td>64</td>
<td>74</td>
<td>78</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>78</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>---</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>18</td>
<td>67</td>
<td>55</td>
<td>40</td>
<td>54</td>
</tr>
<tr>
<td>19</td>
<td>60</td>
<td>75</td>
<td>36</td>
<td>57</td>
</tr>
<tr>
<td>20</td>
<td>32</td>
<td>68</td>
<td>65</td>
<td>55</td>
</tr>
<tr>
<td>21</td>
<td>65</td>
<td>55</td>
<td>45</td>
<td>55</td>
</tr>
<tr>
<td>22</td>
<td>60</td>
<td>125</td>
<td>40</td>
<td>75</td>
</tr>
<tr>
<td>23</td>
<td>67</td>
<td>80</td>
<td>42</td>
<td>63</td>
</tr>
<tr>
<td>24</td>
<td>64</td>
<td>80</td>
<td>30</td>
<td>58</td>
</tr>
</tbody>
</table>

GP3G-30 grams nitrogen, 40 grams phosphorus, 50 grams potassium

Blank space indicated by a dash, the rose coco plant died, so no yield was obtained.

Below is a figure 4.3 that shows the summary graph for the twenty four rose coco beans average yield for the three groups.

![Figure 4.3: The average yield of rose coco bean for the three groups.](image)

Figure 4.3: The average yield of rose coco bean for the three groups.
We saw in figure 4.3 that GP3G has the highest yield, the two groups GP1G and GP2G are more less the same, GP1G being slightly higher than GP2G.

### 4.2.5 Parameter estimation for the first order model

This section gives linear parameters of first order models for GP1G, GP2G and GP3G. The design matrix $X$ is given as in equation (3.14) and the equation (4.1) is obtained.

$$X = \begin{bmatrix}
1 & 1.1072569 & 1.1072569 & 0 \\
1 & -1.1072569 & 1.1072569 & 0 \\
1 & 1.1072569 & -1.1072569 & 0 \\
1 & -1.1072569 & -1.1072569 & 0 \\
1 & 1.1072569 & 0 & 1.1072569 \\
1 & -1.1072569 & 0 & 1.1072569 \\
1 & 1.1072569 & 0 & -1.1072569 \\
1 & -1.1072569 & 0 & -1.1072569 \\
1 & 0 & 1.1072569 & 1.1072569 \\
1 & 0 & -1.1072569 & 1.1072569 \\
1 & 0 & 1.1072569 & -1.1072569 \\
1 & 0 & -1.1072569 & -1.1072569 \\
1 & 0 & 0 & 0.7829487 \\
1 & 0 & -0.7829487 & 0 \\
1 & 0 & 0 & 0.7829487 \\
1 & 0 & 0 & -0.7829487 \\
1 & 0 & 0.7829487 & 0 \\
1 & 0 & -0.7829487 & 0 \\
1 & 0 & 1.2735263 & 0 \\
1 & 0 & -1.2735263 & 0 \\
1 & 0 & 0 & 1.2735263 \\
1 & 0 & 0 & -1.2735263 \\
1 & 0 & 1.2735263 & 0 \\
1 & 0 & -1.2735263 & 0
\end{bmatrix}$$  

(4.1)
The first-order model with three independent variables was expressed as in equation (3.9). On fitting the model in equation (3.9) to the data, the parameter estimates are derived.

### 4.2.5.1 Linear parameters of first-order model for GP1G

The model parameter estimates are calculated using the least squares method. In matrix notation, the least squares estimates of the model parameters are:

$$\hat{\beta} = (X'X)^{-1} X'Y.$$  \hspace{1cm} (4.2)

In order to find the least squares estimates, calculate \((XX)^{-1}\) and \(X'y\):

\[
(X'X) = \begin{bmatrix}
24 & 0 & 0 & 0 \\
0 & 14.2779 & 0 & 0 \\
0 & 0 & 14.2779 & 0 \\
0 & 0 & 0 & 14.2779 \\
\end{bmatrix} \\
(X'X)^{-1} = \begin{bmatrix}
0.0417 & 0 & 0 & 0 \\
0 & 0.0700 & 0 & 0 \\
0 & 0 & 0.0700 & 0 \\
0 & 0 & 0 & 0.0700 \\
\end{bmatrix} \\
(X'y) = \begin{bmatrix}
1228 \\
-72 \\
90.6 \\
16.5 \\
\end{bmatrix}  \hspace{1cm} (4.3)
\]

Thus, the least squares estimates of the parameters for the experimental model are:

$$\hat{B} = (X'X)^{-1} X'Y_{GP1G} = \begin{bmatrix}
51.1667 \\
-5.0452 \\
6.3461 \\
1.1566 \\
\end{bmatrix}$$  \hspace{1cm} (4.4)

### 4.2.5.2 Linear parameters of first-order model for GP2G

We fit first-order model of GP2G with three independent variables. On fitting the model to the data, we derive the parameter estimates as indicated below. In order to find the least square estimates, we calculated \((XX)^{-1}\) (given in equation 4.3) and \(X'y\) using equation (4.1), then equation (4.2) is given as follows:
\[
\begin{bmatrix}
1097 \\
20.6 \\
-10 \\
39.3
\end{bmatrix}
\quad \hat{B} = (X'X)^{-1}X'y_{GP2G} =
\begin{bmatrix}
45.7083 \\
1.4399 \\
-0.6980 \\
2.7503
\end{bmatrix}
\] (4.5)

**4.2.5.3 Linear parameters of first-order model for GP3G**

In order to find the parameter estimates for first-order model of GP3G, we calculated \((XX)^{-1}\) (given in equation 4.3) and \(X'y\) using equation (4.1), then equation (4.2) is given as follows:

\[
\begin{bmatrix}
1519 \\
27.3 \\
28.4 \\
-92.8
\end{bmatrix}
\quad \hat{B} = (X'X)^{-1}X'y_{GP3G} =
\begin{bmatrix}
63.2917 \\
1.9127 \\
1.9876 \\
-6.5028
\end{bmatrix}
\] (4.6)

**4.2.6 Parameter estimation for the second order model.**

This section found the second order parameter estimates for GP1G, GP2G and GP3G. The method of least squares was also used to estimate the regression coefficients; the \(X\) below is the design matrix for *rose coco* bean.
\[
X =
\begin{bmatrix}
1 & 1.1072569 & 1.1072569 & 0 & 1.226018 & 1.226018 & 0 & 1.226018 & 0 & 0 \\
1 & -1.1072569 & 1.1072569 & 0 & 1.226018 & 1.226018 & 0 & -1.226018 & 0 & 0 \\
1 & 1.1072569 & -1.1072569 & 0 & 1.226018 & 1.226018 & 0 & -1.226018 & 0 & 0 \\
1 & -1.1072569 & -1.1072569 & 0 & 1.226018 & 1.226018 & 0 & 1.226018 & 0 & 0 \\
1 & 1.1072569 & 0 & 1.1072569 & 1.226018 & 0 & 1.226018 & 0 & 1.226018 & 0 \\
1 & -1.1072569 & 0 & 1.1072569 & 1.226018 & 0 & 1.226018 & 0 & -1.226018 & 0 \\
1 & 1.1072569 & 0 & -1.1072569 & 1.226018 & 0 & 1.226018 & 0 & -1.226018 & 0 \\
1 & -1.1072569 & 0 & -1.1072569 & 1.226018 & 0 & 1.226018 & 0 & 1.226018 & 0 \\
1 & 0 & 1.1072569 & 1.1072569 & 0 & 1.226018 & 1.226018 & 0 & 0 & 1.226018 \\
1 & 0 & -1.1072569 & 1.1072569 & 0 & 1.226018 & 1.226018 & 0 & 0 & -1.226018 \\
1 & 0 & 1.1072569 & -1.1072569 & 0 & 1.226018 & 1.226018 & 0 & 0 & -1.226018 \\
1 & 0 & -1.1072569 & -1.1072569 & 0 & 1.226018 & 1.226018 & 0 & 0 & 1.226018 \\
1 & 0.7829487 & 0 & 0 & 0.613009 & 0 & 0 & 0 & 0 & 0 \\
1 & -0.7829487 & 0 & 0 & 0.613009 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0.7829487 & 0 & 0 & 0.613009 & 0 & 0 & 0 \\
1 & 0 & 0 & -0.7829487 & 0 & 0 & 0.613009 & 0 & 0 & 0 \\
1 & 0 & 0.7829487 & 0 & 0 & 0.613009 & 0 & 0 & 0 & 0 \\
1 & 0 & -0.7829487 & 0 & 0 & 0.613009 & 0 & 0 & 0 & 0 \\
1 & 1.2735263 & 0 & 0 & 1.621869 & 0 & 0 & 0 & 0 & 0 \\
1 & -1.2735263 & 0 & 0 & 1.621869 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1.2735263 & 0 & 0 & 1.621869 & 0 & 0 & 0 \\
1 & 0 & 0 & -1.2735263 & 0 & 0 & 1.621869 & 0 & 0 & 0 \\
1 & 0 & 1.2735263 & 0 & 0 & 1.621869 & 0 & 0 & 0 & 0 \\
1 & 0 & -1.2735263 & 0 & 0 & 1.621869 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

(4.7)
\[
(X'X) = \begin{bmatrix}
0 & 14.2779 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 14.2779 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
14.2779 & 0 & 0 & 0 & 18.0374 & 6.0125 & 6.0125 & 0 & 0 & 0 \\
14.2779 & 0 & 0 & 0 & 6.0125 & 18.0374 & 6.0125 & 0 & 0 & 0 \\
14.2779 & 0 & 0 & 0 & 6.0125 & 6.0125 & 18.0374 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 6.0125 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 6.0125 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.0125 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.0125 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.0125 \\
\end{bmatrix}
\]

\[
(X'X)^{-1} = \begin{bmatrix}
0.2735 & 0 & 0 & 0 & -0.1299 & -0.1299 & -0.1299 & 0 & 0 & 0 \\
0 & 0.0700 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.0700 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.1299 & 0 & 0 & 0 & 0.1282 & 0.0451 & 0.0451 & 0 & 0 & 0 \\
-0.1299 & 0 & 0 & 0 & 0.0451 & 0.1282 & 0.0451 & 0 & 0 & 0 \\
-0.1299 & 0 & 0 & 0 & 0.0451 & 0.0451 & 0.1282 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.1663 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1663 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1663 \\
\end{bmatrix}
\]

4.2.6.1 The parameter estimation for GP1G

\[
(X'Y_{GP1G}) = \begin{bmatrix}
122.8 \\
-72 \\
90.6 \\
16.5 \\
762.3 \\
710 \\
748.4 \\
15.9 \\
-34.3 \\
-15.9 \\
\end{bmatrix}
\]

\[
\hat{B} = (X'X)^{-1}X'Y_{GP1G} = \begin{bmatrix}
47.3911 \\
-5.0452 \\
6.3461 \\
1.1566 \\
3.9485 \\
-0.3983 \\
2.7962 \\
2.6509 \\
-5.7095 \\
-2.6509 \\
\end{bmatrix}
\]
4.2.6.2 The parameter estimation for GP2G

The equation (4.8) is obtained using the design matrix \( \mathbf{X} \) given in equation (4.7) then GP2G is given as:

\[
\begin{bmatrix}
1097 \\
20.6 \\
-10 \\
39.3 \\
620.8 \\
692.1 \\
718.7 \\
36.8 \\
17.2 \\
6.1
\end{bmatrix}
\]

hence

\[
\hat{\mathbf{B}} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}_{gp2g} =
\begin{bmatrix}
36.1296 \\
1.4399 \\
-0.6980 \\
2.7503 \\
6.6025 \\
8.8179 \\
6.1174 \\
2.8548 \\
1.0196
\end{bmatrix}
\]  
(4.10)

4.2.6.3 The parameter estimation for GP3G

The equation (4.8) is obtained using the design matrix \( \mathbf{X} \) given in equation (4.7) then GP3G is given as:

\[
\begin{bmatrix}
1519 \\
27.3 \\
28.4 \\
-92.8 \\
894.6 \\
920.2 \\
985.1 \\
-65 \\
-1.2 \\
-28.2
\end{bmatrix}
\]

hence

\[
\hat{\mathbf{B}} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}_{gp3g} =
\begin{bmatrix}
51.7514 \\
1.9127 \\
1.9876 \\
-6.5028 \\
3.2471 \\
5.3785 \\
10.7726 \\
-10.8073 \\
-0.2039 \\
-4.6900
\end{bmatrix}
\]  
(4.11)
4.3 Obtaining the generalized variance of the estimated linear parameters

This section deals with the residual variance, first order model variance-covariance and the second order model variance-covariance.

4.3.1 Residual variance

This section gives the residual variance for GP1G, GP2G and GP3G.

4.3.1.1 The residual variance for GP1G

Using table 4.4 the residuals are given as; \( e = y_i - \hat{y} = y_i - X\beta \).

Residuals = estimated errors = \( e = y_i - \hat{y} = y_i - X\beta \) =

\[
\begin{bmatrix}
6.5658 & 9.8932 & 0.8807 & 2.4467 & 0.7354 & 0.7527 & 0.8807 & 2.4467 & 0.3883 & 0.1651 & 0.3270 & 1.2265 & 0.28615 & 0.4238 \\
7.9892 & 0.12068 & 1.0341 & 0.3883 & 4.1651 & -6.3270 & 1.2265 & -2.8615 & 4.2383 \\
\end{bmatrix}
\]

MSE = var(e) = \( (e^*e)/n) \) = 665.8369/14 = 47.5598 (4.12)

4.3.1.2 The residual variance for GP2G

Using table 4.6 the residual variance for GP2G was calculated as below.

Residuals = estimated errors = \( e = y_i - \hat{y} = y_i - X\beta \) =

\[
\begin{bmatrix}
-4.3804 & -5.1918 & 4.0740 & 3.2626 \\
3.0855 & 7.2740 & -5.8240 & -1.6354 & 5.4422 & -1.6034 & 7.0327 & -0.0129 & -3.6742 & 3.5805 \\
\end{bmatrix}
\]

MSE = var(e) = \( (e^*e)/n) \) = 612.7118/14 = 43.7651 (4.13)

4.3.1.3 The residual variance for GP3G

Using table 4.8 residual variance for GP3G was calculated as below.
Residuals = estimated errors = $e = y_i - \hat{y}_i = y_i - X\hat{\beta}$ =

\[-6.3952 \quad 6.3406 \quad -7.4937 \quad 5.2421 \quad 5.3926 \]

\[-6.8717 \quad 8.4921 \quad -3.7722 \quad 6.1965 \quad 5.0980 \quad 2.7041 \quad 0.38025 \quad 4.7605 \quad -1.2443 \quad -2.2638 \]

\[8.5535 \quad -5.6047 \quad 0.5076 \quad -2.4536 \quad 0.4182 \quad -5.9417 \quad -2.5047 \quad -0.0059 \quad 0.0565]'

Residual variance: $MSE = \tilde{\text{var}}(e) = (e^*e)/(n-p) = 595.4813/14 = 42.5344$ \hspace{1cm} (4.14)

### 4.3.2 First order model variance-covariance

This section gives the first order variance for the parameter estimates and variance-covariance for the three groups.

#### 4.3.2.1 The variance-covariance for GP1G

$$SSE = y'y - \hat{y}'\hat{y} = y'y - \hat{\beta}'X'X\hat{\beta} = y'y - \hat{\beta}'X'y = 64920 - 63790 = 1130$$

\[\hat{\sigma}^2 = MSE = \frac{SSE}{n-p} = \frac{1130}{24-4} = 56.5 \hspace{1cm} (4.15)\]

\[\text{var}(\hat{\beta}) = \hat{\sigma}^2 [X'X]^{-1} = \begin{bmatrix} 2.3542 & 0 & 0 & 0 \\ 0 & 3.9572 & 0 & 0 \\ 0 & 0 & 3.9572 & 0 \\ 0 & 0 & 0 & 3.9572 \end{bmatrix} \hspace{1cm} (4.16)\]

The variance of all regression model parameter estimates is found as follows:

$$V(\hat{\beta}) = \hat{\sigma}^2 (\text{diagonal elements of } (XX)^{-1})$$

Therefore, $V(\hat{\beta}_0) = 2.3542$, $V(\hat{\beta}_1) = V(\hat{\beta}_2) = V(\hat{\beta}_3) = 3.9572$
4.3.2.2 The variance-covariance for GP2G

\[ \text{SSE} = y' y - \hat{y}' \hat{y} = y' y - \hat{\beta}' X' X \hat{\beta} = y' y - \hat{\beta}' X' y = 52001 - 50287 = 1714 \]
\[ \hat{\sigma}^2 = \text{MSE} = \frac{\text{SSE}}{n - p} = \frac{1714}{24 - 4} = 85.7 \]

\[ \text{var}(\beta) = \sigma^2 [X'X]^{-1} = \begin{bmatrix} 3.5708 & 0 & 0 & 0 \\ 0 & 6.0023 & 0 & 0 \\ 0 & 0 & 6.0023 & 0 \\ 0 & 0 & 0 & 6.0023 \end{bmatrix} \]

The variance of all regression model parameter estimates is found as follows:

\[ V(\hat{\beta}) = \sigma^2 (\text{diagonal elements of } (XX)^{-1}) \]

Therefore, \( V(\hat{\beta}_0) = 3.5708 \), \( V(\beta_1) = V(\beta_2) = V(\beta_3) = 6.0023 \).

4.3.2.3 The variance-covariance for GP3G

\[ \text{SSE} = y' y - \hat{y}' \hat{y} = y' y - \hat{\beta}' X' X \hat{\beta} = y' y - \hat{\beta}' X' y = 99219 - 96852 = 2367 \]
\[ \hat{\sigma}^2 = \text{MSE} = \frac{\text{SSE}}{n - p} = \frac{2367}{24 - 4} = 118.35 \]

\[ \text{var}(\beta) = \sigma^2 [X'X]^{-1} = \begin{bmatrix} 4.9312 & 0 & 0 & 0 \\ 0 & 8.2890 & 0 & 0 \\ 0 & 0 & 8.2890 & 0 \\ 0 & 0 & 0 & 8.2890 \end{bmatrix} \]

The variance of all regression model parameter estimates was found as follows:

\[ V(\hat{\beta}) = \sigma^2 (\text{diagonal elements of } (XX)^{-1}) \]
Therefore, \( V(\beta_0) = 4.9312 \), \( V(\beta_1) = V(\beta_2) = V(\beta_3) = 8.2890 \).

### 4.3.3 The second order model variance-covariance

This section deals with the second order standard error and variance-covariance for the three groups.

#### 4.3.3.1 The standard error, variance-covariance for GP1G

\[
SSE = y'y - \hat{y}'\hat{y} = y'y - \hat{\beta}'X'y = 64920 - 64254 = 666 \\
\sigma^2 = \frac{SSE}{n-p} = \frac{666}{24-10} = 47.5714 \tag{4.21}
\]

The covariance matrix of the least squares estimator depends on the variance of the disturbances and on the derivative matrix \( X \).

\[
\text{var}(\beta) = \sigma^2 [X'X]^{-1} = \begin{bmatrix}
13.0102 & 0 & 0 & 0 & -6.1791 & -6.1791 & -6.1791 & 0 & 0 & 0 \\
0 & 3.3318 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 3.3318 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 3.3318 & 0 & 0 & 0 & 0 & 0 & 0 \\
-6.1791 & 0 & 0 & 0 & 6.0996 & 2.1435 & 2.1435 & 0 & 0 & 0 \\
-6.1791 & 0 & 0 & 0 & 2.1435 & 6.0996 & 2.1435 & 0 & 0 & 0 \\
-6.1791 & 0 & 0 & 0 & 2.1435 & 2.1435 & 6.0996 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 7.9121 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7.9121 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7.9121
\end{bmatrix} \tag{4.22}
\]

The variance of all regression model parameter estimates was found as follows:

\[
V(\hat{\beta}) = \sigma^2 \text{(diagonal elements of } (XX)^{-1})
\]
Therefore, \( V(\hat{\beta}_0) = 13.0102 \), \( V(\hat{\beta}_1) = V(\hat{\beta}_2) = V(\hat{\beta}_3) = 3.3318 \)

\[
V(\hat{\beta}_{ij}) = V(\hat{\beta}_{22}) = V(\hat{\beta}_{33}) = 6.0996, \quad V(\hat{\beta}_{12}) = V(\hat{\beta}_{13}) = V(\hat{\beta}_{23}) = 7.9121
\]  

(4.23)

4.3.3.2 The standard error, variance-covariance for GP2G

\[
SSE = y'y - \hat{y}'\hat{y} = y'y - \hat{\beta}'X'y = 52001 - 51388 = 613
\]

\[
\hat{\sigma}^2 = \frac{SSE}{n-p} = \frac{613}{24-10} = 43.7857
\]  

(4.24)

\[
\text{var}(\beta) = \hat{\sigma}^2 [X'X]^{-1} = \\
\begin{bmatrix}
11.9748 & 0 & 0 & 0 & -5.6874 & -5.6874 & -5.6874 & 0 & 0 & 0 \\
0 & 3.0667 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 3.0667 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 3.0667 & 0 & 0 & 0 & 0 & 0 & 0 \\
-5.6874 & 0 & 0 & 0 & 5.6142 & 1.9729 & 1.9729 & 0 & 0 & 0 \\
-5.6874 & 0 & 0 & 0 & 1.9729 & 5.6142 & 1.9729 & 0 & 0 & 0 \\
-5.6874 & 0 & 0 & 0 & 1.9729 & 1.9729 & 5.6142 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 7.2825 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7.2825 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7.2825 \\
\end{bmatrix}
\]  

(4.25)

The variance of all regression model parameter estimates was found as follows:

\[
V(\hat{\beta}) = \sigma^2 (\text{diagonal elements of } (XX)^{-1})
\]

Therefore, \( V(\hat{\beta}_0) = 11.9748 \), \( V(\hat{\beta}_1) = V(\hat{\beta}_2) = V(\hat{\beta}_3) = 3.0667 \),

\[
V(\hat{\beta}_{11}) = V(\hat{\beta}_{22}) = V(\hat{\beta}_{33}) = 5.6142, \quad V(\hat{\beta}_{12}) = V(\hat{\beta}_{13}) = V(\hat{\beta}_{23}) = 7.2825
\]
4.3.3.3 The standard error, variance-covariance for GP3G

\[ SSE = y' - \hat{y}' = y' - \hat{\beta}'X'X\hat{\beta} = y' - \hat{\beta}'X'y = 99219 - 98624 = 595 \]

\[ \hat{\sigma}^2 = \frac{SSE}{n - p} = \frac{595}{24 - 10} = 42.5 \]

\[ \text{var}(\beta) = \hat{\sigma}^2 [X'X]^{-1} = \]

\[ \begin{bmatrix}
11.6232 & 0 & 0 & 0 & -5.5204 & -5.5204 & -5.5204 & 0 & 0 & 0 \\
0 & 2.9766 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2.9766 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 2.9766 & 0 & 0 & 0 & 0 & 0 & 0 \\
-5.5204 & 0 & 0 & 0 & 5.4493 & 1.9150 & 1.9150 & 0 & 0 & 0 \\
-5.5204 & 0 & 0 & 0 & 1.9150 & 5.4493 & 1.9150 & 0 & 0 & 0 \\
-5.5204 & 0 & 0 & 0 & 1.9150 & 1.9150 & 5.4493 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 7.0686 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7.0686 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7.0686
\end{bmatrix} (4.27) \]

The variance of all regression model parameter estimates was found as follows:

\[ V(\hat{\beta}) = \hat{\sigma}^2 (\text{diagonal elements of } (XX)^{-1}) \]

Therefore, \( V(\hat{\beta}_0) = 11.6232 \), \( V(\hat{\beta}_1) = V(\hat{\beta}_2) = V(\hat{\beta}_3) = 2.9766 \), \( V(\hat{\beta}_{11}) = V(\hat{\beta}_{22}) = V(\hat{\beta}_{33}) = 5.4493 \), \( V(\hat{\beta}_{12}) = V(\hat{\beta}_{13}) = V(\hat{\beta}_{23}) = 7.0686 \).

4.4 Fitting and testing the three models adequacies

This section fits first order and second order models for the three groups, tests the significance of parameter estimates, analyses the variance, lack of fit and coefficient of determinations.

4.4.1 The first-Order Model

The first-order model with three independent variables is expressed as:
\[
\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3; \tag{4.28}
\]

where:

\(\hat{y}\) represents the predicted response where, in this research, is the yield amount of rose coco bean in grams.

\(\beta_0\) is the mean response, that is the amount of rose coco beans when all the explanatory factors are zero.

\(\beta_1\) is the parameter associated with the nitrogen fertilizer.

\(\beta_2\) is the parameter associated with the phosphorus fertilizer.

\(\beta_3\) is the parameter associated with the potassium fertilizer.

\(x_1\) represents the observed yield on the rose coco bean under nitrogen at seven levels.

\(x_2\) is the effect of phosphorus on rose coco beans, which has been controlled at seven levels.

\(x_3\) is the effect of potassium on rose coco beans, which has been controlled at seven levels.

When there was a curvature in the response surface, then a higher degree polynomial was used. In any model, the levels of each factor are independent of the levels of other factors. The method of least square was used to estimate the parameters in the polynomials in which the response surface analysis was performed by using the fitted surface.

### 4.4.1.1 The first-order model fitting for GP1G

The first order regression equation for GP1G is given as in equation (4.4).

\[
\hat{Y}_{gp1G} = 51.1667 - 5.0452 X_1 + 6.3461 X_2 + 1.1566 X_3 \tag{4.29}
\]

In GP1G the first order model was significant (p=0.0057) at 1%. The nitrogen (p=0.0196) and phosphorus (p=0.0046) were also significant factors, potassium fertilizer at 30 grams is not contributing significantly to the rose coco yield. The coefficient of determination, \(R^2=45.87\%\) is
below 50% so the response is not well explained by the factors chosen though the combined linear factors are significant.

4.4.1.2 The first-order model fitting for GP2G

Therefore the first order regression equation for GP2G is given as equation (4.5);

\[
\hat{Y}_{GP2G} = 45.7083 + 1.4399X_1 - 0.6980X_2 + 2.7503X_3
\]  

(4.30)

In GP2G the first order model is insignificant (p=0.6462). The nitrogen, phosphorus and potassium are not significant factors. The \(R^2=7.78\%\), this is very low far much below, so the response is not well explained by the factors chosen in the first order model.

4.4.1.3 The first-order model fitting for GP3G

Therefore the first order regression equation for GP3G is given as

\[
\hat{Y}_{GP3G} = 63.2917 + 1.9127X_1 + 1.9876X_2 - 6.5028X_3
\]

(4.31)

In GP3G the first order model is insignificant (p=0.1454). The nitrogen and phosphorus are insignificant factors, potassium (p=0.0352) is significant. The \(R^2=23.14\%\), this is below 50% so the response is not well explained by the factors chosen in the first order model confirming its insignificance.

4.4.2 The Analysis of variance of GP1G

The total sum of squares is given as;

\[
SST = \sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} y_i^2 - \frac{\sum_{i=1}^{n} y_i}{n} - n\bar{y}^2
\]
and the sum of squares due to regression is given by

\[ SSR = \hat{\beta}' X'y - \bar{y}^2 \]

\[ = 63790 - 62832.7485 = 957.2515 \]  

(4.33)

The sum of squares due to error is

\[ SSE = y'y - \hat{\beta}' X'y \]

\[ = 64920 - 63790 = 1130 \]  

(4.34)

which verifies that

\[ SST = SSR + SSE. \]  

(4.35)

Using the result of the equations (4.32), (4.33) and (4.34) the analysis of variance table generated from this model is as follows:

Table 4.10: The Analysis of Variance table first order for GP1G.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Mean sum of squares</th>
<th>Variance ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>3</td>
<td>957.2515</td>
<td>319.0838</td>
<td>5.65</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>1130</td>
<td>56.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>2087.2515</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sum of Squared Residuals: 1129.79190

Predicted Residual SS (PRESS): 1672.86356

4.4.2.1 The test of significance in regression-GP1G

A good estimated regression model should explain the variation of the dependent variable in the sample. However, there are certain tests of hypotheses about the model parameters that could
help the experimenter in measuring the effectiveness of the model. The first of all these tests requires the error term $e_i$’s to be normally, identically and independently distributed with mean zero and variance $\sigma^2$. In order to check this assumption, the normal probability of residuals for the rose coco beans yield is graphed as shown in figure 4.4.

4.4.2.2 Normal probability plot of the residuals GP1G

![Normal probability plot for GP1G](image)

Figure 4. 4: The Normal Probability Plot of GP1G.

The residuals plot in figure 4.4 is approximately along a straight line, thus the normality assumption is satisfied. It is important to note that the error term is the difference between the observed value $y_i$ and the corresponding fitted value $\hat{y}_i$, that is, $e_i = y_i - \hat{y}_i$ as shown in table 4.4. As a result of this assumption, observations $y_i$ are also normally, identically and independently distributed.
4.4.2.3 The test of significance of the model

As a result of the normality assumption being satisfied, observations $y_i$ are also normally, identically and independently distributed. Therefore, the test for the significance of the regression could be applied to determine the relationship between the dependent variable $y$ and independent variables $X_1, X_2, X_3$. The hypotheses are;

$H_0: \beta_1 = \beta_2 = \beta_3 = 0$, against

$H_1: \beta_i \neq 0$ for at least one $i$.

From the analysis of variance table 4.10, $F_c = 5.65$. Comparing this value with the $F$-table value $F_{0.05, 3, 20} = 3.10$, we found out that there was a significant statistical evidence to reject the null hypothesis. It implies that at least one of the independent variables, nitrogen, phosphorus or potassium, contributes significantly to the model, therefore the first order model was adequate for GP1G. Further tests were carried out on the parameters $\beta_1, \beta_2$ and $\beta_3$ in order to identify the variable that significantly contributes to the model.

4.4.2.4 The test of significance of parameter estimates in first order model

The hypothesis test was used to gain a rough idea of the importance of the treatment effects. In order to determine whether given variables were justified to be included or excluded from the model, the test of hypotheses for the individual regression coefficients were undertaken as follows:

4.4.2.5 Test for $\beta_1$

Hypothesis

$H_0: \beta_1 = 0$, against

$H_1: \beta_1 \neq 0$. 
The standard error for $\hat{\beta}_1$ (S.E $\hat{\beta}_1$), was found by use of the fact that $\text{Cov}(\hat{\beta}) = \text{MSE}(X'X)^{-1}$.

Thus, using the results in table 4.10 and the diagonal element of $(X'X)^{-1}$ corresponding to this parameter estimate in equation (4.16), the standard error is given as,

$$\text{S.E } \hat{\beta}_1 = (3.9572)^{1/2} = 1.98927$$

$$t = \frac{\hat{\beta}_1}{\sqrt{\text{var}(\hat{\beta}_1)}} \sim t_{(\alpha, n-p)}$$  \hspace{1cm} (4.36)

$$t_c = -5.0452/1.9887 = -2.5362$$

while $t_{0.025,20} = 2.086$.

Since $-t_{\alpha/2} < t_c < t_{\alpha/2}$, we reject the null hypothesis, thus the parameter was significant and the predictor variable-nitrogen ($X_1$) was required in explaining the variation of the rose coco yield at $\alpha = 0.05$ in the first order model.

**4.4.2.6 Test for $\beta_2$**

Hypothesis

$H_0: \beta_2 = 0$, against

$H_1: \beta_2 \neq 0$.

The standard error of the parameter estimate, $\hat{\beta}_2$ given by,

$$\text{S.E } \hat{\beta}_2 = (3.9572)^{1/2} = 1.98927$$

$$t = \frac{\hat{\beta}_2}{\sqrt{\text{var}(\hat{\beta}_2)}} \sim t_{(\alpha, n-p)}$$  \hspace{1cm} (4.37)

$$t_c = 6.3461/1.9887 = 3.1902$$
Since \(-t_{a/2} < t_c < t_{a/2}\), we reject the null hypothesis, thus the parameter was significant and the predictor variable-phosphorus \((X_2)\) was required in explaining the variation of the *rose coco* yield at \(\alpha = 0.05\) in the first order model.

**4.4.2.7 Test for \(\beta_3\)**

Hypothesis

\[ H_0: \beta_3 = 0, \quad \text{against} \quad H_1: \beta_3 \neq 0. \]

The standard error of the parameter estimate, \(\hat{\beta}_3\) is given as;

\[ S.E(\hat{\beta}_3) = (3.9572)^{1/2} = 1.98927 \]

\[
t = \frac{\hat{\beta}_3}{\sqrt{\text{var}(\hat{\beta}_3)}} \sim t_{(n-p)} \]

(4.38)

\[ t_c = 1.1566/1.9887 = 0.5814 \]

Since \(-t_{a/2} < t_c < t_{a/2}\), we do not reject the null hypothesis, thus the parameter was insignificant and the predictor variable-potassium \((X_3)\) was not individually important in explaining the variation of the *rose coco* yield at \(\alpha = 0.05\) in the first order model.

The coefficient of multiple determination is given as;

\[ R^2 = 1 - \frac{SS_E}{SS_T} \quad \text{or} \]

\[ R^2 = \frac{SSR}{SST} = \frac{957.2515}{2087.2515} = 0.4586 \]

(4.39)

which indicates that 45.86 % of the variation of the *rose coco* yield was accounted for by the model, which was below half (50%) to justify the correct relationship between the predictors and the response. The \(R^2\) measures how well estimated model fits the data. When \(R^2\) is closer to
the 100%, the better the estimation of the regression equation fits the sample data. The first order model for GP1G \( \hat{Y}_{GP1G} = 51.1667 - 5.0452X_1 + 6.3461X_2 + 1.1566X_3 \) was significant (p<0.05). The N and P contribute significantly to the model. K was insignificant. So the model with significant terms only is \( \hat{Y}_{GP1G} = 51.1667 - 5.0452X_1 + 6.3461X_2 \)

### 4.4.3 The analysis of variance of GP2G

The sum of squares due to total is given as

\[
SST = y'y - n\bar{y}^2
\]

\[
= 52001 - 24 \times 45.708^2 = 1859.6897 \quad (4.40)
\]

and the sum of squares due to regression was given as

\[
SSR = \hat{\beta}'X'y - n\bar{y}^2
\]

\[
= 50287 - 50141.3103 = 145.6897 \quad (4.41)
\]

The sum of squares due to error is

\[
SSE = y'y - \hat{\beta}'X'y
\]

\[
= 52001 - 50287 = 1714 \quad (4.42)
\]

which verifies that

\[
SST = SSR + SSE. \quad (4.43)
\]

Using the result of the equations (4.40), (4.41) and (4.42) the analysis of variance table generated from this model is as follows:
Table 4.11: The Analysis of Variance table first order for GP2G.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Mean sum of squares</th>
<th>Variance ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>3</td>
<td>145.6897</td>
<td>48.5632</td>
<td>0.567</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>1714</td>
<td>85.7</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>1859.6897</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.4.3.1 The test of significance in regression-GP2G

In order to check this assumption, the normal probability of residuals for the *rose coco* beans yield is graph as shown in Figure 4.5.

4.4.3.2 Normal probability plot of the residuals GP2G

![Normal probability plot for GP2G](image)

Figure 4.5: The Normal Probability Plot of GP2G.
The residuals plot in figure 4.5 is approximately along a straight line, thus the normality assumption is satisfied.

4.4.3.3 The test of significance of the model

As a result of the normality assumption being satisfied, observations $y_i$ are also normally, identically and independently distributed. Therefore, the test for the significance of the regression could be applied to determine if the relationship between the dependent variable $y$ and independent variables $X_1, X_2, X_3$ exists. The hypotheses are,

$H_0: \beta_1 = \beta_2 = \beta_3 = 0$, against

$H_1: \beta_i \neq 0$ for at least one $i$.

From the analysis of variance table 4.11, $F_c = 0.567$. Comparing this value with the $F$-table value $F_{0.05, 3, 20} = 3.10$, we find that we do not reject the null hypothesis. It implies that the independent variables, nitrogen, phosphorus or potassium, contributes insignificantly to the model, therefore the first order model is inadequate for GP2G. We now do not carry out further tests on the parameters $\beta_1, \beta_2$ and $\beta_3$ in order to identify the variable that significantly contributes to the model.

The coefficient of multiple determination is given as;

$$R^2 = \frac{145.6897}{1859.6897} = 0.0783$$  \hspace{1cm} (4.44)

which indicates that 7.83 % of the variation of the rose coco yield was accounted for by the model, which was a very low value to justify the correct relationship between the predictors and the response. We test the first order model and it was insignificant ($p=0.6462$). The first order model was given by this equation:-

$$\hat{Y}_{GP2G} = 45.7083 + 1.4399X_1 - 0.6980X_2 + 2.7503X_3$$  \hspace{1cm} (4.45)
4.4.4 The analysis of variance of GP3G

The sum of squares due to total was given as

\[ \text{SST} = y'y - n\bar{y}^2 \]

\[ = 99219-24\times63.2917^2 = 3078.8571 \]  \hspace{0.5cm} (4.46)

and the sum of squares due to regression was given as;

\[ \text{SSR} = \hat{\beta}'X'y - n\bar{y}^2 \]

\[ = 96852-24\times63.2917^2 = 711.8571 \]  \hspace{0.5cm} (4.47)

The sum of squares due to error is

\[ \text{SSE} = y'y - \hat{\beta}'X'y \]

\[ = 99219-96852 = 2367 \]  \hspace{0.5cm} (4.48)

which verifies that

\[ \text{SST} = \text{SSR} + \text{SSE} \].  \hspace{0.5cm} (4.49)

Using the result of the equations (4.46), (4.47) and (4.48) the analysis of variance table generated from this model was as follows:

Table 4. 12: The Analysis of Variance table first order for GP3G.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Mean sum of squares</th>
<th>Variance ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>3</td>
<td>711.8571</td>
<td>237.2857</td>
<td>2.005</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>2367</td>
<td>118.35</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>3078.8571</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.4.4.1 The test for significance in regression-GP3G

In order to check this assumption, the normal probability of residuals for the rose coco beans yield was graph as shown in figure 4.6.

4.4.4.2 Normal probability plot of the residuals GP3G

Figure 4.6: The Normal probability plot for GP3G.

The residuals plot in figure 4.6 was approximately along a straight line, thus the normality assumption is satisfied.

4.4.4.3 The test of significance of the model

As a result of the normality assumption being satisfied, observations $y_i$ are also normally, identically and independently distributed. Therefore, the test for the significance of the regression could be applied to determine if the relationship between the dependent variable $y$ and independent variables $X_1, X_2, X_3$ exists. The hypotheses are,
\[ H_0: \beta_1 = \beta_2 = \beta_3 = 0, \text{ against} \]

\[ H_1: \beta_i \neq 0 \text{ for at least one } i. \]

From the analysis of variance table 4.12, \( F_c = 2.005 \). Comparing this value with the F-table value \( F_{0.05, 3, 20} = 3.10 \), we found out that we do not reject the null hypothesis. It implies that the independent variables, nitrogen, phosphorus or potassium, contributes insignificantly to the model, therefore the first order model is inadequate for GP3G.

Therefore, there was no need to go further tests on the parameters \( \beta_1, \beta_2 \) and \( \beta_3 \) in order to identify the variable that significantly contribute to the model.

The coefficient of multiple determination was given as;

\[
R^2 = \frac{711.8571}{3078.8571} = 0.2312
\]  

(4.50)

which indicates that 23.12 % of the variation of the rose coco yield was accounted for by the first order model, which was a rather low value to justify the correct relationship between the predictors and the response. If we test the first order model it was insignificant (\( p=0.1454 \)). The first order model was given by this equation:-

\[
\hat{Y}_{GP3G} = 63.2917 + 1.9127X_1 + 1.9876X_2 - 6.5028X_3
\]  

(4.51)

\subsection*{4.4.5 Design for fitting the second-order model-GP1G}

This section fits the second order model, test the significance of parameter estimates in second order model and lack of fit for the three replicates of GP1G.

\subsubsection*{4.4.5.1 Fitting of the second-order model}

The second-order polynomial model for three factors was used to fit the data given below:
\[
\hat{Y} = \beta_0 + \sum_{i=1}^{3} \beta_i x_i + \sum_{i=1}^{3} \beta_{ii} x_i^2 + \sum_{i=1}^{3} \sum_{j=i+1}^{3} \beta_{ij} x_i x_j
\]  

(4.52)

where \( \hat{Y} \) is the predicted response used as a dependent variable; \( x_i \) are the coded levels of independent variables; \( \beta_0 \) intercept; \( \beta_i \) the linear terms coefficients; \( \beta_{ii} \) the quadratic terms coefficients and \( \beta_{ij} \) the interaction terms coefficients.

We explore the second order model fit to verify and explore the curvature if it exists. The second-order model includes all the terms in the first-order model plus all quadratic terms and all cross product terms. The model was expressed as with three independent variables:

\[
Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + e
\]  

(4.53)

The fitted second order model in terms of coded factors for GP1G in equation (4.9) is given as:

\[
\hat{Y}_{gp1g} = 47.3911 - 5.0452 X_1 + 6.3461 X_2 + 1.1566 X_3 + 3.9485 X_1^2 - 0.3983 X_2^2 + 2.7962 X_3^2
\]
\[
+ 2.6509 X_1 X_2 - 5.7095 X_1 X_3 - 2.6509 X_2 X_3
\]  

(4.54)

### 4.4.5.2 The test of significance of parameter estimates in second order model

We undertook the test of hypotheses for the individual regression coefficients as follows to see its significance on the model:

#### 4.4.5.3 Test for \( \beta_i \)

Hypothesis

\( H_0: \beta_i = 0, \) for \( i=1,2,3 \) against

\( H_1: \beta_i \neq 0. \)

The standard error for \( \hat{\beta}_i \), \( S.E \hat{\beta}_i \), was found by using the

\[
\text{Cov}(\hat{\beta}) = \text{MSE}(X'X)^{-1}. \text{ Thus, using the equation (4.22) the diagonal element of } \sigma^2 (X'X)^{-1} \text{ the standard error is given as, } S.E \hat{\beta}_i = (3.3318)^{1/2} = 1.8253
\]
\[ t = \frac{\hat{\beta}_i}{\sqrt{\text{var}(\hat{\beta}_i)}} \sim t_{(\alpha, n-p)} \]

\[ t_n = -5.0452/1.8253 = -2.7640 \]
\[ t_p = 6.3461/1.8253 = 3.4767 \]
\[ t_k = 1.1566/1.8253 = 0.6336 \]

Since \(-t_{a/2} < t_c < t_{a/2}\), we reject the null hypothesis of nitrogen and phosphorus, thus the parameters are significant and the predictor variables-nitrogen and phosphorus are required in explaining the variation of the *rose coco* yield at \(\alpha = 0.05\). The potassium fertilizer was insignificant.

**4.4.5.4 Test for \(\beta_{ii}\)**

Hypothesis

\[ H_0 : \beta_{ii} = 0, \text{ against} \]
\[ H_1 : \beta_{ii} \neq 0. \]

The standard error of the parameter estimate, \(\hat{\beta}_{ii}\) given as;

\[ \text{S.E} \hat{\beta}_{ii} = (6.0996)^{1/2} = 2.4697 \]

\[ t = \frac{\hat{\beta}_i}{\sqrt{\text{var}(\hat{\beta}_i)}} \sim t_{(\alpha, n-p)} \]

\[ t_n^2 = 3.9485/2.4697 = 1.5988 \]
\[ t_p^2 = 0.3983/2.4697 = 0.1613 \]
\[ t_k^2 = 2.7962/2.4697 = 1.1322 \]

Since \(-t_{a/2} < t_c < t_{a/2}\), we accept the null hypothesis, thus the parameters were insignificant and the predictor quadratic variables of nitrogen, phosphorus and potassium were not required in explaining the variation of the *rose coco* yield at \(\alpha = 0.05\).
4.4.5.5 Test for $\beta_{ij}$

Hypothesis

$H_0: \beta_{ij} = 0$, against $H_1: \beta_{ij} \neq 0$.

The standard error of the parameter estimate, $\hat{\beta}_{ij}$ given as;

$$S.E \hat{\beta}_{ij} = (7.9121)^{1/2} = 2.8128$$

$$t_c = \frac{\hat{\beta}_{ij}}{\sqrt{\text{var}(\hat{\beta}_{ij})}} \sim t_{(n,p)}$$

$t_{np} = 2.6509/2.8128 = 0.9424$

$t_{nk} = 5.7095/2.8128 = 2.0298$ (4.57)

$t_{pk} = 2.6509/2.8128 = 0.9424$

Since $-t_{u/2} < t_c < t_{u/2}$, we accept the null hypothesis, thus the parameters were insignificant and the predictor interactions variables of nitrogen and phosphorus, nitrogen and potassium, phosphorus and potassium were not required in explaining the variation of the rose coco yield at $\alpha = 0.05$.

Table 4.13: The ANOVA results on the three replicate of rose coco beans-GP1G.

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FO(N, P, K)</td>
<td>3</td>
<td>2816</td>
<td>938.57</td>
<td>1.3969</td>
<td>0.2522</td>
</tr>
<tr>
<td>TWI(N, P, K)</td>
<td>3</td>
<td>842</td>
<td>280.50</td>
<td>0.4175</td>
<td>0.7411</td>
</tr>
<tr>
<td>PQ(N, P, K)</td>
<td>3</td>
<td>925</td>
<td>308.30</td>
<td>0.4589</td>
<td>0.7120</td>
</tr>
<tr>
<td>Residuals</td>
<td>62</td>
<td>41656</td>
<td>671.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of fit</td>
<td>14</td>
<td>1962</td>
<td>140.13</td>
<td>0.1694</td>
<td>0.9996</td>
</tr>
<tr>
<td>Pure error</td>
<td>48</td>
<td>39695</td>
<td>826.97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.13 comes from the three replicates in GP1G to show the subdivision of residuals into lack of fit and pure error.
Table 4.14: The ANOVA results in the fertilizer concentration on *rose coco* beans-GP1G

<table>
<thead>
<tr>
<th>Type I Sum</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Order</td>
<td>3</td>
<td>957.541428</td>
<td>319.18</td>
<td>0.4587</td>
<td>6.71</td>
<td>0.0049</td>
</tr>
<tr>
<td>Pure Quadratic</td>
<td>3</td>
<td>183.455066</td>
<td>61.15</td>
<td>0.0879</td>
<td>1.29</td>
<td>0.3178</td>
</tr>
<tr>
<td>Two-Factor Interaction</td>
<td>3</td>
<td>280.500000</td>
<td>93.50</td>
<td>0.1344</td>
<td>1.97</td>
<td>0.1655</td>
</tr>
<tr>
<td>Residuals</td>
<td>14</td>
<td>665.84</td>
<td>47.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of fit</td>
<td>14</td>
<td>665.84</td>
<td>47.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pure error</td>
<td>0</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Model</td>
<td>9</td>
<td>1421.496494</td>
<td>0.6810</td>
<td>3.32</td>
<td>0.0219</td>
<td></td>
</tr>
</tbody>
</table>

Coefficient of variation (CV)=13.48, coefficient determination ($R^2$)=0.6810, Adjusted $R$-squared =0.4759 correlation coefficient ($r$)=0.8252, root MSE=6.896359, response mean=51.166667, PRESS=2488.850781

Table 4.14 shows the results of the averaged data of the three replicates in GP1G that is why the pure error is zero because pure error emanates from replicated values (center points). The statistical testing of the model was done by the Fisher's statistical test for analysis of variance (ANOVA) and the results as shown in table 4.14 indicates that there was no significant effect of the quadratic components and cross products. However, the total second order model was significant at 0.05 (p=0.0219). The analysis of variance (F-test) showed that the first model fits well with the experimental data. The Coefficient of variation (C.V.) is a measure expressing the standard deviation as a percentage of the mean (Thomas and Nelson, 1996). It indicates the degree of precision with which the treatments are compared. The higher the value of the coefficient of variation (CV), the lower the reliability of the experiment, lower values of CV indicated a very high degree of precision and a good deal of reliability of the experimental values. Here, a lower value of CV (13.48) indicates the greater reliability of the experiment. The goodness of fit of the model was checked by the coefficient of determination ($R^2$) and correlation coefficient (R). The closer value of R (correlation coefficient) to 1, the better the
correlation between the experimental and predicted values. Here, the value of R (0.8252) for equation (4.54) being close to 1 indicated a close agreement between the experimental results and the theoretical values predicted by the model equation. The $R^2$ of quadratic equation was 68.10% with mean response of 51.1667 grams of *rose coco* beans, implying that $R^2$ in GP1G of second order model explain 68.10% of the variation in the model than $R^2=45.87\%$ of the first order model. The model was adequate to express the actual relationship between the response and significant variables, with a satisfactory coefficient of determination ($R^2=0.6810$), which indicated 68.10% of the variability in the response could be explained by the second-order polynomial predictive equation (4.54), which means that the model was unable to explain only 31.9% of the total variations, quadratic model was insignificant having N and P significant factors.

Table 4.15: The estimated effects & coefficients of the empirical model for GP1G

| Parameter | DF | Estimate | Error | t Value | Pr > |t| |
|-----------|----|----------|-------|---------|-------|
| Intercept | 1  | 47.391125| 3.606523| 13.14   | <.0001*** |
| N         | 1  | -5.045204 | 1.825105| -2.76   | 0.0152*  |
| P         | 1  | 6.346081  | 1.825105| 3.48    | 0.0037** |
| K         | 1  | 1.156615  | 1.825105| 0.63    | 0.5365   |
| N*N       | 1  | 3.948538  | 2.469426| 1.60    | 0.1321   |
| P*N       | 1  | 2.650859  | 2.812504| 0.94    | 0.3619   |
| P*P       | 1  | -0.398323 | 2.469426| -0.16   | 0.8742   |
| K*N       | 1  | -5.709542 | 2.812504| -2.03   | 0.0618● |
| K*P       | 1  | -2.650859 | 2.812504| -0.94   | 0.3619   |
| K*K       | 1  | 2.796168  | 2.469426| 1.13    | 0.2765   |

Signif. codes:‘***’significant at 0.001, ‘**’significant at 0.01, ‘*’significant at 0.05, ‘●’significant at 0.1, ‘ ‘ significant at 1.
Table 4.15 lists the regression coefficients and the corresponding p-values for the second-order polynomial model given as in equation (4.54) as:

\[
\hat{Y}_{GP1G} = 47.3911 - 5.0452X_1 + 6.3461X_2 + 1.1566X_3 + 3.9485X_1^2 - 0.3983X_2^2 + 2.7962X_3^2
+ 2.6509X_1X_2 - 5.7095X_1X_3 - 2.6509X_2X_3
\]

The \( \hat{Y}_{GP1G} \) is the predicted response for *rose coco* beans in group 1. The results presented in Table 4.15 suggest that linear effects of nitrogen and phosphorus were primary determining factors on the *rose coco* beans yield as these had the largest coefficients. The interaction effect \( N*K \) was a secondary determining factor and those other terms of the model showed no significant effect on the yield. The positive coefficient, \( P \), enhances the yield the most. However, all the other terms \( N \) and \( N*K \) had negative coefficients. The \( X_1 \) represent Nitrogen (N), \( X_2 \) represent Phosphorus (P), \( X_3 \) represents Potassium (K).

The significance of each coefficient was determined by the student’s t-test and p-value, which are in table 4.15. The larger the magnitude of t-test and smaller p-value, the more significant the corresponding coefficients. It can be seen in table 4.15 that the regression coefficients of the linear terms of nitrogen and phosphorus \( N(0.0152) \) and \( P(0.0037) \), respectively) had significant effects on the yield (p-value <0.05) and the interaction terms in \( K*N \) (potassium and nitrogen-(0.0618)) is marginally significant. Among these, \( P \) was significant at the 1% significance level, while \( N \) was significant at the 5% level.

Table 4.16: The Analysis of Variance Table for N,P,K in 2nd order model- GP1G.

<table>
<thead>
<tr>
<th>Factor</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>4</td>
<td>723.277182</td>
<td>180.819295</td>
<td>3.80</td>
<td>0.0270</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>660.747662</td>
<td>165.186916</td>
<td>3.47</td>
<td>0.0360</td>
</tr>
<tr>
<td>K</td>
<td>4</td>
<td>318.328520</td>
<td>79.582130</td>
<td>1.67</td>
<td>0.2118</td>
</tr>
</tbody>
</table>

We saw that in GP1G ANOVA nitrogen and phosphorus are significant factors at \( p<0.05 \).

Potassium was not significant at \( \alpha=0.05 \) level of significance.
4.4.5.6 The lack-of-fit test for the three replicates GP1G

Lack-of-fit tests were performed on the fitted models. The lack-of-fit test compares the variation around the model with pure variation within replicated observations using the results of table 4.1. The lack of fit measures the adequacy of the quadratic response surface model.

Table 4.17: The lack of fit test for second order model-GP1G.

<table>
<thead>
<tr>
<th>Sum of</th>
<th>Residual</th>
<th>DF</th>
<th>Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of Fit</td>
<td>14</td>
<td>2073.344319</td>
<td>148.096023</td>
<td>0.19</td>
<td>0.9993</td>
<td></td>
</tr>
<tr>
<td>Pure Error</td>
<td>48</td>
<td>37619</td>
<td>783.722222</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Error</td>
<td>62</td>
<td>39692</td>
<td>640.193726</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The lack of fit of the model was insignificant indicating that the second order model fits the data adequately. This is when the three replicates in group one (GP1G) is used.

4.4.6 Design for fitting the second-order model-GP2G

A second-order polynomial model for three factors was used to fit the data of GP2G using the model in equation (4.52). We explore the second order model fit to verify and explore the curvature if it exists. The second-order model includes all the terms in the first-order model plus all quadratic terms and all cross product terms. The model was expressed as with three independent variables as given by equation (4.52).

4.4.6.1 The test of significance of parameter estimates in second order model

We undertook the test of hypotheses for the individual regression coefficients as follows to see its significance on the model:
4.4.6.2 Test for $\beta_i$

Hypothesis

$H_0: \beta_i = 0$, for $i=1,2,3$ against

$H_1: \beta_i \neq 0$.

The standard error for $\hat{\beta}_i$, $S.E. \hat{\beta}_i$, is found by use of the fact that $\text{Cov}(\hat{\beta}) = \text{MSE}(X'X)^{-1}$. Thus, using the equation (4.25) the diagonal element of $\hat{\sigma}^2 (X'X)^{-1}$ the standard error is given as, $S.E. \hat{\beta}_i = (3.0667)^{1/2} = 1.7512$

\[ t = \frac{\hat{\beta}_i}{\sqrt{\text{var}(\hat{\beta}_i)}} \sim t_{(\alpha,n-p)} \]  

(4.58)

\[ t_a = 1.4399/1.7512 = 0.8222 \]

\[ t_p = -0.6980/1.7512 = -0.3986 \]  

(4.59)

\[ t_k = 2.7503/1.7512 = 1.5705 \]

Since $-t_{a/2} < t < t_{a/2}$, we do not reject the null hypothesis, thus the parameters are insignificant and the predictor variables-nitrogen, phosphorus and potassium are not required in explaining the variation of the rose coco yield at $\alpha = 0.05$.

4.4.6.3 Test for $\beta_{ii}$

Hypothesis

$H_0: \beta_{ii} = 0$, against

$H_1: \beta_{ii} \neq 0$.

The standard error of the parameter estimate, $\hat{\beta}_{ii}$ is using equation (4.25),

$S.E. \hat{\beta}_{ii} = (5.6142)^{1/2} = 2.3694$
\[
t = \frac{\hat{\beta}_i}{\sqrt{\text{var}(\hat{\beta}_i)}} \sim t_{(n-n-p)}
\]

\[t_n^2 = 0.6806/2.3694 = 0.2872\]

\[t_p^2 = 6.6025/2.3694 = 2.7865\]  (4.60)

\[t_k^2 = 8.8179/2.3694 = 3.7215\]

Since \(-t_{n/2} < t_c < t_{n/2}\), we reject the null hypothesis of quadratic of P and K, thus the parameter are significant and the predictor quadratic variables of phosphorus and potassium are required in explaining the variation of the rose coco yield at \(\alpha = 0.05\). The quadratic of nitrogen was insignificant.

**4.4.6.4 Test for \(\beta_{ij}\)**

Hypothesis

\[H_0: \beta_{ij} = 0, \text{ against} \]

\[H_1: \beta_{ij} \neq 0.\]

The standard error for the parameter estimate, \(\hat{\beta}_{ij}\), is using equation (4.25),

\[S.E \hat{\beta}_{ij} = (7.2825)^{1/2} = 2.6986\]

\[
t = \frac{\hat{\beta}_i}{\sqrt{\text{var}(\hat{\beta}_i)}} \sim t_{(n-n-p)}
\]

\[t_{np} = 6.1174/2.6986 = 2.2669\]

\[t_{nk} = 2.8548/2.6986 = 1.0579\]  (4.61)

\[t_{pk} = 1.0196/2.6986 = 0.3778\]

Since \(-t_{n/2} < t_c < t_{n/2}\), we reject the null hypothesis of N*P, thus the parameter was significant and the predictor variable-N*P was important in explaining the variation of the rose coco yield at \(\alpha = 0.05\) others are insignificant.
Table 4.18: The ANOVA results on the three replicate of *rose coco* beans-GP2G.

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FO(N, P, K)</td>
<td>3</td>
<td>85.5</td>
<td>28.49</td>
<td>0.1041</td>
<td>0.95740</td>
</tr>
<tr>
<td>TWI(N, P, K)</td>
<td>3</td>
<td>717.8</td>
<td>239.28</td>
<td>0.8739</td>
<td>0.45953</td>
</tr>
<tr>
<td>PQ(N, P, K)</td>
<td>3</td>
<td>2354.8</td>
<td>784.92</td>
<td>2.8669</td>
<td>0.04364</td>
</tr>
<tr>
<td>Residuals</td>
<td>62</td>
<td>16975.1</td>
<td>273.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of fit</td>
<td>14</td>
<td>1962.4</td>
<td>140.17</td>
<td>0.4482</td>
<td>0.94877</td>
</tr>
<tr>
<td>Pure error</td>
<td>48</td>
<td>15012.7</td>
<td>312.76</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table 4.18 comes from the three replicates in GP2G to show the subdivision of residuals into lack of fit and pure error.

Table 4.19: The ANOVA results in the fertilizer concentration on *rose coco* beans-GP2G.

<table>
<thead>
<tr>
<th>Regression</th>
<th>DF</th>
<th>Type I Sum of Squares</th>
<th>Mean sq</th>
<th>R-Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Order</td>
<td>3</td>
<td>144.555104</td>
<td>48.185</td>
<td>0.0778</td>
<td>1.10</td>
<td>0.3815</td>
</tr>
<tr>
<td>Pure Quadratic</td>
<td>3</td>
<td>821.441245</td>
<td>273.814</td>
<td>0.4419</td>
<td>6.26</td>
<td>0.0065</td>
</tr>
<tr>
<td>Two-Factor Interaction</td>
<td>2</td>
<td>280.250000</td>
<td>93.417</td>
<td>0.1508</td>
<td>2.13</td>
<td>0.1417</td>
</tr>
<tr>
<td>Residuals</td>
<td>14</td>
<td>612.71</td>
<td>43.765</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of fit</td>
<td>14</td>
<td>612.71</td>
<td>43.765</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pure error</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Model</td>
<td>9</td>
<td>1246.246349</td>
<td>0.6704</td>
<td>3.16</td>
<td>0.0264</td>
<td></td>
</tr>
</tbody>
</table>

Coefficient of variation (CV)=14.47, coefficient determination (R²)=0.6704, Adjusted R-squared =0.4585, correlation coefficient (r)=0.8188, root MSE=6.615523, response mean=45.708333, PRESS=2091.1898012

Table 4.19 is the results of the averaged data of the three replicates in GP2G. In the table, the combined quadratic terms are significant at 1% with a p-value of 0.0065 also the total model was significant at 5% with a p value of 0.0264. We tested the first order model it was insignificant (p=0.3815). Linear and cross product terms were not significant, which corresponds to the fact that the quadratic model for GP2G was appropriate. Therefore, the analysis of variance (F-test) showed that the second order model fits well with the experimental data in GP2G. The R² of quadratic equation was 67.04% with mean response of 45.7083 grams
of *rose coco* beans, implying that $R^2$ in GP2G of second order model explain 67.04% of the variation in the model than $R^2=7.83\%$ of the first order model. The model was adequate to express the actual relationship between the response and significant variables, with a satisfactory coefficient of determination ($R^2=0.6704$), which indicated 67.04% of the variability in the response could be explained by the second-order polynomial predictive equation (4.62), meaning that the model was unable to explain 32.96% of the total variations. The value of $R$ (0.8188) for equation (4.62) being close to 1 indicated a close agreement between the experimental results and the theoretical values predicted by the model equation.

Table 4.20: The estimated effects & coefficients of the empirical model for GP2G.

| Parameter | DF | Estimate   | Standard Error | t Value | Pr > |t| |
|-----------|----|------------|----------------|---------|------|---|
| Intercept | 1  | 36.129616  | 3.459657       | 10.44   | <.0001*** |
| N         | 1  | 1.439859   | 1.750782       | 0.82    | 0.4246 |
| P         | 1  | -0.697954  | 1.750782       | -0.40   | 0.6962 |
| K         | 1  | 2.750284   | 1.750782       | 1.57    | 0.1385 |
| N*N       | 1  | 0.680614   | 2.368865       | 0.29    | 0.7781 |
| P*N       | 1  | 6.117366   | 2.697972       | 2.27    | 0.0397* |
| P*P       | 1  | 6.602530   | 2.368865       | 2.79    | 0.0145* |
| K*N       | 1  | 2.854771   | 2.697972       | 1.06    | 0.3079 |
| K*P       | 1  | 1.019561   | 2.697972       | 0.38    | 0.7112 |
| K*K       | 1  | 8.17910    | 2.368865       | 3.72    | 0.0023** |

Signif. codes: `'***'` significant at 0.001, `'**'` significant at 0.01, `'*'` significant at 0.05, `'●'` significant at 0.1, `' '` significant at 1.

Table 4.20 lists the regression coefficients and the corresponding p-values for the second-order polynomial model given as in equation 4.62 as:

$$\hat{Y}_{GP2G} = 36.1296 + 1.4399X_1 - 0.6980X_2 + 2.7503X_3 + 0.6806X_1^2 + 6.6025X_2^2 + 8.8179X_3^2 + 6.1174X_1X_2 + 2.8548X_1X_3 + 1.0196X_2X_3$$

(4.62)

The $\hat{Y}_{GP2G}$ is the predicted response for *rose coco* beans in group 2. The results presented in Table 4.20 suggest that quadratic effects of phosphorus ($p=0.0145$) and potassium ($p=0.0023$), the interaction effect N*P ($p=0.0397$) are significant factors on the rose coco bean, among these, K$^2$ is significant at the 1% significance level, while P$^2$ and N*P are significant at the 5% level, those other terms of the model showed no significant effect on the yield. The positive
coefficient of $P^2$ (quadratic phosphorus) $K^2$ (quadratic potassium), N*P enhance the yield since they are the largest coefficients in the model equation (4.62). The $X_1$ represent Nitrogen (N), $X_2$ represent Phosphorus (P), $X_3$ represents Potassium (K). Table 4.20 suggests that quadratic effects of phosphorus ($P^2$), potassium ($K^2$) and interaction N*P were the determining significant factors on the rose coco bean yield as these had the more or less the same coefficients and those other terms of the model showed no significant effect on the yield. The $P^2$, $K^2$ and N*P showed the positive coefficient, meaning all the three (quadratic phosphorus, quadratic potassium and nitrogen and phosphorus combined) enhance the rose coco yield.

4.4.6.5 The lack of fit for the three replicates in GP2G

Lack-of-fit tests were performed on the fitted models. The lack-of-fit test compares the variation around the model with pure variation within replicated observations using the results of table 4.5.

Table 4.21: The lack of fit test for GP2G.

<table>
<thead>
<tr>
<th>Sum of</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual</td>
<td>DF</td>
<td>Squares</td>
<td>Mean Square</td>
<td>F Value</td>
<td>Pr &gt; F</td>
</tr>
<tr>
<td>Lack of Fit</td>
<td>14</td>
<td>1838.135953</td>
<td>131.295425</td>
<td>0.50</td>
<td>0.9236</td>
</tr>
<tr>
<td>Pure Error</td>
<td>48</td>
<td>12692</td>
<td>264.416667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Error</td>
<td>62</td>
<td>14530</td>
<td>234.357031</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The lack of fit of the model was insignificant, so the second order model was adequate for the experiments of GP2G.

Table 4.22: The Analysis of Variance Table for N,P,K in 2nd order model- GP2G.

<table>
<thead>
<tr>
<th>Sum of</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>DF</td>
<td>Squares</td>
<td>Mean Square</td>
<td>F Value</td>
<td>Pr &gt; F</td>
</tr>
<tr>
<td>N</td>
<td>4</td>
<td>307.213714</td>
<td>76.803429</td>
<td>1.75</td>
<td>0.1941</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>578.197324</td>
<td>144.549331</td>
<td>3.30</td>
<td>0.0420</td>
</tr>
<tr>
<td>K</td>
<td>4</td>
<td>769.676867</td>
<td>192.419217</td>
<td>4.40</td>
<td>0.0165</td>
</tr>
</tbody>
</table>
In table 4.22, the phosphorus and potassium were significant factors at 5% level of significance. Nitrogen was not significant at the 5% level of significance.

### 4.4.7 Design for fitting the second-order model GP3G

The second-order polynomial model for three factors was used to fit the data of GP3G using the model equation (4.52). The model was expressed as with three independent variables as given by equation (4.53).

#### 4.4.7.1 The test of significance of parameter estimates in second order model

The hypothesis test was used to gain a rough idea of the importance of the treatment effects. In order to determine whether given variables are justified to be included or excluded from the model, we undertook the test of hypotheses for the individual regression coefficients as follows:

#### 4.4.7.2 Test for $\beta_i$

**Hypothesis**

$H_0: \beta_i = 0, \text{ for } i=1,2,3$ against $H_1: \beta_i \neq 0.$

The standard error for $\hat{\beta}_i$, $\text{S.E} \hat{\beta}_i$, was found by use of the fact that $\text{Cov}(\hat{\beta}) = \text{MSE}(X'X)^{-1}$. Thus, using the equation (4.27) the diagonal element of $\hat{\sigma}^2 (X'X)^{-1}$ the standard error was given as, $\text{S.E} \hat{\beta}_i = \text{S.E} \hat{\beta}_i = (2.9766)^{1/2} = 1.7253$

$$t = \frac{\hat{\beta}_i}{\sqrt{\text{var}(\hat{\beta}_i)}} \sim t_{(\alpha, n-p)}$$

(4.63)

$\text{t}_n = 1.9127/1.7253 = 1.1086$
\[ t_p = \frac{1.9876}{1.7253} = 1.1520 \]  

\[ t_k = \frac{-6.5028}{1.7253} = -3.7691 \]

Since \(-t_{u/2} < t_c < t_{u/2}\), we do not reject the null hypothesis of N and P, thus the parameters were insignificant and the predictor variables-nitrogen and phosphorus was not required in explaining the variation of the *rose coco* yield at \(\alpha = 0.05\). We reject the null hypothesis of K, thus the parameter was significant and the predictor variables-potassium was required in explaining the variation of the *rose coco* yield at \(\alpha = 0.05\).

### 4.4.7.3 Test for \(\beta_{ii}\)

Hypothesis

\[ H_0 : \beta_{ii} = 0, \text{ against} \]

\[ H_1 : \beta_{ii} \neq 0. \]

The standard error of the parameter estimate, \(\hat{\beta}_{ii}\) using equation (4.27),

\[ \text{S. E} \hat{\beta}_{ii} = (5.4493)^{1/2} = 2.3344 \]

\[ t = \frac{\hat{\beta}_{ii}}{\sqrt{\text{var}(\hat{\beta}_{ii})}} \sim t_{(\alpha, n-p)} \]

\[ t_n^2 = \frac{3.2471}{2.3344} = 0.2872 \]

\[ t_p^2 = \frac{5.3785}{2.3344} = 2.7865 \]

\[ t_k^2 = \frac{10.7726}{2.3344} = 3.7215 \]  

(4.65)

Since \(-t_{u/2} < t_c < t_{u/2}\), we reject the null hypothesis of quadratic of P and K, thus the parameters were significant and the predictor quadratic variables of phosphorus and potassium were required in explaining the variation of the *rose coco* yield at \(\alpha = 0.05\). The quadratic effect of nitrogen was insignificant.
4.4.7.4 Test for $\beta_{ij}$

Hypothesis

$H_0: \beta_{ij} = 0$, against

$H_1: \beta_{ij} \neq 0$.

The standard error of the parameter estimate, $\hat{\beta}_{ij}$ using equation (4.27),

$$S.E \hat{\beta}_{ij} = (7.0686)^{1/2} = 2.6587$$

$$t = \frac{\hat{\beta}_{ij}}{\sqrt{\text{var}(\hat{\beta}_{ij})}} \sim t_{(n-p)}$$

$t_{np} = -10.8073/2.6587 = -4.0649$

$t_{nk} = -0.2039/2.6587 = -0.0767$ \hspace{1cm} (4.66)\hspace{1cm}$t_{pk} = -4.6900/2.6587 = -1.7640$

Since $-t_{n/2} < t < t_{n/2}$, we reject the null hypothesis of NP, thus the parameter was significant and the predictor variable-N*P was important in explaining the variation of the *rose coco* yield at $\alpha = 0.05$, others were insignificant.

Table 4.23: The ANOVA results on the three replicate of rose coco beans-GP3G.

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FO(N, P, K)</td>
<td>3</td>
<td>1724.2</td>
<td>574.75</td>
<td>1.1641</td>
<td>0.3307</td>
</tr>
<tr>
<td>TWI(N, P, K)</td>
<td>3</td>
<td>3648.8</td>
<td>1216.28</td>
<td>2.4634</td>
<td>0.0707</td>
</tr>
<tr>
<td>PQ(N, P, K)</td>
<td>3</td>
<td>1875.1</td>
<td>625.05</td>
<td>1.2659</td>
<td>0.2938</td>
</tr>
<tr>
<td>Residuals</td>
<td>62</td>
<td>30612.4</td>
<td>493.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of fit</td>
<td>14</td>
<td>2925.7</td>
<td>208.98</td>
<td>0.3623</td>
<td>0.9793</td>
</tr>
<tr>
<td>Pure error</td>
<td>48</td>
<td>27686.7</td>
<td>576.81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table 4.23 comes from the three replicates in GP3G to show the subdivision of residuals into lack of fit and pure error.

Table 4.24: The ANOVA results in the fertilizer concentration on rose coco bean-GP3G.
Table 4.24 is the results of the averaged data of the three replicates in GP3G. The statistical 
testing of the model was done by the Fisher’s statistical test for analysis of variance (ANOVA) 
and the results are shown in table 4.24. In the table 4.24, all the linear ($p=0.0099$), quadratic 
($p=0.0034$) and cross product ($p=0.0054$) terms were significant at 1%; therefore the total model 
was significant with $p$ values of 0.0011. The second order model for GP3G was highly 
significant. The analysis of variance (F-test) showed that the second model fits well with the 
experimental data. The goodness of fit of the model can be checked by the determination 
coefficient ($R^2$) and correlation coefficient ($r$). The determination coefficient ($R^2$) implies that 
the sample variation of 80.66% with mean response of 63.2917gms of rose coco bean 
production was attributed to the independent variables, and about 19.34% of the total variation 
couldn’t be explained by the model, implying that $R^2$ in GP3G of second order model explain 
80.66% of the variation in the model than $R^2=23.12\%$ of first order model in equation (4.50). 
The closer the value of R (correlation coefficient) to one, the better the correlation between the 
experimental and predicted values. Here, the value of R (0.8981) for equation (4.67) being close 
to 1 indicated a close agreement between the experimental results and the theoretical values 
predicted by the model equation.
Table 4.25 lists the regression coefficients and the corresponding p-values for the second-order polynomial model given as in eq. 4.67 as:

\[
\hat{Y}_{GP3G} = 51.7514 + 1.9127X_1 + 1.9876X_2 - 6.5028X_3 + 3.2471X_1^2 + 5.3785X_2^2 + 10.7726X_3^2 \\
- 10.8073X_1X_2 - 0.2039X_1X_3 - 4.6900X_2X_3
\] (4.67)

The \(\hat{Y}_{GP3G}\) is the predicted response for rose coco beans in group 3. The regression coefficients of the linear term for potassium (p=0.0021) have significant effects on the yield (p-value <0.05), the quadratic P^2 (p=0.0371) and K^2 (p=0.0004) have significant effects on the yield and the interaction terms in N*P (nitrogen and phosphorus, p=0.0012) is significant. Among these, K, K^2 (quadratic potassium), N*P was significant at the 1% significance level, while P^2 (quadratic phosphorus) was significant at the 5% level, those other terms of the model showed no significant effect on the yield. The positive coefficients of P^2, K^2 enhance the yield since they are the largest coefficients in the model equation (4.67). The largest negative coefficient of K and NP minimizes the yield of rose coco at their respective fertilizer input. The \(X_1\) represents Nitrogen (N), \(X_2\) represents Phosphorus (P), \(X_3\) represents Potassium (K).

This suggests that Potassium (K), quadratic effects of phosphorus (P^2), quadratic effects of potassium (K^2) and interaction N*P were the determining significant factors on the rose coco...
beans yield as these had the largest coefficients and those other terms of the model showed no significant effect on the yield.

4.4.7.5 The lack of Fit for the three replicates in GP3G

Lack-of-fit tests were performed on the fitted models. The lack-of-fit test compares the variation around the model with pure variation within replicated observations using the results of table 4.9.

Table 4.26: The lack of fit test for GP3G.

<table>
<thead>
<tr>
<th>Sum of</th>
<th>Residual</th>
<th>DF</th>
<th>Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of Fit</td>
<td>14</td>
<td>1786.444300</td>
<td>127.603164</td>
<td>0.25</td>
<td>0.9968</td>
<td></td>
</tr>
<tr>
<td>Pure Error</td>
<td>48</td>
<td>24694</td>
<td>514.458333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Error</td>
<td>62</td>
<td>26480</td>
<td>427.103940</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The lack of fit of the model was insignificant, so the second order model was adequate for the experiments of GP3G.

Table 4.27: The Analysis of Variance Table for N,P,K in 2nd order model- GP3G.

<table>
<thead>
<tr>
<th>Sum of</th>
<th>Factor</th>
<th>DF</th>
<th>Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>4</td>
<td>836.965298</td>
<td>209.241325</td>
<td>4.92</td>
<td>0.0109</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>1116.523540</td>
<td>279.130885</td>
<td>6.56</td>
<td>0.0034</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>4</td>
<td>1641.346675</td>
<td>410.336669</td>
<td>9.65</td>
<td>0.0006</td>
<td></td>
</tr>
</tbody>
</table>

We saw in the table 4.27 that all the three N, P, K are significant factors at 5% level of significance.
4.5 Finding the settings of the experimental factors that produces the optimal response

This section gives the three dimensional surfaces and contour plots, canonical analysis, stationary points and the ridge analysis.

4.5.1 Three dimensional surfaces and contour plots

This section illustrates the three dimensional surfaces and contour plots for the three groups.

4.5.1.1 Three dimensional surfaces and contour plots of GP1G

Graphical visualization of contour plots helps in understanding the second-order response surface. Specifically, three dimensional surface plots and their accompanying contour plots help characterize the shape of the surface and through this we are able to approximately locate the optimum response. Using the fit of the second-order model we illustrate quadratic response surfaces such as minimum, maximum, ridge, and saddle point in the case that an optimum exits, then this point is a stationary point.
Figure 4.7 a: The response surface plots for the treatments of nitrogen and phosphorus fertilizer concentrations in GP1G.

The N and P are significant, maximum yield of 71.81 grams was achieved with lower levels of N and higher P, also high N and high P still gave higher yield of around 60 grams.
The contour plot is a two-dimensional and the third design variable must be held constant to construct the graph. The contour plot for nitrogen and phosphorus when potassium is held constant was given as:

![Contour Plot](image)

**Figure 4.7a 1**: The contour plot for nitrogen and phosphorus fertilizers in GP1G of 4.7a.

In the figure 4.7a 1 the maximum yield of 60 gram per *rose coco* yield was obtained by reducing the nitrogen input from the center of 10 grams and increasing the phosphorus to the maximum scale from the centre point of 20 grams.
In figure 4.7b, high yield of 62.56 grams was noticed when we have low levels of nitrogen and high levels of potassium, as well as high yield with high levels of nitrogen and low levels of potassium.
The contour plot for nitrogen and potassium when phosphorus was held constant is given as:

Figure 4.7a 2: The contour plot of nitrogen and potassium fertilizers in GP1G of 4.7b.
In the figure 4.7a 2 the maximum yield of 70 grams was obtained by reducing the nitrogen fertilizer input to the lowest scale from the centre point 10 grams and increasing the potassium to the maximum scale from the centre point of 30 grams.
Figure 4.7 c: The response surface plots for the treatments of phosphorus and potassium fertilizer concentrations in GP1G.

In the figure 4.7c, we saw high yield of 55.54 grams under high levels of phosphorus from the center point of 20 grams and under low levels of potassium fertilizer from the center point 30 grams, we saw also high yield with high level of phosphorus with high level of potassium.
The contour plot for phosphorus and potassium when nitrogen was held constant is given as:

Figure 4.7a 3: The contour plot of phosphorus and potassium fertilizers in GP1G of 4.7c.

In the figure 4.7a 3 the maximum yield of 60 grams is obtained by increasing the phosphorus to the maximum scale from the center point 20 grams and reducing the potassium to the lowest scale from the centre point of 30 grams.
4.5.1.2 Three dimensional surfaces and contour plots of GP2G

Figure 4.8a: The response surface plots for the treatments of nitrogen and phosphorus fertilizer concentrations in GP2G.

In the figure 4.8a, maximum yield of 53.95 grams is achieved with lower levels of N and lower levels of P, also a yield of about 52 grams with high N and high P.
The contour plot is a two-dimensional and the third design variable must be held constant to construct the graph. The contour plot for nitrogen and phosphorus when potassium was held constant is given as:

Figure 4.8 b 1: The contour plot of nitrogen and phosphorus fertilizers in GP2G of 4.8a.
In the figure 4.8 b 1 the maximum yield of 56 grams per rose coco plant was obtained by lowering the nitrogen input from the center of 20 grams and reducing the phosphorus to the lower scale from the centre point of 30 grams, also a yield of 58 grams is obtained by getting nitrogen and phosphorus at maximum scale.
Figure 4.8 b: The response surface plots for the treatments of nitrogen and potassium fertilizer concentrations in GP2G.

In figure 4.8b, high yield of 60.88 grams was noticed when we have high levels of nitrogen from the center point of 20 grams and high level of potassium from the center point of 40 grams.
The contour plot for nitrogen and potassium when phosphorus was held constant is given as:

Figure 4.8 b 2: The contour plot of nitrogen and potassium fertilizers in GP2G of 4.8b.

In the figure 4.8 b 2 the maximum yield of 60 grams was obtained by increasing the nitrogen fertilizer input to the highest scale from the centre point 20 grams and increasing the potassium to the maximum scale from the centre point of 40 grams.
Figure 4.8 c: The response surface plots for the treatments of phosphorus and potassium fertilizer concentrations in GP2G.

In the figure 4.8c maximum yield of 64.72 grams of *rose coco* beans was achieved by high levels from the centre point of 30 grams of phosphorus and higher levels from the centre point of 40 grams of potassium fertilizers.
The contour plot for phosphorus and potassium when nitrogen was held constant is given as:

Figure 4.8 b 3: The contour plot of phosphorus and potassium fertilizers in GP2G of 4.8c.
In the figure 4.8 b 3, the maximum yield of 62 grams was obtained by increasing the phosphorus to the maximum scale from the center point 30 grams and increasing the potassium to the highest scale from the center point of 40 grams.
4.5.1.3 Three dimensional surfaces and contour plots of GP3G

Figure 4.9a: The response surface plots for the treatments of nitrogen and phosphorus fertilizer concentrations in GP3G.

In the figure 4.9a, maximum yield of 78 grams was achieved with lower levels of nitrogen and higher levels of phosphorus, also a yield of about 71 grams with high nitrogen and low phosphorus.
The contour plot is a two-dimensional and the third design variable must be held constant to construct the graph. The contour plot for nitrogen and phosphorus when potassium was held constant is given as:

Figure 4.9 c 1: The contour plot of nitrogen and phosphorus fertilizers in GP3G of 4.9a.
In the figure 4.9 c 1 the maximum yield of 80 gram per rose coco plant was obtained by lowering the nitrogen input from the center of 30 grams and increasing the phosphorus to the higher scale from the centre point of 40 grams, the same maximum yield of 80 grams was obtained by higher nitrogen and lower phosphorus.
Figure 4.9 b: The response surface plots for the treatments of nitrogen and potassium fertilizer concentrations in GP3G.

In figure 4.9b, high yield of 84.42 grams was obtained when we have high level of nitrogen from the center point of 30 grams and low level of potassium from the center point of 50 grams.
The contour plot for nitrogen and potassium when phosphorus was held constant is given as:

Figure 4.9 c 2: The contour plot of nitrogen and potassium fertilizers in GP3G of 4.9b.

In the figure 4.9 c 2, the maximum yield of 85 grams was obtained by increasing the nitrogen fertilizer input to the highest scale from the centre point 30 grams and reducing the potassium to the lowest scale from the centre point of 50 grams.
Figure 4.9 c: The response surface plots for the treatments of phosphorus and potassium fertilizer concentrations in GP3G.

In the figure 4.9c maximum yield of 88.83 grams of *rose coco* beans was achieved by high levels of phosphorus from the centre point of 40 grams and low level of potassium fertilizers from the centre point of 50 grams.
The contour plot for phosphorus and potassium when nitrogen was held constant is given as:

Figure 4.9 c 3: The contour plot of phosphorus and potassium fertilizers in GP3G 4.9c.

In the figure 4.9 c 3, the maximum yield of 95 grams was obtained by increasing the phosphorus to the maximum scale from the center point 40 grams and reducing the potassium to the lowest scale from the centre point of 50 grams.
4.5.2 The canonical analysis

This section gives the canonical analysis for the three groups.

4.5.2.1 The canonical analysis of GP1G

We determined the setting for optimal and efficient production of *rose coco* beans using canonical analysis. Canonical analysis is used to predict the shape of the curve generated by the model. The canonical analysis in table 4.28 and figure 4.7a, 4.7b and 4.7c of the GP1G response surface indicates that the predicted response surfaces are saddle points. The coded eigenvalues are $\lambda_1=6.770341$, $\lambda_2=0.483456$ and $\lambda_3=-0.907415$. The eigenvalue of 6.770341 shows that the valley orientation of the saddle was more curved, eigenvalue 0.4834566 shows that the valley orientation of the saddle was less curved, the less hilly orientation was with an eigenvalue of –0.967415. The coefficients of the associated eigenvectors show that the valley was more aligned with nitrogen, a less valley with potassium and the hill with phosphorus. Because the canonical analysis resulted in a saddle point, the estimated surface does not have a unique optimum. The surface was more sensitive to the changes of nitrogen compared to fertilizers of K. The results from these saddle point indicated that under natural values 10.978780gms of N, 20.991461gms of P and 33.358334gms of K fertilizer were needed to achieve the saddle point bean yield of 54.88gms per *rose coco* plant.

Table 4. 28: The Canonical Analysis of Response Surface, Eigenvectors and Eigenvalues-GP1G.

<table>
<thead>
<tr>
<th>Critical Factor</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1.957550</td>
</tr>
<tr>
<td>P</td>
<td>3.304884</td>
</tr>
<tr>
<td>K</td>
<td>3.358325</td>
</tr>
</tbody>
</table>

Predicted value at stationary point: 54.881682
Stationary point is a saddle point

4.5.2.2 The canonical analysis of GP2G

Table 4.29: The Canonical Analysis of Response Surface, Eigenvectors and Eigenvalues-GP2G.

<table>
<thead>
<tr>
<th>Critical Factor</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>0.769867</td>
</tr>
<tr>
<td>P</td>
<td>-0.283395</td>
</tr>
<tr>
<td>K</td>
<td>-0.264186</td>
</tr>
</tbody>
</table>

Predicted value at stationary point: 36.419471

The canonical analysis table 4.29 and figure 4.8a, 4.8b and 4.8c of response surfaces indicates that the predicted response surface was shaped a saddle point. The coded eigenvalues are $\lambda_1=9.570182$, $\lambda_2=7.278125$ and $\lambda_3=-0.747253$. The eigenvalue of 9.570182 shows that the valley orientation of the saddle was more curved, eigenvalue 7.278125 shows that the valley orientation of the saddle was less curved, the less hilly orientation was with an eigenvalue of −0.747253. The coefficients of the associated eigenvectors show that the valley was more aligned with potassium, a less valley with phosphorus and the hill with nitrogen. Because the canonical analysis resulted in a saddle point, the estimated surface does not have a unique optimum. The results from these indicated that under natural values 20.384936gms of N, 29.914981gms of P and 39.735813gms of K fertilizer were needed to achieve the saddle point of *rose coco* bean yield of 36.42gms per plant. The positive signs of the eigenvalues for P and K fertilizers...
indicated the directions of upward curvature. The largest eigenvalue (in absolute) for the N fertilizer, means that the N fertilizer was more pronounced and the curvature of the response surface was in the associated direction. So in GP1G and GP2G nitrogen shows mixed reaction of upward curvature and downward curvature respectively.

4.5.2.3 The canonical analysis of GP3G

Table 4.30: The Canonical Analysis of Response Surface, Eigenvectors and Eigenvalues-GP3G.

<table>
<thead>
<tr>
<th>Critical Factor</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>0.444188</td>
</tr>
<tr>
<td>P</td>
<td>0.436329</td>
</tr>
<tr>
<td>K</td>
<td>0.401005</td>
</tr>
</tbody>
</table>

Predicted value at stationary point: 51.306032

<table>
<thead>
<tr>
<th>Eigenvectors</th>
<th>N</th>
<th>P</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12.224838</td>
<td>-0.511793</td>
<td>0.805437</td>
</tr>
<tr>
<td></td>
<td>8.574872</td>
<td>0.567316</td>
<td>0.578120</td>
</tr>
<tr>
<td></td>
<td>-1.401508</td>
<td>0.645152</td>
<td>0.130574</td>
</tr>
</tbody>
</table>

Stationary point is a saddle point

The canonical analysis table 4.30 and figure 4.9a, 4.9b and 4.9c of response surface indicates that the predicted response surface was shaped a saddle point. The coded eigenvalues are \( \lambda_1=12.224838 \), \( \lambda_2=8.574872 \) and \( \lambda_3=-1.401508 \). The eigenvalue of 12.224838 shows that the valley orientation of the saddle was more curved, eigenvalue 8.574872 shows that the valley orientation of the saddle was less curved, the less hilly orientation was with an eigenvalue of 1.401508. The coefficients of the associated eigenvectors show that the valley in both of the first two was more aligned with potassium and the hilly with nitrogen. Because the canonical analysis resulted in a saddle point, the estimated surface does not have a unique optimum. The surface was more sensitive to the changes in amount of K & P, compared to fertilizers of N. The
results from these indicated that under natural values 30.222094gms of N, 40.130899gms of P and 50.401005gms of K fertilizer were needed to achieve the saddle point of rose coco bean yield of 51.31gms per plant.

4.5.3 The stationary points

This section gives the stationary points for the three groups.

4.5.3.1 The stationary point-GPIG

The stationary point can be found by using matrix algebra. The fitted second-order model in matrix form is:

\[ \hat{y} = \hat{\beta}_0 + x'b + x'Bx \]

The derivative of \( \hat{y} \) with respect to the elements of the vector \( x \) is

\[ \frac{\partial \hat{y}}{\partial x} = b + 2Bx \]

Therefore, setting the derivative vector to 0 yields the stationary point:

\[ x_s = -\frac{1}{2} B^{-1}b \]

This could be a maximum, minimum, or a saddle point of the fitted surface. The eigenvalues (call them \( \lambda_s \) here) and eigenvectors of \( B \) are the key to characterizing the shape. The \( x_s \) is a point of maximum if all \( \lambda \)'s are negative, the point of minimum if all \( \lambda \)'s are positive and saddle point if \( \lambda \)'s are of mixed sign.
Where \( \mathbf{b} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} \) and \( \mathbf{B} = \begin{bmatrix} \hat{\beta}_{11} & \frac{1}{2} \hat{\beta}_{12} & \frac{1}{2} \hat{\beta}_{13} \\ \frac{1}{2} \hat{\beta}_{12} & \hat{\beta}_{22} & \frac{1}{2} \hat{\beta}_{23} \\ \text{sym} & \frac{1}{2} \hat{\beta}_{23} & \hat{\beta}_{33} \end{bmatrix} \) (4.68)

\( \mathbf{b} \) is a (3x1) vector of the first-order regression coefficients and \( \mathbf{B} \) is a (3x3) symmetric matrix whose main diagonal elements are the quadratic coefficients and whose off-diagonal elements are one-half the mixed quadratic coefficients (Montgomery, 2005). The estimated response value at the stationary point is

\[
\hat{y}_s = \hat{\beta}_0 + \frac{1}{2} \mathbf{x}' \mathbf{b}
\]  

(4.69)

Therefore, locating stationary points for the yield-GP1G

\[
\mathbf{b} = \begin{bmatrix} -5.0452 \\ 6.3461 \\ 1.1566 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3.9485 & 1.32545 & -2.85475 \\ 1.32545 & -0.3983 & -1.32545 \\ -2.85475 & -1.32545 & 2.7962 \end{bmatrix}, \quad \mathbf{B}^{-1} = \begin{bmatrix} 0.9663 & -0.0261 & 0.9742 \\ -0.0313 & -0.9733 & -0.4932 \\ 0.9717 & -0.4880 & 1.1184 \end{bmatrix}
\]  

(4.70)

The stationary point using the equation

\[
\mathbf{x}_s = -\frac{1}{2} \mathbf{B}^{-1} \mathbf{b}
\]  

is

\[
\mathbf{x}_s = \begin{bmatrix} 1.9571 \\ 3.2946 \\ 3.3530 \end{bmatrix}
\]  

(4.71)

We can find the natural values of nitrogen, phosphorus and potassium using the stationary points from table 4.28.

\[
1.9571 = \frac{\psi_{1u} - 10}{0.5}, \quad \psi_{1u} = 10.97855, \quad 3.2946 = \frac{\psi_{2u} - 20}{0.3}, \quad \psi_{2u} = 20.98838
\]  

(4.72)
Using the equation \( \hat{y} = \hat{\beta}_0 + x'\hat{\beta} + x'\hat{B}x \), we found out that estimated maximum response yield of *rose coco* bean at the stationary point was

\( \hat{Y} = 54.88 \) grams

Thus, it could be concluded that this level of main factors setting resulted in saddle optimum yield for the *rose coco* bean for the given amount of predictor variables. The maximum yield of *rose coco* bean is obtained when under this combination of fertilizers in GP1G

\[
-0.850004 = \frac{\psi_{1u} - 10}{0.5}, \quad \psi_{1u} = 9.574998, \quad 0.157611 = \frac{\psi_{2u} - 20}{0.3}, \quad \psi_{2u} = 20.0472833 \quad (4.73)
\]

\[
0.502645 = \frac{\psi_{3u} - 30}{1}, \quad \psi_{3u} = 30.502645
\]

Therefore, we can obtain a maximum yield of 58.68gms per *rose coco* bean with a combination of 9.57gms nitrogen, 20.05gms phosphorus and 30.50gms potassium fertilizers.

### 4.5.3.2 The stationary point-GP2G

Locating stationary points for the yield-GP2G was given as;

\[
b = \begin{bmatrix} 1.4399 \\ -0.6980 \\ 2.7503 \end{bmatrix}, \quad B = \begin{bmatrix} 0.6806 & 3.0587 & 1.4274 \\ 3.0587 & 6.6025 & 0.5098 \\ 1.4274 & 0.5098 & 8.8179 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} -1.1135 & 0.5042 & 0.1511 \\ 0.5042 & -0.0762 & -0.0772 \\ 0.1511 & -0.0772 & 0.0934 \end{bmatrix} (4.74)
\]

The stationary point using the equation

\[
x_s = -\frac{1}{2}B^{-1}b \quad \text{is}
\]

\[
x_s = \begin{bmatrix} 0.7699 \\ -0.2834 \\ -0.2642 \end{bmatrix} \quad (4.75)
\]
This was a saddle point since it showed mixed signs of $\lambda$’s.

We could find the natural values of nitrogen, phosphorus and potassium fertilizers using the stationary point from table 4.29.

\[
0.7699 = \frac{\psi_{2u} - 20}{0.5}, \quad \psi_{1u} = 20.38495, \quad -0.2834 = \frac{\psi_{2u} - 30}{0.3}, \quad \psi_{2u} = 29.91498
\]

\[-0.2642 = \frac{\psi_{3u} - 40}{1}, \quad \psi_{3u} = 39.7358 \tag{4.76}\]

Using the equation $\hat{y} = \hat{\beta}_0 + x'\hat{b} + x'Bx$, we can find that estimated maximum response yield of rose coco bean at the stationary point was

$\hat{Y} = 36.4gm$

Thus, it could be concluded that this level of main factors setting results in best optimum yield for the rose coco beans. The maximum yield of rose coco beans was obtained when under this combination of fertilizers in GP2G

\[
0.270935 = \frac{\psi_{1u} - 20}{0.5}, \quad \psi_{1u} = 20.1354675, \quad 0.226583 = \frac{\psi_{2u} - 30}{0.3}, \quad \psi_{2u} = 30.0679749
\]

\[
0.935550 = \frac{\psi_{3u} - 40}{1}, \quad \psi_{3u} = 40.935550 \tag{4.77}\]

Therefore, we can obtain a maximum yield of 48.36gms per rose coco beans with a combination 20.14gms nitrogen, 30.07gms phosphorus and 40.94gms potassium fertilizers.

4.5.3.3 The stationary point-GP3G

Locating stationary points for the yield-GP3G

\[
b = \begin{bmatrix} 1.9127 \\ 1.9876 \\ -6.5028 \end{bmatrix}, \quad B = \begin{bmatrix} 3.2471 & -5.40365 & -0.10195 \\ -5.40365 & 5.3785 & -2.3450 \\ -.10195 & -2.3450 & 10.7726 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} -0.3570 & -0.3979 & -0.0900 \\ -0.3979 & -0.2380 & -0.0556 \\ -0.0900 & -0.0556 & 0.0799 \end{bmatrix}
\]
The stationary point using the equation

\[ x_s = -\frac{1}{2} B^{-1}b \]

is

\[ x_s = \begin{bmatrix} 0.4442 \\ 0.4363 \\ 0.4010 \end{bmatrix} \]  

(4.79)

We could find the natural values of nitrogen, phosphorus and potassium using the stationary points from table 4.30.

\[ 0.4442 = \frac{\psi_{2u} - 30}{0.5}, \quad \psi_{1u} = 30.2221, \quad 0.4363 = \frac{\psi_{2u} - 40}{0.3}, \quad \psi_{2u} = 40.13089 \]  

(4.80)

\[ 0.4010 = \frac{\psi_{3u} - 50}{1}, \quad \psi_{3u} = 50.4010 \]

Using the equation \( \hat{\psi} = \hat{\beta}_0 + x'\beta + x'Bx \), we can find that estimated maximum response yield of *rose coco* beans at the stationary point was

\[ \hat{\psi} = 51.31 \text{ grams} \]

Thus, it can be concluded that this level of main factors setting results in best optimum yield for the *rose coco* bean.

The maximum yield of *rose coco* bean was obtained when under this combination of fertilizers in GP3G

\[ -0.077342 = \frac{\psi_{1u} - 30}{0.5}, \quad \psi_{1u} = 29.961329, \quad 0.365469 = \frac{\psi_{2u} - 40}{0.3}, \quad \psi_{2u} = 40.1096407 \]

\[ -0.927605 = \frac{\psi_{3u} - 50}{1}, \quad \psi_{3u} = 49.072395 \]  

(4.81)
Therefore, we obtained a maximum yield of 70.25gms per *rose coco* bean with a combination of 29.96gms nitrogen, 40.11gms phosphorus and 49.07gms potassium fertilizers.

### 4.5.4 The ridge analysis

A RIDGE statement computes the ridge of optimum response. The ridge starts at a given point $x_0$, and the point on the ridge at radius $R$ from $x_0$ is the collection of factor settings that optimizes the predicted response at this radius. The ridge analysis can be used as a tool to help interpret an existing response surface or to indicate the direction in which further experimentation should be performed. The default starting point, $x_0$, has each coordinate equal to the point midway between the highest and lowest values of the factor in the design. The default radii at which the ridge is computed are 0, 0.1, ..., 0.9, 1. If the ridge analysis is based on the response surface fit to coded values for the factor variables, then this result in a ridge that starts at the point with a coded zero value for each coordinate and extends toward, but not beyond, the edge of the range of experimentation. Alternatively, both the center point of the ridge and the radii at which it is to be computed can be specified. The starting point should be well inside the range of experimentation. The coded radii give the distances from the ridge starting point at which to compute the optimal (SAS User Guide 2013). Ridge analysis was performed to determine the critical levels of the design variables that produce the maximum response.

#### 4.5.4.1 The ridge analysis –GP1G

The ridge analysis was performed to determine the critical levels of the design variables that produce the maximum and minimum response.
Table 4.31: The Estimated Ridge of Minimum Response for Variable YIELD-GP1G.

<table>
<thead>
<tr>
<th>Radius</th>
<th>Estimated Response</th>
<th>Standard Error</th>
<th>N</th>
<th>P</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>47.391125</td>
<td>3.606523</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
<td>46.573422</td>
<td>3.594075</td>
<td>0.057896</td>
<td>-0.080662</td>
<td>-0.011904</td>
</tr>
<tr>
<td>0.2</td>
<td>45.753978</td>
<td>3.557499</td>
<td>0.109534</td>
<td>-0.166041</td>
<td>-0.020805</td>
</tr>
<tr>
<td>0.3</td>
<td>44.928378</td>
<td>3.499176</td>
<td>0.156344</td>
<td>-0.254479</td>
<td>-0.028228</td>
</tr>
<tr>
<td>0.4</td>
<td>44.093828</td>
<td>3.423340</td>
<td>0.199358</td>
<td>-0.345000</td>
<td>-0.035092</td>
</tr>
<tr>
<td>0.5</td>
<td>43.248466</td>
<td>3.336814</td>
<td>0.239315</td>
<td>-0.437000</td>
<td>-0.041945</td>
</tr>
<tr>
<td>0.6</td>
<td>42.390993</td>
<td>3.247865</td>
<td>0.276746</td>
<td>-0.530094</td>
<td>-0.049111</td>
</tr>
<tr>
<td>0.7</td>
<td>41.520461</td>
<td>3.170067</td>
<td>0.312049</td>
<td>-0.624021</td>
<td>-0.056779</td>
</tr>
<tr>
<td>0.8</td>
<td>40.636149</td>
<td>3.119186</td>
<td>0.345525</td>
<td>-0.718596</td>
<td>-0.065057</td>
</tr>
<tr>
<td>0.9</td>
<td>39.737496</td>
<td>3.114224</td>
<td>0.377413</td>
<td>-0.813685</td>
<td>-0.073998</td>
</tr>
<tr>
<td>1.0</td>
<td>38.824051</td>
<td>3.175047</td>
<td>0.407901</td>
<td>-0.909188</td>
<td>-0.083625</td>
</tr>
</tbody>
</table>

Table 4.32: The Estimated Ridge of Maximum Response for Variable YIELD-GP1G.

<table>
<thead>
<tr>
<th>Radius</th>
<th>Estimated Response</th>
<th>Standard Error</th>
<th>N</th>
<th>P</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>47.391125</td>
<td>3.606523</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
<td>48.214348</td>
<td>3.594075</td>
<td>40.065957</td>
<td>0.073143</td>
<td>0.017314</td>
</tr>
<tr>
<td>0.2</td>
<td>49.055109</td>
<td>3.557499</td>
<td>40.160534</td>
<td>0.134455</td>
<td>0.043284</td>
</tr>
<tr>
<td>0.3</td>
<td>49.931933</td>
<td>3.499176</td>
<td>40.258543</td>
<td>0.179101</td>
<td>0.080692</td>
</tr>
<tr>
<td>0.4</td>
<td>50.868599</td>
<td>3.423340</td>
<td>40.354944</td>
<td>0.205031</td>
<td>0.129395</td>
</tr>
<tr>
<td>0.5</td>
<td>51.886533</td>
<td>3.348681</td>
<td>40.450083</td>
<td>0.214920</td>
<td>0.186128</td>
</tr>
<tr>
<td>0.6</td>
<td>53.009694</td>
<td>3.247865</td>
<td>40.544396</td>
<td>0.213572</td>
<td>0.247237</td>
</tr>
<tr>
<td>0.7</td>
<td>54.243860</td>
<td>3.170067</td>
<td>40.638159</td>
<td>0.205018</td>
<td>0.310362</td>
</tr>
<tr>
<td>0.8</td>
<td>55.598337</td>
<td>3.114224</td>
<td>40.731544</td>
<td>0.191884</td>
<td>0.374295</td>
</tr>
<tr>
<td>0.9</td>
<td>57.077682</td>
<td>3.076061</td>
<td>40.824658</td>
<td>0.175760</td>
<td>0.438473</td>
</tr>
<tr>
<td>1.0</td>
<td>58.684816</td>
<td>3.175047</td>
<td>40.917572</td>
<td>0.157611</td>
<td>0.502645</td>
</tr>
</tbody>
</table>

4.5.4.2 The ridge analysis-GP2G

The ridge analysis was performed to determine the critical levels of the design variables that produce the maximum and minimum response.

Table 4.33: The Estimated Ridge of Minimum Response for Variable YIELD-GP2G.

<table>
<thead>
<tr>
<th>Radius</th>
<th>Estimated Response</th>
<th>Standard Error</th>
<th>N</th>
<th>P</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>36.129616</td>
<td>3.459567</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
<td>35.873343</td>
<td>3.447716</td>
<td>-0.065440</td>
<td>0.034219</td>
<td>-0.067429</td>
</tr>
<tr>
<td>0.2</td>
<td>35.682881</td>
<td>3.412629</td>
<td>-0.160534</td>
<td>0.079533</td>
<td>-0.088901</td>
</tr>
<tr>
<td>0.3</td>
<td>35.501614</td>
<td>3.356681</td>
<td>-0.258543</td>
<td>0.122718</td>
<td>-0.089978</td>
</tr>
<tr>
<td>0.4</td>
<td>35.313422</td>
<td>3.323933</td>
<td>-0.354944</td>
<td>0.163818</td>
<td>-0.084724</td>
</tr>
<tr>
<td>0.5</td>
<td>35.113822</td>
<td>3.200611</td>
<td>-0.450083</td>
<td>0.203742</td>
<td>-0.076907</td>
</tr>
<tr>
<td>0.6</td>
<td>34.901149</td>
<td>3.115604</td>
<td>-0.544396</td>
<td>0.242973</td>
<td>-0.067807</td>
</tr>
<tr>
<td>0.7</td>
<td>34.674645</td>
<td>3.040974</td>
<td>-0.638159</td>
<td>0.281766</td>
<td>-0.057969</td>
</tr>
<tr>
<td>0.8</td>
<td>34.433913</td>
<td>2.992165</td>
<td>-0.731544</td>
<td>0.320266</td>
<td>-0.047669</td>
</tr>
<tr>
<td>0.9</td>
<td>34.178724</td>
<td>2.987405</td>
<td>-0.824658</td>
<td>0.358560</td>
<td>-0.037059</td>
</tr>
<tr>
<td>1.0</td>
<td>33.908940</td>
<td>3.045751</td>
<td>-0.917572</td>
<td>0.396704</td>
<td>-0.026231</td>
</tr>
</tbody>
</table>

Table 4.34: The Estimated Ridge of Maximum Response for Variable YIELD-GP2G.
The ridge analysis was performed to determine the critical levels of the design variables that produce the maximum and minimum response.

Table 4.35: The Estimated Ridge of Minimum Response for Variable YIELD-GP3G.

<table>
<thead>
<tr>
<th>Radius</th>
<th>Estimated Response</th>
<th>Standard Error</th>
<th>N</th>
<th>Factor Values</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>51.751435</td>
<td>3.410664</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
<td>51.142673</td>
<td>3.398892</td>
<td>0.038479</td>
<td>-0.031963</td>
<td>0.086589</td>
</tr>
<tr>
<td>0.2</td>
<td>50.696721</td>
<td>3.364302</td>
<td>-0.106285</td>
<td>-0.087059</td>
<td>0.148935</td>
</tr>
<tr>
<td>0.3</td>
<td>50.357171</td>
<td>3.309246</td>
<td>-0.139312</td>
<td>-0.147243</td>
<td>0.175264</td>
</tr>
<tr>
<td>0.4</td>
<td>49.998655</td>
<td>3.237429</td>
<td>-0.282585</td>
<td>-0.218021</td>
<td>0.180589</td>
</tr>
<tr>
<td>0.5</td>
<td>49.653355</td>
<td>3.155286</td>
<td>-0.368113</td>
<td>-0.288100</td>
<td>0.177458</td>
</tr>
<tr>
<td>0.6</td>
<td>49.290454</td>
<td>3.071483</td>
<td>-0.451052</td>
<td>-0.357014</td>
<td>0.170567</td>
</tr>
<tr>
<td>0.7</td>
<td>48.905533</td>
<td>2.997910</td>
<td>-0.532192</td>
<td>-0.424991</td>
<td>0.161723</td>
</tr>
<tr>
<td>0.8</td>
<td>48.496347</td>
<td>2.949793</td>
<td>-0.612079</td>
<td>-0.492276</td>
<td>0.151737</td>
</tr>
<tr>
<td>0.9</td>
<td>48.061651</td>
<td>2.945100</td>
<td>-0.691071</td>
<td>-0.559047</td>
<td>0.141024</td>
</tr>
<tr>
<td>1.0</td>
<td>47.600701</td>
<td>3.002620</td>
<td>-0.769404</td>
<td>-0.625432</td>
<td>0.129819</td>
</tr>
</tbody>
</table>

Table 4.36: The Estimated Ridge of Maximum Response for Variable YIELD-GP3G.

<table>
<thead>
<tr>
<th>Radius</th>
<th>Estimated Response</th>
<th>Standard Error</th>
<th>N</th>
<th>Factor Values</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>51.751435</td>
<td>3.410664</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.1</td>
<td>52.563118</td>
<td>3.398892</td>
<td>0.019147</td>
<td>0.027378</td>
<td>-0.094254</td>
</tr>
<tr>
<td>0.2</td>
<td>53.593652</td>
<td>3.364302</td>
<td>0.026607</td>
<td>0.055878</td>
<td>-0.190184</td>
</tr>
<tr>
<td>0.3</td>
<td>54.850777</td>
<td>3.309146</td>
<td>0.026256</td>
<td>0.087039</td>
<td>-0.285893</td>
</tr>
<tr>
<td>0.4</td>
<td>56.339123</td>
<td>3.237429</td>
<td>0.020297</td>
<td>0.121001</td>
<td>-0.380719</td>
</tr>
<tr>
<td>0.5</td>
<td>58.061911</td>
<td>3.155286</td>
<td>0.010145</td>
<td>0.157476</td>
<td>-0.474445</td>
</tr>
<tr>
<td>0.6</td>
<td>60.020980</td>
<td>3.071483</td>
<td>-0.003227</td>
<td>0.196868</td>
<td>-0.567045</td>
</tr>
<tr>
<td>0.7</td>
<td>62.218302</td>
<td>2.997910</td>
<td>-0.019119</td>
<td>0.234737</td>
<td>-0.658571</td>
</tr>
<tr>
<td>0.8</td>
<td>64.654972</td>
<td>2.949793</td>
<td>-0.037012</td>
<td>0.278327</td>
<td>-0.749109</td>
</tr>
<tr>
<td>0.9</td>
<td>67.331915</td>
<td>2.945100</td>
<td>-0.056519</td>
<td>0.321394</td>
<td>-0.838756</td>
</tr>
<tr>
<td>1.0</td>
<td>70.249853</td>
<td>3.002620</td>
<td>-0.077342</td>
<td>0.365469</td>
<td>-0.927605</td>
</tr>
</tbody>
</table>

In general the ridge analysis estimated maximum response for the variable yield was 58.68gms, 48.36gms and 70.25gms for GP1G, GP2G and GP3G respectively at nitrogen of 9.57gms, phosphorus of 20.05gms and potassium of 30.50gms for GP1G, nitrogen of 20.14gms,
phosphorus of 30.07gms and potassium of 40.94gms for GP2G and nitrogen of 29.96gms, phosphorus of 40.11gms and potassium of 49.07gms for GP3G.
CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter presents conclusions and recommendations in line with the four specific objectives.

5.2 Estimation of the linear parameters in an existing A-optimum and D-efficient calculus optimum value second order rotatable design

The study was able to estimate the parameter coefficient successfully for the three groups.

Below is the table 5.1 showing the parameter coefficients obtained and their p-values.

Table 5.1: Regression coefficients of predicted second order polynomial models for the response of GP1G, GP2G and GP3G.

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>GP1G (p-value)</th>
<th>GP2G (p-value)</th>
<th>GP3G (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$ intercept</td>
<td>47.391125 (0.0001)</td>
<td>36.129616 (0.0001)</td>
<td>51.751435 (0.0001)</td>
</tr>
<tr>
<td>Linear</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$ (N)</td>
<td>-5.045204 (0.0152)</td>
<td>1.439859 (0.4246)</td>
<td>1.912726 (0.2865)</td>
</tr>
<tr>
<td>$\beta_2$ (P)</td>
<td>6.346081 (0.0035)</td>
<td>-0.697954 (0.6962)</td>
<td>1.987578 (0.2688)</td>
</tr>
<tr>
<td>$\beta_3$ (K)</td>
<td>1.156615 (0.5365)</td>
<td>2.750284 (0.1385)</td>
<td>-6.502796 (0.0021)</td>
</tr>
<tr>
<td>Quadratic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$ (N$^2$)</td>
<td>3.948538</td>
<td>0.680614</td>
<td>3.247054</td>
</tr>
<tr>
<td></td>
<td>(0.1321)</td>
<td>(0.7781)</td>
<td>(0.1861)</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>-0.398323</td>
<td>6.02530</td>
<td>5.378536</td>
</tr>
<tr>
<td></td>
<td>(0.8742)</td>
<td>(0.0145)</td>
<td>(0.0371)</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>2.796168</td>
<td>8.17910</td>
<td>10.772613</td>
</tr>
<tr>
<td></td>
<td>(0.2765)</td>
<td>(0.0023)</td>
<td>(0.0004)</td>
</tr>
</tbody>
</table>

Cross product

<table>
<thead>
<tr>
<th></th>
<th>(0.3619)</th>
<th>(0.0397)</th>
<th>(0.0012)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{12}$</td>
<td>2.650859</td>
<td>6.117366</td>
<td>-10.807347</td>
</tr>
<tr>
<td></td>
<td>(0.3619)</td>
<td>(0.0397)</td>
<td>(0.0012)</td>
</tr>
<tr>
<td>$\beta_{13}$</td>
<td>-5.709542</td>
<td>2.854771</td>
<td>-0.203912</td>
</tr>
<tr>
<td></td>
<td>(0.0618)</td>
<td>(0.3079)</td>
<td>(0.9400)</td>
</tr>
<tr>
<td>$\beta_{23}$</td>
<td>-2.650859</td>
<td>1.019561</td>
<td>-4.689981</td>
</tr>
<tr>
<td></td>
<td>(0.3619)</td>
<td>(0.7112)</td>
<td>(0.0997)</td>
</tr>
</tbody>
</table>

| R²               | 68.10%   | 67.04%   | 80.66%   |
| Adjusted R²      | 47.59%   | 45.85%   | 68.23%   |
| Sum of squared   | 666      | 613      | 595      |
| residuals        |          |          |          |
| Coefficient of   | 13.4782  | 14.4733  | 10.3044  |
| variation        |          |          |          |
| PRESS            | 2488     | 2091     | 2645     |
| p -the probability | 0.0219  | 0.0264  | 0.0011  |

N=nitrogen fertilizer; P=phosphorus fertilizer; K=potassium fertilizer

We saw in the three groups, that is, GP1G, GP2G and GP3G there was a large difference between ordinary residual and PRESS residual meaning the second order models fitted the data well.
5.3 Obtaining the generalized variance of the estimated linear parameters

In the table 5.2 the variance under first order models for the three groups (GP1G, GP2G, GP3G) are increasing, but for the second order models they are decreasing implying that second order models fitted the data appropriately.

Table 5.2: The eigenvalues, predicted *rose coco* yield at the stationary points, and uncoded critical values of fertilizer levels.

<table>
<thead>
<tr>
<th>Fertilizer</th>
<th>Eigenvalue (λ)</th>
<th>Natural critical value</th>
<th>Variance, $\sigma^2$.First order model</th>
<th>Variance, $\sigma^2$.Second order model</th>
<th>Predicted rose coco yield at stationary point</th>
<th>Estimated Ridge of Maximum Response for rose coco YIELD</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP1G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>6.770341</td>
<td>10.978780</td>
<td></td>
<td></td>
<td>54.88gms</td>
<td>58.78gms</td>
</tr>
<tr>
<td>P</td>
<td>0.483456</td>
<td>20.991461</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>-0.907415</td>
<td>33.358334</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP2G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>9.570182</td>
<td>20.384936</td>
<td></td>
<td></td>
<td>36.42gms</td>
<td>48.36gms</td>
</tr>
<tr>
<td>P</td>
<td>7.278125</td>
<td>29.914981</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>-0.747253</td>
<td>39.735813</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP3G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>12.224838</td>
<td>30.222094</td>
<td></td>
<td></td>
<td>51.31gms</td>
<td>70.25gms</td>
</tr>
<tr>
<td>P</td>
<td>8.574872</td>
<td>40.130899</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>-1.401508</td>
<td>50.401005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5.4 Fitting and testing the three model adequacies

The analysis of variance (ANOVA) of response surface for *rose coco* yield showed that the twenty four second order rotatable designs was adequate due to satisfactory levels of coefficient of determinations, $R^2$ (0.68, 0.67, 0.81) for the three cases GP1G, GP2G and GP3G respectively and coefficient of variations (CV, 13.48, 14.47 and 10.30 respectively). Generally, high values of CV indicate that experimental design developed was inadequate. In addition, linear, quadratic and cross product terms were all found to be significant at 1% for the GP3G. On the other hand, quadratic was highly significant for GP2G whereas GP1G was significant only on the linear terms. However, for the three groups total model was significant at 5% ($p = 0.0219, 0.0264$ and $0.0011$, respectively). The results showed GP2G and GP3G fitted the second order models well for the *rose coco* yield using the three fertilizer treatments, nitrogen (N), phosphorus (P) and potassium (K).

Table 5.3: The analysis of variance of the factors and the critical values of the *rose coco* yield response.

<table>
<thead>
<tr>
<th>Independent variables$^a$</th>
<th>Analysis of variance</th>
<th>Critical values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>d.f</td>
<td>Sum of squares</td>
</tr>
<tr>
<td>GP1G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>4</td>
<td>723.277182</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>660.747662</td>
</tr>
<tr>
<td>K</td>
<td>4</td>
<td>318.328520</td>
</tr>
<tr>
<td>GP2G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>4</td>
<td>307.213714</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>578.197324</td>
</tr>
<tr>
<td>K</td>
<td>4</td>
<td>769.676867</td>
</tr>
</tbody>
</table>
In GP3G the coefficient of determination, $R^2$ for the fitted model was 80.03%, which means that 80.03% of the total variability of the system was explained by the chosen factors N, P, K meaning investigation further need to be done on other pertinent factor components of fertilizers required by *the rose coco plant* for the coefficient of determination to increase to say 99.9%.

### 5.5 Finding the settings of the experimental factors that produce the optimal response

Graphics and visualization techniques are some of the best tools for understanding response surfaces, for which this research work utilized in the expounding of the nature and shape of the response surface generated from the fitted first order and second order models for the three groups. We visualize mountain and valleys of the second order response surface of *rose coco* beans for GP1G, GP2G and GP3G each at two combinations of fertilizers N, P, K. The curvature showed in GP2G and GP3G validated the second order model because it fitted the data well. Contour plots showing the contours of the surface, that is, curves of N, P, K pairs that have the same response value were generated which depicts the pattern and nature of the combination of the predictive factors, nitrogen, phosphorus and potassium fertilizer concentration of the *rose coco* beans yield.

The eigenvalues were used to determine whether the solution gave a maximum, minimum or saddle point on the response curve. The 3-dimensional response surface plots were a good way to visualize the fertilizer interaction. The canonical analysis of response surface for all groups,
that is, GP1G, GP2G and GP3G indicated that the stationary points were saddle points, implying that in the experimental region there were no maximum or minimum points. The study uses the ridge analysis as an alternative solution to overcome the saddle point problem. Ridge analyses were performed (table 5.1) to determine the critical levels of the design variables that could produce a maximum response.

Table 5. 4: Canonical analysis based on coded and actual values, eigenvalues and eigenvectors for the three groups.

<table>
<thead>
<tr>
<th></th>
<th>Canonical Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coded values</td>
</tr>
<tr>
<td></td>
<td>Eigenvalues</td>
</tr>
<tr>
<td>GP1G</td>
<td>Coded values</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uncoded/natural values</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>GP2G</td>
<td>Coded values</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uncoded/natural values</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>GP3G</td>
<td>Coded values</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uncoded/natural values</td>
</tr>
</tbody>
</table>
Table 5.4 displays the eigenvalues and eigenvectors. The canonical analysis indicates that the directions of principal orientation for the predicted response surface are along the axes associated with the three factors. In GP1G uncoded values, the largest eigenvalue (21.655306) corresponds to the eigenvector (0.869293, 0.353108, -0.345898), the largest component of which (0.869293) is associated with N; similarly, the second-largest eigenvalue in absolute (-8.231376) is associated with P. The third eigenvalue (0.740548), associated with K. The coded form of the canonical analysis indicates that the estimated response surface is at a saddle point, in uncoded terms, the model predicts that the yield of rose coco saddles when N=10.978780 grams, P=20.991461 grams, and K=33.358334 grams of fertilizers. In GP2G uncoded values, the largest eigenvalue (78.906109) corresponds to the eigenvector (0.259596, 0.965120, 0.033973), the largest component of which (0.965120) is associated with P; similarly, the second-largest eigenvalue (9.186905) is associated with K. The third eigenvalue (-3.191124), associated with N. The coded form of the canonical analysis indicates that the estimated response surface is at a saddle point, in uncoded terms, the model predicts that the yield of rose coco saddles when N=20.384936 grams, P=29.914981 grams, and K=39.735813 grams. In GP3G uncoded values, the largest eigenvalue (79.989081) corresponds to the eigenvector (-0.471069, 0.876678, -0.097616), the largest component of which (0.876678) is associated with P; similarly, the second-largest eigenvalue (10.972726) is associated with K. The third eigenvalue (-7.439426), associated with N. The coded form of the canonical analysis indicates that the estimated response surface is at a saddle point, in uncoded terms, the model predicts that the yield of rose coco saddles when N=30.222094 grams, P=40.130899 grams, and
K=50.401005 grams. In this canonical analysis, we saw that in GP1G the curvature is upward, was much more sensitive to nitrogen at 10gms for both coded and uncoded, in GP2G the upward curvature, that is, valley orientation was inclined to potassium at 40gms for coded but inclined to phosphorus for uncoded and in the GP3G was a valley orientation inclines towards potassium but inclined to phosphorus for uncoded. On the hilly or downward curvature is in GP1G, GP2G and GP3G are inclined to phosphorus at 20grams, nitrogen at 20grams and nitrogen at 30grams respectively. The optimal values all the three variables (nitrogen, phosphorus and potassium) for optimal production of rose coco beans predicted in GP1G (54.88 grams) by the model were: nitrogen fertilizers at 10.98 grams; phosphorus fertilizer at 20.99 grams; potassium fertilizer at 33.36 grams. The production of 36.4 grams was achieved in GP2G with 20.38 grams nitrogen fertilizer, 29.91 gram phosphorus fertilizer and 39.74 grams potassium fertilizer. In GP3G the production of 51.3 grams was achieved with 30.22 grams nitrogen, 40.13 grams phosphorus and 50.40 grams potassium fertilizers, this means that we have to be cautious with the specific amount of each fertilizer applied to rose coco plant because it might be affecting the effectiveness of the other fertilizers variable component as shown by the mean yield for the three groups after steadily increasing the fertilizers linearly.

The stationary points of all the three groups were the saddle points, indicating that optimum conditions for GP1G, GP2G and GP3G did not exist in the experimental range. It was found in the current study that three factors, nitrogen concentration, phosphorus concentration and potassium concentration, were important fertilizers for the rose coco beans yield. Since analysis of the surface response revealed that the stationary point for all the three groups was saddled points. Hence, the study used ridge analysis to offer an alternative solution for the saddle problem. Under the three combinations of the fertilizers, nitrogen 9.57 grams, phosphorus
20.05 grams and potassium 30.50 grams, 58.68 grams of *rose coco* yield was obtained in GP1G, 
nitrogen 20.14 grams, phosphorus 30.07 grams and potassium 40.94 grams, 48.36 grams of *rose coco* yield was obtained in GP2G and nitrogen 29.96 grams, phosphorus 40.11 grams and potassium is 49.07 grams, 70.25 grams of *rose coco* yield was obtained in GP3G. The results from this study show that a twenty four-points second order rotatable design was one of the suitable methods to optimize the best operating conditions in multi-factor operating environment for the purpose of obtaining maximum *rose coco* beans yield. The study has demonstrated the applicability of the twenty four points, second order rotatable design for the optimal and efficient production of *rose coco* beans. Among the three tests that are GP1G, GP2G and GP3G, we saw that GP3G is more desirable for further researches and acted as a starting point for farmers to plant *rose coco* beans so as to achieved the optimal/maximum potential yield. We therefore recommend the use of GP3G since it gave the above board the required coefficient of determination, $R^2$ of 80.66%, and the maximum yield of 70.25 grams was achieved. With the use of N, P and K fertilizers in this research study design, if adopted the farmers have a high potentiality to increase their yields or even triple the *rose coco* yield in the study area. Furthermore, this research can be extended to the area where beans are grown and a preliminary research of this kind should be conducted to ascertain the initial starting point for the inorganic fertilizers. The ministry of agriculture in Kenya should focus more researches of this kind, such that production using inorganic fertilizers (N, P, K) concentration matches the crop and the soil requirements of the *rose coco* beans in the growing areas for maximum yield to be realized. The research needs to be tried in the open field with the found optimal fertilizer combinations to see its performance. The yield of *rose coco* beans decreases as we increase the phosphorus fertilizer, more investigations need to be done to ascertain the amount of nitrogen
and potassium fertilizer needed for the effectiveness of the phosphorus fertilizer. The investigation also needs to be done if the certain amount of each component of the fertilizer affects the effectiveness of each other on the yield of *rose coco* beans.

5.6 Recommendation for further research

The randomization of the initial fertilizer application is advisable because in this research; a linear increment of the fertilizers was used in the order of N, P then K, therefore, thorough repetitive screening of all the fertilizer components needs to be done to ascertain the right initial amount of fertilizers that could achieve maximum yield. The optimal output of *rose coco* beans needs to be obtained when the free/letter parameter (f) assumes a unit value, then a comparison can be done with the current study of calculus optimum value and determine which one gives the best estimates among the two.
REFERENCES


Montgomery D. C. (2005). *Design and Analysis of Experiments response surface methods and design*. John Willey and Sons Inc. New Jersey,


APPENDIX 1

1.1 The codes for generating the response surfaces using SAS software

GP1G

DATA RP;
INPUT N P K YIELD;
title 'Surface plot of rose coco yield vs Nitrogen, Phosphorus';
CARDS;
1.1072569 1.1072569 0 63
-1.1072569 1.1072569 0 71
1.1072569 -1.1072569 0 35
-1.1072569 -1.1072569 0 56
1.1072569 0 1.1072569 37
-1.1072569 0 -1.1072569 62
1.1072569 0 -1.1072569 57
-1.1072569 0 -1.1072569 54
0 1.1072569 1.1072569 52
0 -1.1072569 1.1072569 52
0 1.1072569 -1.1072569 53
0 -1.1072569 -1.1072569 40
0.7829487 0 0 43
-0.7829487 0 0 58
0 0 0.7829487 58
0 0 -0.7829487 42
0 0.7829487 0 48
0 -0.7829487 0 44
1.2735263 0 0 50
-1.2735263 0 0 53
0 0 1.2735263 60
0 0 -1.2735263 56
0 1.2735263 0 52
0 -1.2735263 0 32;

proc g3grid data=rp out=out1;
grid N*P=yield / join spline
smooth=.1
axis1=-1.28 to 1.28 by 0.05
axis2=-1.28 to 1.28 by 0.05;
run; quit;

proc g3grid data=rp grid=rpgrid;
grid N*P=yield/join spline
smooth=.1
axis1=-1.28 to 1.28 by 0.1
axis2=-1.28 to 1.28 by 0.1;
run;

proc g3d data=out1;
note j=N f=K "fertilizer used: "
j=n c=red "Nitrogen ";
j=n c=red "Phosphorus ";
plot N*P=Yield/ctext=blue cbottom=red
cTop=green rotate=45 zmin=0 zmax=1.0 xticknum=5 yticknum=5 style=3 grid;
label N='Nitrogen' P='Phosphorus' K='Potassium';
run; quit;
GP1G
DATA RP;
INPUT N P K YIELD;
title 'Surface plot of rose coco yield vs Phosphorus, Potassium';
CARDS;
  1.1072569 1.1072569 0 63
  1.1072569 1.1072569 0 71
  -1.1072569 -1.1072569 0 35
  -1.1072569 -1.1072569 0 56
  1.1072569 0 1.1072569 37
  -1.1072569 0 1.1072569 62
  1.1072569 0 -1.1072569 57
  -1.1072569 0 -1.1072569 54
  0 1.1072569 1.1072569 52
  0 -1.1072569 1.1072569 52
  0 1.1072569 -1.1072569 53
  0 -1.1072569 -1.1072569 40
  0.7829487 0 0 43
  -0.7829487 0 0 58
  0 0 0.7829487 58
  0 0 -0.7829487 42
  0 0.7829487 0 48
  0 -0.7829487 0 44
  1.2735263 0 0 50
  -1.2735263 0 0 53
  0 0 1.2735263 60
  0 0 -1.2735263 56
  0 1.2735263 0 52
  0 -1.2735263 0 32
;
proc g3grid data=rp out=out1;
grid P*K=yield / join spline
  smooth=.1
  axis1=-1.28 to 1.28 by 0.05
  axis2=-1.28 to 1.28 by 0.05;
run; quit;

proc g3grid data=rp out=rpgrid;
grid P*K=yield/join spline
  smooth=.1
  axis1=-1.28 to 1.28 by 0.1
  axis2=-1.28 to 1.28 by 0.1;
run;

proc g3d data=out1;
note j=P f=N "fertilizer used: "
j=n c=red "Phosphorus"
j=n c=red "Potassium"
plot P*K=Yield/ctext=blue cbottom=red
cstop=green rotate=45 zmin=0 zmax=1.0 xticknum=5 yticknum=5 style=3 grid;
label N='Nitrogen' P='Phosphorus' K='Potassium';
run; quit;
PROC G3GRID DATA=RP OUT=OUT1;
GRID P*K=YIELD / JOIN SPLINE
   SMOOTH=.1
   AXIS1=-1.28 TO 1.28 BY 0.01
   AXIS2=-1.28 TO 1.28 BY 0.01;
RUN; QUIT;

PROC G3GRID DATA=RP OUT=RPGRID;
GRID P*K=YIELD / JOIN SPLINE
   SMOOTH=.1
   AXIS1=-1.28 TO 1.28 BY 0.1
   AXIS2=-1.28 TO 1.28 BY 0.1;
RUN;

PROC G3D DATA=OUT1;
NOTE J=P F=N "fertilizer used:"
   J=N C=RED "Phosphorus"
   J=N C=RED "Potassium";
PLOT P*K=YIELD / CTXT=BLUE CBOTTOM=RED
   CTOP=GREEN ROTATE=45 ZMIN=0 ZMAX=1.0 XTIKNUM=5 YTIKNUM=5 STYLE=3 GRID;
LABEL N='Nitrogen' P='Phosphorus' K='Potassium';
RUN; QUIT;
1.2 The codes for generating the contour plots using R software

GP3G

\[
\begin{array}{ccccccc}
& x_1 \ (\text{or } N) & x_2 \ (\text{or } P) & x_3 \ (\text{or } K) & y_1 & y_2 & y_3 \\
1.1072569 & 1.1072569 & 0 & 63 & 49 & 47 \\
-1.1072569 & 1.1072569 & 0 & 71 & 30 & 82 \\
1.1072569 & -1.1072569 & 0 & 35 & 44 & 68 \\
-1.1072569 & -1.1072569 & 0 & 56 & 55 & 50 \\
1.1072569 & 0 & 1.1072569 & 37 & 59 & 69 \\
-1.1072569 & 0 & 1.1072569 & 62 & 53 & 53 \\
1.1072569 & 0 & -1.1072569 & 57 & 37 & 87 \\
-1.1072569 & 0 & -1.1072569 & 54 & 45 & 70 \\
0 & 1.1072569 & 1.1072569 & 52 & 64 & 67 \\
0 & -1.1072569 & 1.1072569 & 52 & 56 & 73 \\
0 & 1.1072569 & -1.1072569 & 53 & 57 & 84 \\
0 & -1.1072569 & -1.1072569 & 40 & 54 & 67 \\
0.7829487 & 0 & 0 & 43 & 34 & 60 \\
-0.7829487 & 0 & 0 & 58 & 39 & 51 \\
0 & 0 & 0.7829487 & 58 & 37 & 51 \\
0 & 0 & -0.7829487 & 42 & 42 & 72 \\
0 & 0.7829487 & 0 & 48 & 47 & 51 \\
0 & -0.7829487 & 0 & 44 & 47 & 54 \\
1.2735263 & 0 & 0 & 50 & 44 & 57 \\
-1.2735263 & 0 & 0 & 53 & 30 & 55 \\
0 & 0 & 1.2735263 & 60 & 46 & 55 \\
0 & 0 & -1.2735263 & 56 & 46 & 75 \\
0 & 1.2735263 & 0 & 52 & 41 & 63 \\
0 & -1.2735263 & 0 & 32 & 41 & 58
\end{array}
\]

```r

# Reading data from CSV file
data<-read.csv("D:\mut\PhD moi Research\Research Ph.D Thesis\yieldrosecoco.csv")
attach(data)
names(data)

# Loading required libraries
library(rsm)

# Generating models
model1<-rsm(y1~SO(N,P,K),data=data)
summary(model1)
model2<-rsm(y2~SO(N,P,K),data=data)
summary(model2)
model3<-rsm(y3~SO(N,P,K),data=data)
summary(model3)

# Plotting contour plots
par(mfrow=c(1,3))
persp(model1,~N+P+K)
par(mfrow=c(1,3))
persp(model2,~N+P+K)
```


1.3 The codes for generating the Ridge Analysis using SAS software

GP3G
DATA RP;
INPUT N P K YIELD;
CARDS;
1.1072569 1.1072569 0 47
-1.1072569 1.1072569 0 82
1.1072569 -1.1072569 0 68
-1.1072569 -1.1072569 0 50
1.1072569 0 1.1072569 69
-1.1072569 0 1.1072569 53
1.1072569 0 -1.1072569 87
-1.1072569 0 -1.1072569 70
0 1.1072569 1.1072569 67
0 -1.1072569 1.1072569 73
0 1.1072569 -1.1072569 84
0 -1.1072569 -1.1072569 67
0.7829487 0 0 60
-0.7829487 0 0 51
0 0 0.7829487 51
0 0 -0.7829487 72
0 0.7829487 0 51
0 -0.7829487 0 54
1.2735263 0 0 57
-1.2735263 0 0 55
0 0 1.2735263 55
0 0 -1.2735263 75
0 1.2735263 0 63
0 -1.2735263 0 58
;
PROC RSREG;
MODEL YIELD = N P K/LACKFIT NOCODE PRESS;
ridge max min;
title 'Rose coco beans analysis';
RUN;
1.4 Greenhouse photos

- Constructed greenhouse for rose coco beans
- Mr. Elias assisting to fix drip line pipes
- Rose coco beans three weeks after planting
- Rose coco beans four week after planting
- Rose coco beans six weeks after planting
- The rose coco beans starting to flower
The rose coco beans in flowering stages

The rose coco beans in flowering stages

The rose coco beans have some developed pods

A climbing rose coco plant support with a string

Rose coco beans have fully develop pods

Rose coco beans has fully develop pods
Some rose coco beans start to ripe
Some rose coco beans start to dry up
Harvested beans in white labeled polythene bags
Showing each rose coco bean harvested per plant
Rose coco beans on polythene bags
Dried rose coco beans