

Modeling of rose coco beans using twenty four points optimum second order rotatable design

Tum Isaac Kipkosgei *, Mutiso John Muindi, Koske Joseph Kipsigei

Department of Statistics and Computer Science, Moi University, Eldoret, Kenya

*Corresponding author E-mail: isaactum@gmail.com

Abstract

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 the response. The objective of the study was to model the *rose coco* beans (*Phaseolus vulgaris*) through an existing A-optimum and D-efficient second order rotatable design of twenty four points in three dimensions in a greenhouse setting using three inorganic fertilizers, namely, nitrogen, phosphorus and potassium. Thus, the objective of the study was accomplished using the calculus optimum value of the free/letter parameter $f=1.1072569$. This was done by estimating the parameters via least square's techniques, by making available for the yield response of rose coco beans at calculus optimum value design for the first time. The results showed that, the three factors: nitrogen, phosphorus, and potassium contributed significantly on the yield of *rose coco* beans ($p<0.05$). In GP3G, the second-order model was adequate for 1% level of significance with p value of 0.0034. The analysis of variance (ANOVA) of response surface for rose coco yield showed that this design was adequate due to satisfactory level of a coefficient of determination, R^2 , 0.8066 and coefficient variation, CV was 10.30. This study demonstrated the importance of statistical methods in the optimal and efficient production of rose coco beans. We do recommend a randomize screening of all the fertilizer components with which it has influence on *rose coco* beans be done to ascertain the right initial amount of each fertilizer that could achieve maximum yield than this study realized.

Keywords: Response Surface Methodology; Second Order Design; Optimality; Coded Levels; Natural Levels; Calculus Optimum Value; Rose Coco Beans.

1. Introduction

The research was done with a combination of independent treatment factors of inorganic fertilizers to check how the factors of fertilizers influence the yield of *rose coco* beans. The experiment was carried out in a greenhouse of size 15m x 10m during the period of February-July 2016 on the twenty four points second order rotatable design in three dimensions. The focus of this paper was to achieve an optimal and efficient production of *rose coco* beans (*Phaseolus vulgaris*) through an existing A-optimum and D-efficient second order rotatable design of twenty four points in three dimensions in a greenhouse setting using three inorganic fertilizers, namely, nitrogen, phosphorus and potassium. The study was geared to estimate the linear parameters, by making available for the yield response of *rose coco* beans at calculus optimum value design in one of the existing six specific second orders rotatable designs in three dimensions in which Mutiso [9] calculated the calculus optimal value to be 1.1072569 for the free/letter parameter. Koech [7] 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 dwell much 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 inorganic fertilizer components. 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* bean was capable of delivering the bean production efficiency. The study founds how the yield fitted the second-order model and checked its model adequacy. 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. If the fertilizer application rates are greater or smaller than the optimum application rates, they might result in a reduction in the yields. The purpose of implementing this RSM technique is 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 advance the current methods being used in farming in order to maximize the yield using the scarce resources available. It was 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.

2. 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

[8]. 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. In order to get the most efficient result in the approximation of polynomials the proper experimental design must be used to collect data. The representation $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ was given where \mathbf{y} is a vector of observations, $\boldsymbol{\varepsilon}$ is the vector of errors, \mathbf{X} is the design matrix and $\boldsymbol{\beta}$ 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 $\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ was used to estimate parameters in the polynomials by using the least squares method. In this study, we concentrated on the statistical modeling to develop an appropriate approximating model between the response y and independent variables x_1 (nitrogen), x_2 (phosphorus), x_3 (potassium). The second-order model is likely required in these situations. In the case of three variables, the second-order model is given by:

$$\hat{y} = \beta_0 + \sum_i \beta_i x_i + \sum_i \beta_{ii} x_i^2 + \sum_{i < j=2}^k \beta_{ij} x_i x_j \quad (1)$$

Where \hat{y} is the measured *rose coco* response, β_0 is the intercept term, β_i are the linear coefficients, β_{ij} is the logarithmic coefficient, β_{ii} are the quadratic coefficients and x_{iu} denotes the coded level of the i^{th} factor ($i=1,2,3$) in the u^{th} run ($u=1,2,\dots,24$) of the experiment. The parameters of the model $\beta_0, \beta_i, \beta_{ii}, \beta_{ij}$ are estimated by least squares estimation to provide $\hat{\beta}_0, \hat{\beta}_i, \hat{\beta}_{ii}, \hat{\beta}_{ij}$. It is said to be SORD if the variance of the estimate of the response \hat{y}_u is only a function

of the distance $(d^2 = \sum_{i=1}^k x_i^2)$ of the point (x_1, x_2, \dots, x_k) from the origin (centre) of the design [6]. This model would likely be useful as an approximation to the true response surface in a relatively small region. The low and high factor settings are coded as negative and positive, the midpoint coded as 0.

3. Experiment layout of twenty four points calculus optimum value design

The *rose coco* plants were planted once on spacing of 75x30cm with N, P and K fertilizers during planting, and thereafter no other supplements were added either by top dressing or foliar spray. The *rose coco* beans were subjected to inorganic fertilizer's N, P, K at different levels of a twenty four point second order rotatable design. Therefore, GP3G was given three replications; for example, in GP3G we had GP3GA, GP3GB and GP3GC. In GP3G a combination of 30 grams of nitrogen, 40grams of phosphorus and 50grams of potassium where 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 greenhouse, organic matter on the soil surface was cleared. The field was prepared by Jembe ploughing followed by harrowing until fine tillage was obtained. We had three replications of twenty four design points. Certified, viable and uniform seeds of *rose coco* beans were planted in the plots in February 2016. The beans were planted as a pure stand in a greenhouse. Before planting, the bean's seeds were dressed with Aldrin at the rate of 5g per kg of seeds, to control soil pests especially been flown (*Melanargromyza phaseoli*). Furadan (5% carbofuran) was applied in the bean rows at sowing to control cutworms

(*Agrotis ipsilon*). All the plots were given a drip irrigation using drip line pipes. The inorganic fertilizers (nitrogen, phosphorus, potassium) combinations were applied in each experimental unit before planting two seeds of *rose coco* beans, and one week after germination, they were thinned to one plant per experimental unit. First weeding was carried out on the greenhouse at two weeks after emergence. Second weeding was carried out four weeks later. We consider a set of twenty four point's rotatable designs as highlighted by Draper [4], Mutiso J.M (1998) and Koske et al (2008) is given as:

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

4. A practical greenhouse example

The rotatable design (3) was set up in a greenhouse in an area in Saroiyot, Kesses, Uasin-Gishu, Kenya in order to investigate the effects of three fertilizer ingredients N, P, K on the yield of *rose coco* beans. The fertilizer ingredients and actual amount applied were nitrogen (N) $x_{1u}, \Psi_{1.}=30$ grams/hole; phosphorus (P) $x_{2u}, \Psi_{2.}=40$ grams/hole; and potassium (K) $x_{3u}, \Psi_{3.}=50$ grams/hole. The response of interest was the average yield of *rose coco* in mg per plant. The set of twenty four points in (2) forms a second order rotatable arrangement in three factors. The calculations done by Koske [5] showed that $f=1.1072569$, $c_1=0.7829487$, $c_2=1.2735263$, hence from (2) the design D_1 yields an optimum design as:

$$D_2 = [\frac{1}{2} G(1.1072569, 1.1072569, 0) + \frac{1}{4} G(0.7829487, 0, 0) + \frac{1}{4} G(1.2735263, 0, 0)] \quad (3)$$

Let the scale parameters S_i , assume $s_1=0.5$, $s_2=0.3$ and $s_3=1$. According to Box [2] and Box and Wilson [3] it can be reverted to the natural levels denoted by Ψ_{iu} where Bose and Draper [1] scaling condition fixes a particular design when $\lambda_2=1$ where

$$x_{iu} = \frac{\Psi_{iu} - \Psi_i}{S_i}, \Psi_i = \frac{\sum_{u=1}^N \Psi_{iu}}{N}, S_i = \left[\frac{\sum_{u=1}^N (\Psi_{iu} - \Psi_i)^2}{N} \right]^{0.5}, \Psi_i = x_{iu} S_i + \Psi_i, \sum_{u=1}^N x_{iu}^2 = N \text{ and } \sum_{u=1}^N x_{iu} = 0. \quad (4)$$

The experimenter tries to quantify the relationship between a set of 3 predictor variables $\xi' = (\xi_1, \xi_2, \xi_3)$ 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 ξ_i are transformed into coded x_{iu} by $(\Psi_{iu} - \Psi_i)/S_i$, $i=1,2,3$ where Ψ_{iu} and $S_i > 0$ are the centering and scaling constants, respectively. Often, a second order model fit to the experimental data, including all linear, quadratic and cross product terms for the x_{iu} .

The table 1 below shows the yield which was obtained in Tum [10] for twenty four point's rotatable design of coded levels and natural levels with the yield of *rose coco* beans where x_{1u} , x_{2u} and x_{3u} are coded values while Ψ_{1u} , Ψ_{2u} and Ψ_{3u} are natural values-GP3G. The natural values (Ψ_{iu}) of fertilizers at the ratio of 30:40:50 N: P: K fertilizers, were measured using a sensitive weighing scale and planted in a greenhouse which gave the observed yield in grams $-y_i$ and predicted yield \hat{y} using the second order model of GP3G.

$$(X'X)^{-1} = \begin{bmatrix} 0.2735 & 0 & 0 & 0 & -0.1299 & -0.1299 & -0.1299 & 0 & 0 & 0 \\ 0 & 0.0700 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0700 & & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0700 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.1299 & 0 & 0 & 0 & 0.1282 & 0.0451 & 0.0451 & 0 & 0 & 0 \\ -0.1299 & 0 & 0 & 0 & 0.0451 & 0.1282 & 0.0451 & 0 & 0 & 0 \\ -0.1299 & 0 & 0 & 0 & 0.0451 & 0.0451 & 0.1282 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1663 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1663 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1663 \end{bmatrix} \quad (7)$$

$$(X'Y_{GP3G}) = \begin{bmatrix} 1519 \\ 27.3 \\ 28.4 \\ -92.8 \\ 894.6 \\ 920.2 \\ 985.1 \\ -65 \\ -1.2 \\ -28.2 \end{bmatrix} \quad \hat{B} = (X'X)^{-1} X'Y_{GP3G} = \begin{bmatrix} 51.7514 \\ 1.9127 \\ 1.9876 \\ -6.5028 \\ 3.2471 \\ 5.3785 \\ 10.7726 \\ -10.8073 \\ -0.2039 \\ -4.6900 \end{bmatrix} \quad (8)$$

The eq. 8 shows the regression coefficients for the second-order polynomial model given as;

$$\hat{Y}_{GP3G} = 51.7514 + 1.9127X_1 + 1.9876X_2 - 6.5028X_3 + 3.2471X_1^2 + 5.3785X_2^2 + 10.7726X_3^2 - 10.8073X_1X_2 - 0.2039X_1X_3 - 4.6900X_2X_3 \quad (9)$$

The \hat{Y}_{GP3G} is the predicted response for *rose coco* beans in group 3 (GP3G). The regression coefficients of the linear term for potassium (p=0.0021) have significant effects on the yield (p-value <0.05), the quadratic P² (p=0.0371) and K² (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, K², N*P was significant at the 1% significance level, while P² was significant at the 5% level, those other terms of the model showed no significant effect on the yield. The positive coefficients of P², K² enhance the yield since they are the largest coefficients in the model eq. 9. The largest negative coefficient of K and NP minimizes the yield of *rose coco* at their respective fertilizer input. The X₁ represents nitrogen (N), X₂ represents phosphorus (P), X₃ represents potassium (K). This suggests that potassium (K), quadratic effects of phosphorus (P²), potassium (K²) 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.

Table 2: The ANOVA Results in the Fertilizer Concentration on Rose Coco Bean-GP3G

Regression	D F	Type I Sum of Squares	Mean Sq	R-Square	F Value	Pr > F
First Order	3	712.400678	237.467	0.2314	5.58	0.0099
Pure Quadratic	3	936.326222	312.109	0.3041	7.34	0.0034
Two-Factor Interaction	3	834.750000	278.250	0.2711	6.54	0.0054
Residuals	14	595.48	42.534			
Lack of fit	4	595.48	42.534			
Pure error	0	0.00				
Total Model	9	2483.476900		0.8066	6.49	0.0011

Coefficient of variation (CV)= 10.3044, coefficient determination (R²)= 0.8066, Adjusted R-squared =0.6823, correlation coefficient (r)=0.8981, root MSE=6.521839, response mean=63.291667, PRESS=2645.6089489.

Table 2 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 2. In the table 2, 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 (R²) and correlation coefficient (r). The determination coefficient (R²) 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 R² in GP3G of second order model explain 80.66% of the variation in the model. The value of r (0.8981) for eq. 9 being close to 1 indicated a close agreement between the experimental results and the theoretical values predicted by the model equation.

6. Conclusions

In this research work we have been able to successfully utilize response surface methodology to come up with a clear model for the relationship involving nitrogen, phosphorus and potassium as fertilizers variables for the production of *rose coco* beans using the twenty four point second order rotatable design in a greenhouse setting. In this study, the average *rose coco* yield was obtained using three replications for the group GP3G. We were able to estimate the parameter coefficient for the GP3G. 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, R² (0.80) for the GP3G which means that 80.03% of the total variability of the system was explained by the chosen factors N, P, K and coefficient of variations (CV, 10.30). 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. The group GP3Gtotal model was significant at 0.1% (p = 0.0011). The results showed GP3G fitted the second order models well for the *rose coco* yield using the three fertilizer treatments, nitrogen (N), phosphorus (P) and potassium (K). Furthermore, 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.

References

- [1] Bose, R. C. and Draper, N.R.(1959). Second order rotatable designs in three dimensions. *Ann. Math. Stat.*, Vol. 30, (1959): pp. 1097-1112. <https://doi.org/10.1214/aoms/1177706093>.
- [2] Box, G.E.P. (1952). *Multi-factor designs of the first order*. *Journal of Biometrika*: 39, 49-57. <https://doi.org/10.1093/biomet/39.1-2.49>.
- [3] Box, G.E.P, Wilson, K.B. (1951). *On the Experimental Attainment of Optimum Conditions*. *Journal of the Royal Statistical Society B*: 13, 1–45.
- [4] Draper, N.R. (1960). *Second order rotatable designs in four or more dimension*. *Ann. Math. Stat.*, Vol. 31, pp. 23-33. <https://doi.org/10.1214/aoms/1177705984>.
- [5] Koske, J.K., Mutiso J.M., & Kosgei, M.K. (2008). *A specific optimum second order rotatable design of twenty four points with a practical example*. *East African Journal of pure and applied science*, school of science, Moi University, Eldoret Kenya.
- [6] Koske, J.K., Mutiso J.M., & Tum, I.K. (2012). Construction of a Practical Optimum Second Order Rotatable Designs in the Three Dimensions. *Advances and Applications in Statistics*. Vol. 30, No. 1, 2012 pp 31-43, ISSN 0972-3617.
- [7] Koech, F. (2016). Relative efficiency and DT- optimality criteria for the six specific secondorder rotatable designs. Unpublished M.Phil. Thesis, Moi University.
- [8] Montgomery D. C. (2005). *Design and Analysis of Experiments response surface methods and design*. John Willey and Sons Inc. New Jersey.
- [9] Mutiso, J. M. (1998). *Second and Third Order Specific and Sequential Rotatable Designs in K Dimensions*. D.Phil. Thesis, Moi University.
- [10] Tum, I.K (2017). *Optimal and efficient production of rose coco beans through the twenty four points second order rotatable design*. Unpublished Ph.D Thesis, Moi University.