RESPONSE SURFACE OPTIMIZATION OF DIETARY IRON, CALCIUM AND VITAMIN C IN SOYAMILK FOR COMPLEMENTARY FEEDING

Abstract:
Response surface methodology (RSM) is a collection of tools developed in the 1950s for the purpose of determining optimum operating conditions. In this work, a three level three factor (3^3) factorial design that metamorphosed to the response surface design with two augmented central point was employed. In its applications a secondary data from the department of Food Science and Technology (FST), Michael Okpara University of Agriculture, Umudike (MOUAU), containing the mineral components of soymilk for complementary feeding of infants was used. The analysis for the First Order (FO), Two Way Interaction (TWI) and the Polynomial (PQ) model was carried out and the augmented response surface analysis was performed. Following the path of steepest ascent, an optimality condition from the surface and contour lines shows that dietary iron is significant for varying the colour content, while calcium was significant for varying the ash and moisture content. It was then recommended that for optimal colour content in the soymilk, 3.07mg/100ml of Dietary Iron, 154.1mg/100ml of Calcium and 24.23mg/100ml of Vitamin C should be used, for optimal ash content in the soymilk, 2.22mg/100ml of Dietary Iron, 152.03mg/100ml of Calcium and 13.53mg/100ml of Vitamin C should be used while for optimal moisture content in the soymilk, 2.9858mg/100ml of dietary Iron, 335.71mg/100ml of Calcium and 25.48mg/100ml of Vitamin C should be used.

Introduction

Response surface methodology is a procedure and a philosophy for the design, the conduct, the analysis, and the interpretation of experiments performed to determine the quantitative relationship between a dependent variable (the response) and one or more quantitative, continuous independent variables.

The basic approach, first suggested by Box and Wilson (1951), ingeniously combined elements of multiple regression theory and its specialized form in analysis of variance with special features of the factorial designs, including principles of partitioning, confounding and fractional replicates. An important aspect in applying response surface methodology is the design of experiment. These designs differ from one another with respect to their selection of experimental points, number of runs and block. The 3k factorial, the central composite design etc are commonly used in response surface methodology (Solomon, 2021). Central composite design is one of a number of experimental designs developed specially for use in response surface exploration in order that the data collection phase be performed as completely, as cheaply and as efficiently as possible.

Box and Hunter (2005), suggested the characteristics of experimental designs for fitting response surfaces.

Naturally, Soymilk also known as vegetable or imitation milk is produced from whole soybean (glycine max).Soymilk resembles cow's milk in appearance flavor and nutritive value. When properly processed (Iwe, 2003) should contain 8.25% solid-not fat; not less than 3.25% fat, not more than 88.50% water and not less than 11.50% solids including 3.25% fat.

Soymilk or soy drink is a stable off-white emulsion/suspension water extract of whole soybean that contain 2% oil, 88% water, water soluble protein of about 3.5%, 2.9% carbohydrate, 0.5%

ash and others like Calcium Iron Lecithin, Riboflavin, Isoflavones and Vitamin E. The protein content is higher than that of cow's milk by about 2.2%. Addition of Iron, Calcium and Vitamin C in foods including beverages like soymilk will not only meet the nutrient needs of infants and young children, but they also will work in synergy to promote growth performance (Clarke, 1995; Thiers, 2009). Among other sources of vegetable milk, soymilk had received very high research attention as reference vegetable milk (Onweluzo and Nwakalor, 2009).

Infancy is a period of tremendous physical growth characterize by increase in length and weight as well as physiological, immunological and mental development. Yeung, (2011). Furthermore, the composition of human milk varies in magnitude between nutrients in lactating mothers. Lonnerdal, (1985). These and more portend that with time (about 4-6 months to 12 months) the breast milk alone will not be sufficient to meet the child's nutrition for energy needs, it therefore becomes mandatory for an adequate and appropriate nutrition (in terms of calorie, vitamins and minerals) be introduced, else the infant will not achieve the expected growth. A complementary food is therefore introduced to improve both the energy and nutrient intakes since the child will no longer gain weight despite appropriate breast feeding, and will be feeling hungry always despite frequent breast feeding. Rarback, (2011). Some experimental analysis have been presented to ascertain these physio-chemical changes which include determination of PH, viscosity etc.

Factorial experiments are employed in all fields of life; agricultural science, biology, medicine, the physical sciences etc. experiments are usually carried out by researchers either to discover something about a particular process or to compare the effects of several factors on responses.

Factorial experiment is therefore a crossed factor design that usually involves several factors and it is such that every possible combination of the factor is included or observed or examined. Factorial experiment permits the analysis of a number of factors with the same precision (eg. Individual and joint effects) as if the entire experiment had been devoted to the study of only one factor.

Some notable factorial experiments are as follows;

2^k factorial experiments- This involves K factors each at two levels.

3^K factorial experiment- This involves K factors each at three levels.

B^K factorial experiment- This involves K factors each at B levels. (Montgomery 2013).

The problem:

According to Fallon and Enig (2007), Infants are at high risk of iron, protein and calcium deficiencies after six months of exclusive breast feeding. They stated that soymilk can serve as a supplement towards improving on these deficiencies. The problem now lies on determining the composition and quality of soymilk to be used. Furthermore, the composition and quality of soymilk varies with the variety of soybean used and the method of production (Wang et al., 1978). It is therefore necessary to statistically circumvent these problems.

The aim of this study is to apply response surface method in determining the optimal combination of levels of different component of soymilk as a complement for feeding infant after six months of exclusive breast feeding.

The specific objectives Include; To determine the optimal combinations of levels of the different component of soymilk (that is; Dietary Iron (Fe), calcium (Ca) & Vitamin C (C) that is suitable for complementary feeding in infants. To investigate the linear relationship as well as the curvature (quadratic) relationship using the response surface analysis. This study is going to help determine the combination of level of the different components of soymilk (that is; Dietary Iron (Fe) (X_1), calcium (Ca)(X_2) & Vitamin C (C) (X_3) that will be suitable for complementary feeding in infants, and would be a guide to people using this approach. This study is limited to the use of response surface methodology to analyze the different levels of the mineral components that made up soymilk for complementary feeding.

Methodology

The data used for this study are secondary data. It was gotten from the work “Application of response surface methodology to fortification of soymilk from sprouted soybean with Iron, Calcium and Vitamin C for complementary feeding” (FST), published by Dr. I. N. Okwunodulu (2014) from the department of Food Science and Technology, MOUAU.

RESPONSE SURFACE METHODOLOGY (RSM)- The method of response surface analysis was adopted for this study. However RSM has been reported to be the best option since it identifies the effects of the individual process variables, locates optimum process variable. Combinations for multivariable system efficiency, and offers economy of experimental points since it requires least experimental data. (Mullen and Ennia, 1979).

The objective of RSM experiments is to predict the value of a response variable called the dependent variable based on the controlled value of experimental factors called independent variables. As an important subject in the statistical design of experiment, the *Response surface methodology(RSM) is a collection of mathematical or statistical techniques that are 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.*

A response surface is fitted an extension of linear model algorithm, and works almost exactly like it; however, the model formula for response surface must make use of the First Order, Two-Way Interaction, Pure Quadratic or Second Order models where the First Order model in (1.0) is given by:

$$y_{ijk} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon_{ijk} \quad (3.1).$$

Adding the interaction term introduces curvature into the response function. It is necessary to use more than two levels to detect curvature, however, and, in general, to determine the shape of the “response surface.” That is, how does the response vary over different combinations of values of the factors? In what region(s), if any, is the change approximately linear? Are there humps, and valleys, and saddle points, and, if so, where do they occur? These are the types of questions that response surface methodology attempts to answer.

We usually represent the response surface graphically. To help visualize the shape of the response surface, we often plot the contours of the response surface. Each contour corresponds to a particular height of the response surface.

In most response surface methodology, the form of the relationship between the response and the independent variable is unknown. The first step in RSM is to find a suitable approximation for the true functional relationship between the set of independent variables. (Montgomery 2013).

RSA is based on the assumption that when k factors (or independent variables) are being studied in an experiment, the response (or dependent variable) will be a function of the levels at which these factors are combined (xk).

The Sequential Nature of the Response Surface Methodology

Most applications of RSM are sequential in nature.

Phase 0: At first some ideas are generated concerning which factors or variables are likely to be important in response surface study. It is usually called a **screening experiment**. The objective of factor screening is to reduce the list of candidate variables to a relatively few so that subsequent experiments will be more efficient and require fewer runs or tests. The purpose of this phase is the identification of the important independent variables.

Phase 1: The experimenter's objective is to determine if the current settings of the independent variables result in a value of the response that is near the optimum. If the current settings or levels of the independent variables are not consistent with optimum performance, then the experimenter must determine a set of adjustments to the process variables that will move the process toward the optimum. This phase of RSM makes considerable use of the first-order model and an optimization technique called the **method of steepest ascent (descent)**.

Phase 2: Phase 2 begins when the process is near the optimum. At this point the experimenter usually wants a model that will accurately approximate the true response function within a relatively small region around the optimum. Because the true response surface usually exhibits curvature near the optimum, a second-order model (or perhaps some higher-order polynomial) should be used. Once an appropriate approximating model has been obtained, this model may be analyzed to determine the optimum conditions for the process.

MATHEMATICAL MODEL EQUATIONS AND RSM MODEL SOLUTION SEARCH-

RSM is often represented with mathematical models that resemble those of regression equations. These mathematical model equations are determined by the number of process variable cases involved and have the probability of showing significant/non-significant linear, quadratic and cross product (interaction) order effects. Those variables significant ($p>0.05$) at 1% level do not contribute at 5% level ($p\leq 0.05$) (Bradley, 2007).

For two process variable cases, the mathematical model is;

$$y_{ijk} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_{12} + \varepsilon_{ijk} \qquad (3.2)$$

This model equation indicates significant/ non-significant two linear, two quadratic and a single cross product or interaction order effect.

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 will often work well as an approximation to the true response surface.
2. It is easy to estimate the parameters (the β's) in the second-order model. The method of least squares can be used for this purpose.
3. There is considerable practical experience indicating that second-order models work well in solving real response surface problems.

For three process variable cases, the mathematical model is;

$$y_{ijk} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \beta_{12} x_{12} + \beta_{13} x_{13} + \beta_{23} x_{23} + \varepsilon_{ijk} \qquad (3.2.1)$$

This model equation indicates significant/ non-significant three linear, three quadratic and three cross product or interaction order effect.

Where; β_0= Intercept

B_i (ie. B_1, β_2, β_3)= linear regression terms

B_{ii} (ie. B_{11}, β_{22}, β_{33})= quadratic regression terms

B_{ij} (ie. B_{12}, β_{13}, β_{23})= cross product regression terms

ε= Random error term.

y_{ijk} = Dependent response variable

Normality, independent, homogeneity of variance assumptions were justified using the necessary approaches.

3.3 FACTORIAL DESIGNS (THE CCD'S WITH 5 AXIAL POINTS)

Factorial designs are designs in terms of treatment combinations (a factorial treatment design) which is a necessary first step in designing a factorial experiment after the factors and their levels are known. The central composite design is an experimental design, useful in response surface methodology, for building a second order (quadratic) model for the response variable without needing to use a complete three level factorial experiment and also for obtaining optimum conditions. The design consists of three distinct sets of experimental runs;

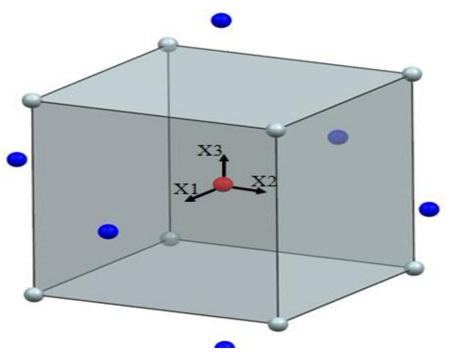
1. A factorial (perhaps fractional) design in the factors studied, each having two levels.
2. A set of center points, experimental runs whose values of each factor are the medians of the values used in the factorial portion. This point is often replicated in order to improve the precision of the experiment.
3. A set of axial points, experimental runs identical to the center points except for one factor, which will take on values both below and above the median of the two factorial levels and typically both outside their range. All factors are varied in this way.

The *central composite designs* (*CCDs*) are the most frequently used response surface designs. These designs permit the estimation of nonlinear effects and are constructed by starting with a two-level full factorial and then adding center points (i.e., at the center of the full factorial) and axial (star) points that lie outside the square formed from connecting the factorial points. The design for two factors is shown in Figure 12.7.The value of α would be selected by the experimenter. Desirable properties of the design include orthogonality and rotatability. A design is rotatable if $Var(\hat{y})$ is the same for all points equidistant from the center of the design. In the case of two factors, rotatability is achieved when α = 1.414. If there are not tight constraints on the number of center points that can be used, both orthogonality and rotatability can be achieved by selecting the number of center points to achieve orthogonality since the number of center points does not affect rotatability.

CCD enables estimation of the regression parameters to fit a second-degree polynomial regression model to a given response. A polynomial, as given by Equation (1.7), quantifies relationships among the measured response y and a number of experimental variables *X1…Xk*, where k is the number of factors considered, β are regressors and ε is an error associated with the model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \ldots + \beta_{kk} x_k^2 + \beta_{12} x_1 x_2 + \ldots + \beta_{k-1,k} x_{k-1} x_k + \varepsilon \quad (3.9)$$

The regressors (β1, β2, β3…) in the various terms of Equation (3.9) provide a quantitative measure of the significance of linear effects, curvilinear effects of factors and interactions between factors. It is worth noting that the model presented by Equation (3.9) is not a model in purely physical sense, but rather it should be understood as a statistical model, *i.e.*, a correlation developed based on regression analysis.
However, this nomenclature is widely used in the field of design of experiments and statistics, and therefore it is used hereafter.



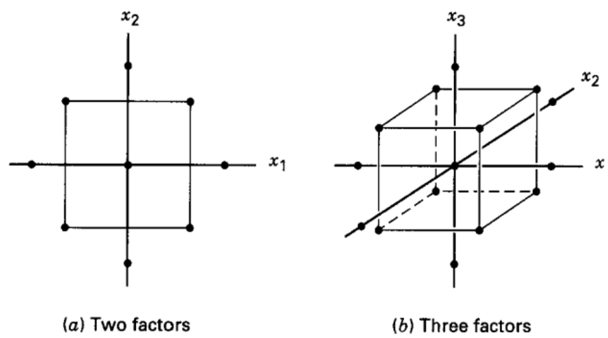


Figure 6-36 Central composite designs.

Figure 3.1. Visualization of original type rotatable CCD for 3 factors X1, X2 and X3

Values at the center point (red point with coordinates 0, 0, 0) that is located in the center of the cube are used to detect curvature in the response, *i.e.*, they contribute to the estimation of the coefficients of quadratic terms. Axial points (six blue points located at a distance α from the center point) are also used to estimate the coefficients of quadratic terms, while factorial points (eight grey points located at corners of the cube with a side length equal to 2) are used mainly to estimate the coefficients of linear terms and two-way interactions. For testing four or more factors in an experiment, Figure 3.1 should be extended to the fourth or more dimensions (Marcin et al. 2015).

3.4 THE METHOD OF STEEPEST ASCENT AND STEEPEST DESCENTS

The method of steepest ascent is a procedure for moving sequentially along the path of steepest ascent, that is, along the direction of maximum increase in the predicted response. The fitted first order model is; $y_{ijk} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i$

(3.9.1)

And the fisrt order response surface, that is the contors of y, is a series of parrallel line such as that shown in figure 1.2 below. The direction of the steepest ascent is the direction in which y increases most rapidly. This direction is parallel to the normal to the fitted resonse surface. We usually take as the path of steepest ascent the line through the centre of the region of interest and normal to the fitted surface

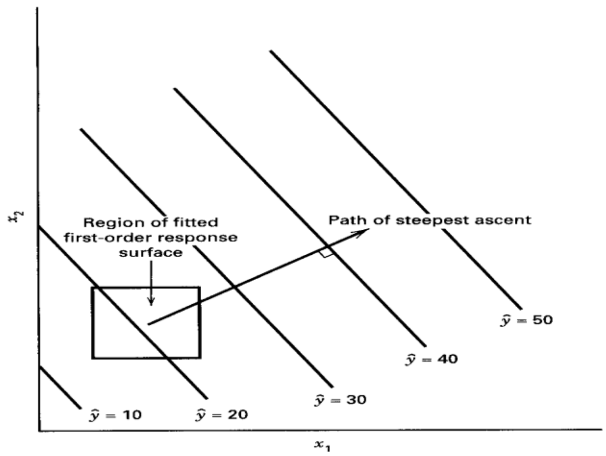


Fig 3. First-order response surface and path of steepest ascent (Montgomery, 2013)

3.4.1 STEPS INVOLVED IN THE METHOD OF STEEPEST ASCENT

The following steps describe the general procedure:

1. The experimenter runs a first-order model. In some restricted region of variables, $X_1, X_2, \ldots, X_k$. The experiment usually contains center-points runs which can be used to perform a lack-of-fit test for curvature.
   (a) If a lack-of-fit for curvature is not significant, then go to step 2,
   (b) If lack-of-fit for curvature is significant and the experimenter is satisfied that little or no additional information can be obtained using the path of steepest ascent (or path of steepest descent) procedure, then we stop applying this method.
2. The fitted first-order model is used to determine a path of steepest ascent (or path of steepest descent).
3. A series of experimental runs is conducted along the path until no additional increase (or decrease) in response is evident.
4. Centered near the location along this path which yields a maximum (or minimum) response, a new experiment is designed.
5. Return to step 1.Once curvature is detected and the method is stopped, a more elaborate experiment to fit a quadratic response surface model should be designed and conducted.

3.5 METHOD OF STEEPEST DESCENT

The method of steepest descent is a procedure for moving sequentially along the path of steepest descent, that is, along the direction of maximum decrease in the predicted response. The steepest descent method, which is based on the gradient of $\varepsilon'\varepsilon$. Where;

$$\varepsilon'\varepsilon = [y - f(x,\Omega)]'[y - f(x,\Omega)]$$

(3.9.2)

Therefore, the gradient of $\varepsilon'\varepsilon$ is given by

$$\frac{\partial(\varepsilon'\varepsilon)}{\partial\Omega} = -2\left[\frac{\partial fx;\Omega}{\partial\Omega}\right]'[y - f(x;\Omega)] \tag{3.9.3}$$

4.0 RESULT AND DISCUSSION

This chapter presents the analysis and results which was done using R. We have discussed the methodologies that were adopted in this study in the previous chapter. The analysis presents a response surface model for the optimal concentration of soya milk using dietary iron, calcium and vitamin c.

4.1 STATISTICAL DATA ANALYSIS See data presentation at appendix (page 22-26)

The Response Surface Methodology (RSM) was used to find the mixture (in quantity) of Iron, Calcium and Vitamin C that gives the optimum soymilk concentration which can be used for feeding. This analysis was presented to optimize three contents of soymilk, colour, ash and moisture content.

4.1.1 COLOUR CONTENT OF SOYMILK

Table 4.1 shows the regression result of the model for getting the optimal colour for soymilk. The regression model contains the linear effect, interaction effect and the polynomial effect. Since we have three independent variable, the model we fitted has three two-way interaction term. From the p-value presented in the table, we do not have enough evidence to suggest that there is a significant interaction effect. That is, the effect of dietary Iron on the colour of soybean do not depend on either Calcium or Vitamin C and vice versa.

Furthermore, the three squared effects presented in the model are not significant since the p-value for each squared effect is greater than 0.05 which is the significant level we chose. Thus, there is enough evidence for us to conclude that there is no curvature in the response surface. This can also be seen from the surface plot which is present in Figure 4.2. Since the square terms were identified as not significant in the analysis of variance, there is no need for looking at each of them individually.

On the regression model and the analysis of variance, the model contains three linear effects (Iron, Calcium and Vitamin C). The p-value of Iron is less than 0.05. Therefore, there is a significant linear effect for Iron. This means that we have enough evidence to conclude that the variation in the response (colour of soymilk) depends on the quantity of Iron present. Calcium and Vitamin C are however not significant for the colour content.

Finally, the constant term was also significant in the model. The p-value for the lack of fit in the analysis of variance table is greater than 0.05. Thus, there is no evidence to that the model does not adequately explain the variation in the responses. The Adjusted R^2 suggests that only 9% variation of the response have been accounted for by the variation in explanatory variable. This is however considerable low, as such some important terms which affects the variation of the response can be included also.

The regression model thus is presented as:

$$\begin{aligned} Y &= 0.63 + 0.071X_1 + 0.025X_2 - 0.003X_3 + 0.014X_1X_2 - 0.00001X_1X_3 - 0.0013X_2X_3 \\ &\quad + 0.004X_1^2 - 0.00333X_2^2 - 0.018X_3^2 \\ &where: X_1, X_2, X_3 = \text{ Iron, Calcium and Vitamin C respectively} \end{aligned} \tag{4.1}$$

The contour plot and surface plot is presented in Figure 4.1 below.

TABLE 4.0 REGRESSION MODEL FOR COLOUR CONTENT

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	6.3046e-01	6.4155e-02	9.8271	3.869e-16	***
x1	7.0659e-02	1.7460e-02	4.0468	0.0001058	***
x2	2.4713e-02	1.7460e-02	1.4154	0.1602236	
x3	-3.5846e-03	1.7460e-02	-0.2053	0.8377786	
x_1:X_2	1.4804e-02	2.2814e-02	0.6489	0.5179860	
x_1:X_3	-1.7857e-05	2.2814e-02	-0.0008	0.9993771	
x2:X_3	-1.3393e-03	2.2814e-02	-0.0587	0.9533113	
x_1^2	4.2221e-03	2.6226e-02	0.1610	0.8724430	
x_2^2	-3.3269e-03	2.6226e-02	-0.1269	0.8993217	
x_3^2	-1.8299e-02	2.6226e-02	-0.6977	0.4870454	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Multiple R-squared: 0.1734, Adjusted R-squared: 0.09504

F-statistic: 2.214 on 9 and 95 DF, p-value: 0.02754

TABLE 4.1 ANALYSIS OF VARIANCE TABLE FOR COLOUR CONTENT

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
FO(x1, x2, x3)	3	0.53696	0.178986	6.1407	0.0007324
TWI(x1, x2, x3)	3	0.01237	0.004124	0.1415	0.9348720
PQ(x1, x2, x3)	3	0.03135	0.010451	0.3586	0.7830624
Residuals	95	2.76901	0.029148		
Lack of fit	5	0.08953	0.017905	0.6014	0.6989489
Pure error	90	2.67949	0.029772		

The contour plots suggests that the optimal quantity for Dietary Iron, Calcium and Vitamin C required to have an optimal colour content in the soya milk is -3.035, -3.041 and 0.0148 coded unit respectively.

This result thus suggests that for optimal colour content in the soymilk, 3.07mg/100ml of dietary Iron, 154.1mg/100ml of Calcium and 24.23mg/100ml of Vitamin C should be used.

The surface plot on the other hand suggests that there is no significant curvature effect in the model.

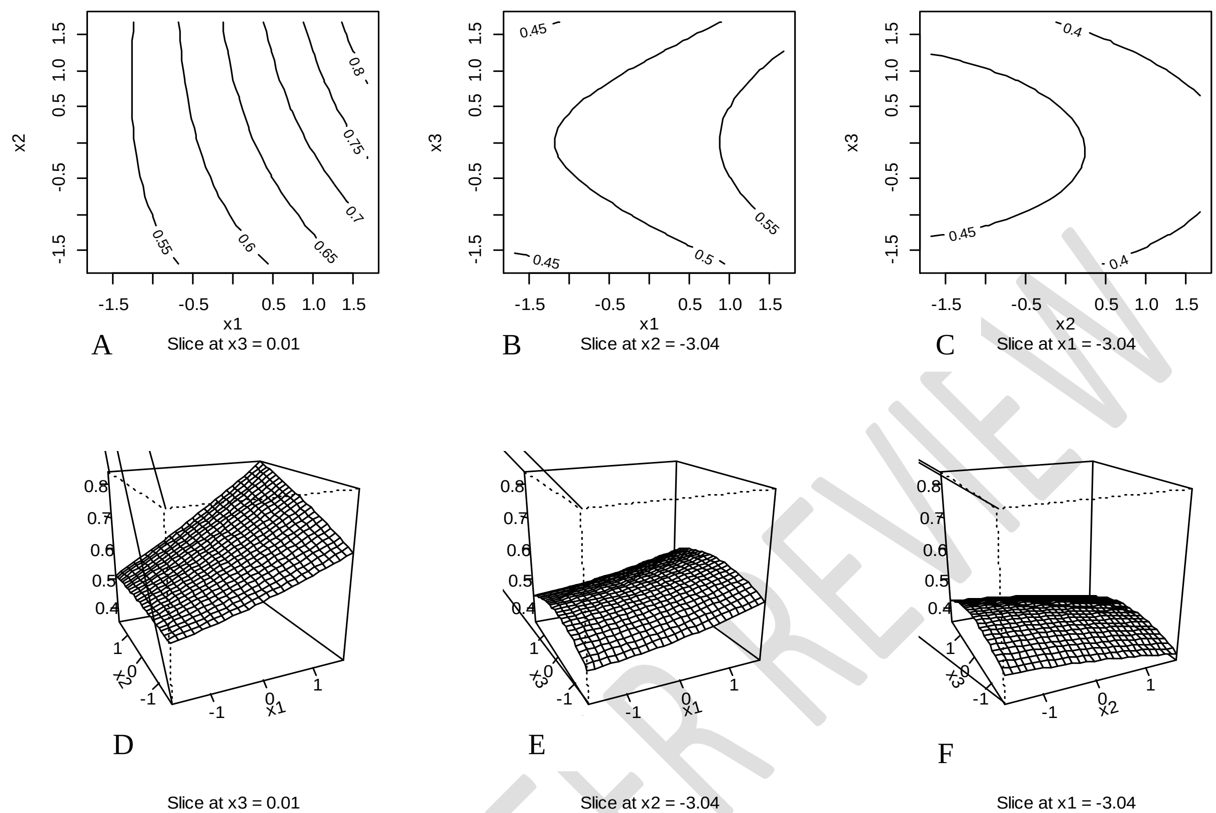


Figure 4.1: Contour and Response Surface plot for Colour content

4.1.2 ASH CONTENT OF SOYMILK

Following the same procedure as the previous, Table 4.3 shows the regression result of the model for getting the optimal ash content for soya milk.

The regression model and the analysis of variance allows to check if there are any significant effect of the linear, interaction and squared term. As the previous model, the model contains three linear effects, interaction effect and squared effect.

The p-value of the linear effect is less than 0.05, which suggest that amongst the three variables at least one has a significant effect. Checking them individual shows that the P-value of Calcium is less than 0.05. Therefore, there is a significant linear effect for Calcium. This means that we have enough evidence to conclude that the variation in the response (ash content of soya milk) depends on the variability on the quantity of Calcium.

From the p-value presented in the table, we have sufficient evidence to conclude that there is no significant interaction effect. That is, the effect of Iron on the ash content of soybean do not depend on either Calcium or Vitamin C.

Furthermore, the three squared effects presented in the model are not significant since the p-value for each squared effect is greater than 0.05 which is the significant level we chose. Thus, there is enough evidence for us to conclude that there is no curvature in the response surface. This can also be seen from the surface plot which is present in Figure 4.2. Since the square terms were identified as not significant in the analysis of variance, there is no need for looking at each of them individually.

Finally, the constant term was also significant in the model. The p-value for the lack of fit in the analysis of variance table is greater than 0.05. Thus, there is no evidence to that the model does not adequately explain the variation in the responses. The Adjusted R^2 suggests that only 9% variation of the response have been accounted for by the variation in explanatory variable. This is however considerable low, as such some important terms which affects the variation of the response can be included also.

The regression model thus is presented as:

$$\begin{aligned} Y = 0.0358 + 0.00008X_1 + 0.00927X_2 - 0.00008X_3 + 0.00002X_1X_2 + 0.0001X_1X_3 - 0.00002X_2X_3 \\ + 0.00008X_1^2 - 0.00005X_2^2 - 0.00028X_3^2 \end{aligned}$$

$where: X_1, X_2, X_3 =$ Iron, Calcium and Vitamin C respectively (4.2)

The contour plot and surface plot is presented in Figure 4.2 below.

TABLE 4.2 REGRESSION MODEL FOR ASH CONTENT

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	3.5825e-02	8.0072e-04	44.7405	<2e-16	***
x1	8.3194e-05	2.1792e-04	0.3818	0.7035	
x2	9.2780e-03	2.1792e-04	42.5743	<2e-16	***
x3	-8.4156e-05	2.1792e-04	-0.3862	0.7002	
$x_1:X_2$	1.7857e-05	2.8475e-04	0.0627	0.9501	
$x_1:X_3$	1.2500e-04	2.8475e-04	0.4390	0.6617	
$x2:X_3$	1.7857e-05	2.8475e-04	0.0627	0.9501	
x_1^2	7.7611e-05	3.2733e-04	0.2371	0.8131	
x_2^2	-4.8627e-05	3.2733e-04	-0.1486	0.8822	
x_3^2	-2.7585e-04	3.2733e-04	-0.8428	0.4015	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Multiple R-squared: 0.9503, Adjusted R-squared: 0.9455

F-statistic: 201.6 on 9 and 95 DF, p-value: < 2.2e-16

TABLE 4.3 ANALYSIS OF VARIANCE TABLE FOR ASH CONTENT

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
FO(x1, x2, x3)	3	0.0082314	0.00274379	604.2883	<2e-16
TWI(x1, x2, x3)	3	0.0000009	0.00000030	0.0669	0.9774
PQ(x1, x2, x3)	3	0.0000076	0.00000252	0.5560	0.6453
Residuals	95	0.0004314	0.00000454		
Lack of fit	5	0.0002948	0.00005896	38.8516	<2e-16
Pure error	90	0.0001366	0.00000152		

The contour plots suggests that the optimal quantity for Dietary Iron, Calcium and Vitamin C required to have an optimal ash content in the soymilk is 0.1118, 0.0203 and -0.6545 coded unit respectively.

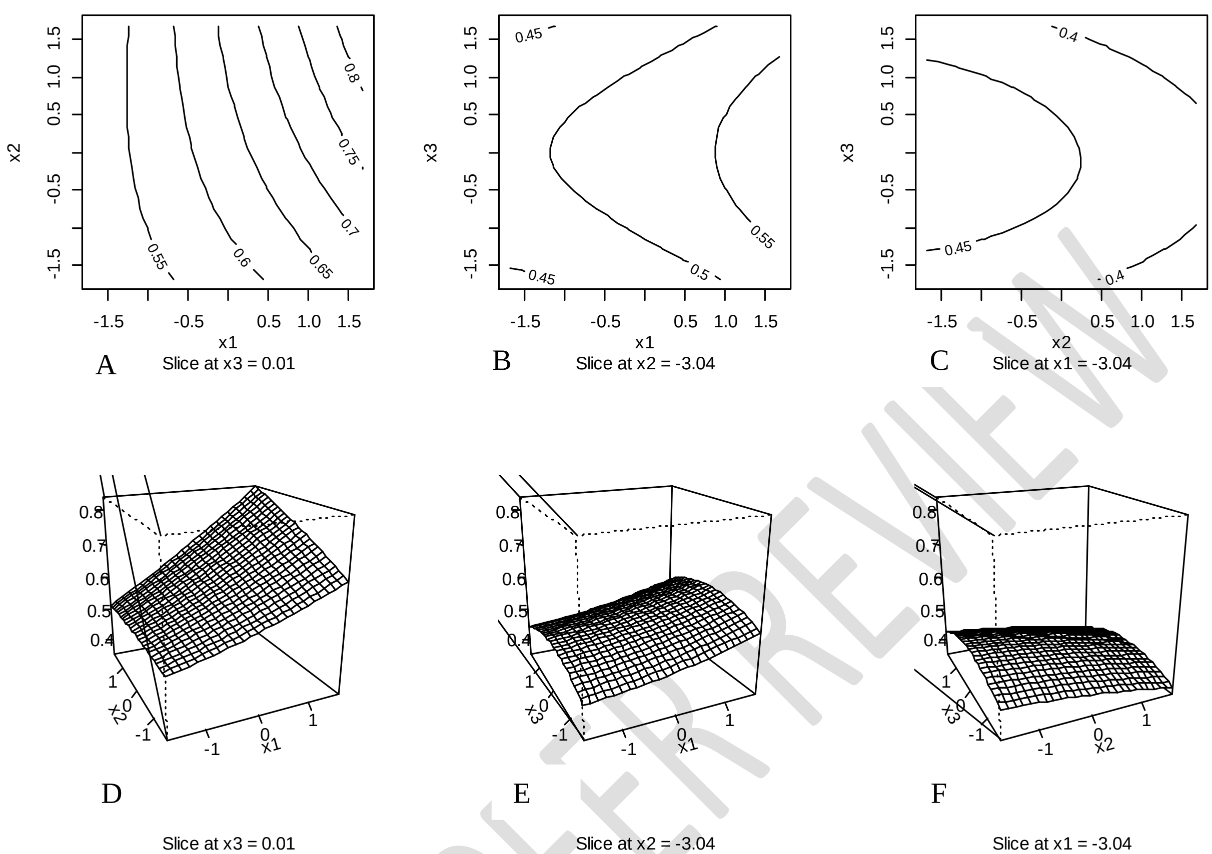


Figure 4.2: Contour and Response Surface plot for ash content

This result thus suggests that for optimal ash content in the soymilk, 2.22mg/100ml of dietary Iron, 152.03mg/100ml of Calcium and 13.53mg/100ml of Vitamin C should be used.

The surface plot on the other hand suggests that there is no significant curvature effect in the model.

4.1.3 MOSITURE CONTENT OF SOYA MILK

Following the same procedure as the previous, Table 4.5 shows the regression result of the model for getting the optimal moisture content for soymilk.

Again, the p-value of the linear effect is less than 0.05, which suggest that amongst the three variables at least one has a significant effect. Checking them individual shows that the P-value of Calcium is less than 0.05. Therefore, there is a significant linear effect for Calcium. This means that we have enough evidence to conclude that the variation in the response (moisture content of soya milk) depends on the variability on the quantity of Calcium.

From the p-value presented in the table, we have sufficient evidence to conclude that there is no significant interaction effect. That is, the effect of Iron on the moisture content of soya bean do not depend on either Calcium or Vitamin C.

Furthermore, the three squared effects presented in the model are not significant since the p-value for each squared effect is greater than 0.05 which is the significant level we chose. Thus, there is enough evidence for us to conclude that there is no curvature in the response surface. This can also be seen from the surface plot which is present in Figure 4.3. Since the square terms were identified as not significant in the analysis of variance, there is no need for looking at each of them individually.

Finally, the constant term was also significant in the model. The p-value for the lack of fit in the analysis of variance table is greater than 0.05. Thus, there is no evidence to that the model does not adequately explain the variation in the responses. The Adjusted R^2 suggests that only 9% variation of the response have been accounted for by the variation in explanatory variable. This is however considerable low, as such some important terms which affects the variation of the response can be included also.

The regression model thus is presented as:

$$\begin{aligned} Y = {} & 0.0358 + 0.00008X_1 + 0.00927X_2 - 0.00008X_3 + 0.00002X_1X_2 + 0.0001X_1X_3 - 0.00002X_2X_3 \\ & + 0.00008X_1^2 - 0.00005X_2^2 - 0.00028X_3^2 \end{aligned} \quad (4.3)$$

$where: X_1, X_2, X_3 =$ Iron, Calcium and Vitamin C respectively

The contour plot and surface plot is presented in Figure 4.2 below.

TABLE 4.4 SUMMARY FOR REGRESSION ANALYSIS FOR MOISTURE CONTENT.

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	95.6709049	0.2109203	453.5879	< 2.2e-16	***
x1	-0.0013597	0.0574040	-0.0237	0.9812	
x2	-0.4239059	0.0574040	-7.3846	5.827e-11	***
x3	-0.0013361	0.0574040	-0.0233	0.9815	
x_1:X_2	0.0016071	0.0750057	0.0214	0.9830	
x_1:X_3	-0.0044643	0.0750057	-0.0595	0.9527	
x2:X_3	-0.0055357	0.0750057	-0.0738	0.9413	
x_1^2	0.0532441	0.0862217	0.6175	0.5384	
x_2^2	0.1350463	0.0862217	1.5663	0.1206	
x_3^2	0.0540016	0.0862217	0.0862217	0.5326	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Multiple R-squared: 0.3762, Adjusted R-squared: 0.3171

F-statistic: 6.365 on 9 and 95 DF, p-value: 5.071e-07

TABLE 4.5 ANALYSIS OF VARIANCE TABLE FOR MOISTURE CONTENT

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
FO(x1, x2, x3)	3	17.1807	5.7269	18.1779	2.111e-09
TWI(x1, x2, x3)	3	0.0030	0.0010	0.0031	0.9998
PQ(x1, x2, x3)	3	0.8649	0.2883	0.9151	0.4367
Residuals	95	29.9296	0.3150		
Lack of fit	5	1.5406	0.3081	0.9768	0.4364
Pure error	90	28.3890	0.3154		

The contour plots suggests that the optimal quantity for Dietary Iron, Calcium and Vitamin C required to have an optimal moisture content in the soymilk is 0.0071, 1.5714 and 0.0926 coded unit respectively.

This result thus suggests that for optimal moisture content in the soymilk, 2.9858mg/100ml of dietary Iron, 335.71mg/100ml of Calcium and 25.48mg/100ml of Vitamin C should be used.

The surface plot on the other hand suggests that there is no significant curvature effect in the model.

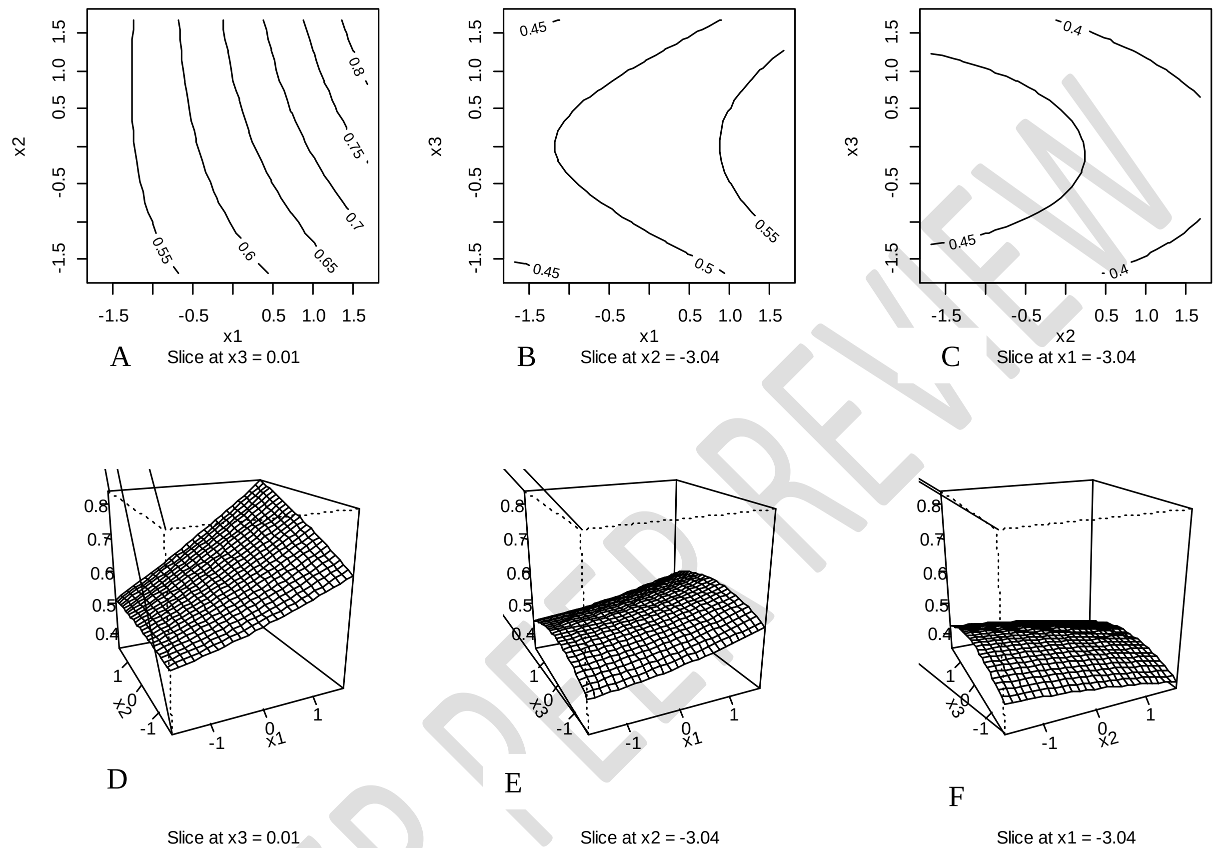


Figure 4.3: Contour and Surface plot for moisture content

The plot showing the normality plot as shown in the Appendix shows that the residual seem to follow a normal distribution.

CONCLUSION

The objective of this study is to determine the optimal combinations of levels of the different minerals (for soymilk) that is suitable for complementary feeding, and also to investigate the linear relationship as well as the curvature (quadratic) relationship using the response surface analysis.

As soon as we are close to the optimum a rotatable central composite design can be used to estimate the second-order response surface. The optimum on that response surface can be determined analytically.

Results in this study indicated that only the first-order process variables had significant ($P < |t|$).

The second-order process variable and the polynomial process variables show non-significance. Response surface analysis of the experimental data on colour content located maximum point for X1 at -3.04, X2 at -3.04 and X3 at 0.01, for the ash content, the maximum point for X1 was at 0.11, X2 at 0.02 and X3 at -0.65, while for the moisture content, the maximum point for X1 was at -0.01, X2 at 1.57 and X3 at 0.09.

RECOMMENDATIONS

In order to raise healthy infants using Soymilk as a complement to breast milk, the following recommendations are made;

(i) Colour content of Soymilk with the mixture compositionsX1at -3.04, X2 at -3.04 and X3 at 0.01 should be encouraged for feeding infants and young children.

(ii) Ash content of Soymilk with the mixture compositions X1 at 0.11, X2 at 0.02 and X3 a t -0.65 should be encouraged for feeding infants and young children.

(iii) Moisture content of Soymilk with the mixture compositions X1 at -0.01, X2 at1.57 and X3 at 0.09 should be encouraged for feeding infants and young children.

REFERENCES

Aguilera J. M. and Kosikowski F. V. (1976). Soybean Extruded Product: A Response Surface Analysis. *Journal of food science.* 41. 647-651.

Box G.E.P.& Wilson K.B. (1951). On the Experimental Attainment of Optimum Conditions. *Journal of Royal Statistical Society.* Series B, 13: 1-45.

Clarke, R. I. (1995). Micronutrient fortification of foods: Technology and quality control FAO *Corporate Doc Repository*. Retrieved from http//www.fao-org/docrep/w284oe/w284oeob.htp.

Conover W. (1999). *Practical Nonparametric Statistical*, 3rd edition, pp. 428-433 (6.1), John Wiley & Sons, Inc. New York.

Exon, D. (2007). *Good reasons why you may need Vitamin supplements*. Nourishing Traditions. In: The cook book that challenges politically correct Nutrition and the Diet Doctocrates by Sally Fallon and Mary Enig. New trends publishing Inc. Retrieved from; www.Newtrendspublishing.com.

Fallon, S. & Enig, M. G. (2007). Nourishing Tradition In: The Cookbook that Challenges Politically Correct Nutrition and the Diet Dictocrates. New Trends Publishing Inc. Retrieved from http://dspace.unijos.edu.ng/bitstream/10485/872/1/nutritional evaluation of cereal legume.pdf.

Graham H. (n.d). Research skill kruskal-wallis handout. Version 1.0, page 1

Gurinder S. and Roopa S. P. (2014). Optimization (Central Composite Design) and Validation of HPLC Method for Investigation of Emtricitabine Loaded Poly(lactic-co-glycolic acid) Nanoparticles: *In Vitro* Drug Release and *In Vivo* Pharmacokinetic Studies. *The Scientific World Journal.* Article ID 583090, 12 pages.

Iwe, M.O. (2003). *The science & Technology of soybean*. Published by Rojoint communication services Ltd, 27-262.

Keharom S., Mahachai R. & Chanthai S. (2016). The optimization study of α-amylase activity based on central composite design-response surface methodology by dinitrosalicylic acid method. *International Food Research Journal* 23, (1):10-17.

Laura C. O. and Eric C. O. (2013). Optimization of composite flour biscuits by mixture response surface methodology. *Food science and technol. International*. 19, (4): 343-350.

Lonnerdal, B. (1985). “Effects of maternal Dietary intake on Human milk Composition”. Retrieved from www.research Diets.com.

Li-Chun Z., Ying H., Xin D., Geng-Liang Y., Jian L. and Qian-Li T. (2012). Response Surface Modeling and Optimization of Accelerated Solvent Extraction of Four Lignans from Fructus Schisandrae. *Molecules17,* 3618-3629.

Mahsa A., MortezaB., Sirous N. and Abdolhosein N. (2012).A central composite design for the optimization of the remova lof the azo dye, methyl orange, from waste water using the Fenton reaction. *J. Serb. Chem. Soc. 77* (2). 235–246.

Marcin D., Mario D. and Terese L. (2015). Application of Central Composite Design for the study of NO_x Emission Performance of a Low NO_x Burner. *Energies.* 8, 3606-3627.

Meyers, Khuri and Carter (1989). Response Surface Methodology: *1966-1988. Technometrics* 31 (2): 137-153 http://www.jstor.org/.

Montgomery D. C. (2013). *Design and Analysis of Experiments.* 8th edition. John Wiley and Sons; NY, USA.

Murray R. Spiegel (1992). *Theory and problems of statistics*. (2nd Edition). M. C. Grow Hill Book.

Nsofor, L.M. Okpara, B.U. and Osuji, C.M. (1997). Tropical storage stability and chemical Properties of soymilk from sprouted soybeans. *J. Fd. Sci. Technol.* 34 (6) 377-384.

Okwunodulu, I. N. & Iwe, M. O. (2015). Micronutrient evaluation of fortified soymilk from Sprouted whole soybean for complementary feeding using response surface methodology. *Journal of Molecular Studies and Medicine Research,* 01(01), 16-25.

Onweluzo J.C and Nwakalor, C. (2009). Development and evaluation of vegetable milk from Treculia Africana (Decne) seeds. *Pakistan J. of Nutri.* 8, (3): 233-238.

Rarback, S. (2011). Nutrition and wellbeing A-Z. Infant Nutrition. Retrieved from http.//www.faqs.org/nutrition/na-irr/infant-nutrition. Html.

Seyyed H. P., Keyhani A., Mostofi-Sarkari M. R. and Jafari A. (2012). Application of response surface methodology for optimization of Picker-Husker harvesting losses in corn seed. *Iranica journal of Energy and environment*. 3, (2): 134-142.

Sharma H. K. Kaur J. Sarkar B. C. and Singh B. (2016). Optimization of pretreatment conditions of carrot to maximize juice recovery by response surface methodology. *Journal of Engineering science and Technology.* 1, (2): 158-165.

Solomon W. K. (2021). Application of Response Surface Methodology in Food Process Modelling and Optimization, DOI: 5772/Intechopen.100113.

Thiers, E.A. (2009). Water and nutrition. Published by Lenntech water treatment and purification Holding B. U. Retrieved from; *http//www.lenntech.com/about/aboutlenntech.en.htp.*

Thomas M. Little and F. J. Hills (1997). *Agricultural Experimentation*. John Wiley and Sons; USA.

Titus U. N. (2008).Growth performance of rats fed raw and extruded African breadfruit-based Complementary diets: A response surface analysis. *Journal of the Science of Food and Agriculture*.88:522–528.

Vipul S., Nishi B. & Swapan K. (2013). Application of response surface methodology and central composite design for the optimization of talc filler and retention aid in papermaking. *Indian Journal of Chemical technology*. 20, 121-127.

Wang H., Mustakas G., Wolt W., Wang C., Hesseltine W.& Bagley E.(1978). Soybean as the human food and simply processed. In *Parkistan Journal of Nutrition.* (14 (12): 898-906). www. R project.com

Yeung L.D. (2011).*Iron and complementary food fortification*. United Nations University.

Zhang X., Jinyan Z., Wen F. U. and Juan Z. (2010). Response surface methodology used for statistical optimization of Jiean-Peptide production by Bacillus Subtilis. *Electric journal of Biotechnology*.13 (4):3-8.

APPENDIX

TABLE 4.6

BOX-WILSON (1951) EXPERIMENTAL DESIGN MATRIX FOR CODED AND REAL INDEPENDENT VARIABLES.

Experimental Runs	**Coded independent process variables**			**Real independent process variables**		
	$\mathbf{X_1}$	$\mathbf{X_2}$	$\mathbf{X_3}$	**Fe(X_1)**	**Ca(X_2)**	**C(X_3)**
1	-1	-1	-1	2	100	16
2	-1	-1	1	2	100	32
3	-1	1	-1	2	200	16
4	-1	1	1	2	200	32
5	1	-1	-1	4	100	16
6	1	-1	1	4	100	32
7	1	1	-1	4	200	16
8	1	1	1	4	200	32
9	1.682	0	0	5	150	24
10	-1.682	0	0	1	150	24
11	0	1.682	0	3	250	24
12	0	-1..682	0	3	50	24
13	0	0	1.682	3	150	40
14	0	0	-1.682	3	150	8
15	0	0	0	3	150	24
16	0	0	0	3	150	24
17	0	0	0	3	150	24
18	0	0	0	3	150	24
19	0	0	0	3	150	24
20	0	0	0	3	150	24
21	0	0	0	3	150	24

X_1 (Fe) represents concentration (mg/100ml) of iron fortificant, X_2 (Ca) represents concentration (mg/100ml) of calcium fortificant and X_3 (Vit C) represents concentration (mg/100ml) of vitamin C fortificant used in the fortification trials. Each row represents a fortification trial run of adjustment levels of the process variable combination at one run. The experimental design has the upper limit of +1.682, intermediate limit of Zero (0) and a lower limit of -1.682 values of process variable combinations. The three process variables (Fe, Ca and Vitamin C) and the five experimental levels coded -1, -1.682, 0, 1 and 1.682 gave 15 variable combination which when replicated six times at center point (0) for estimation of error gave a total of 21 experimental runs shown in Table 4.0.

TABLE 4.7

RESPONSE VALUES OBTAINED ON MOISTURE CONTENT AND THEIR CORRESPONDING TOTAL SOLIDS FROM SOYMILK SAMPLES AT 2 WEEKS INTERVALS FOR 12 WEEKS.

Process variables			**Moisture Content (MC)**							**Row Average**	
X1	**X2**	**X3**	**T0**	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**	**%MC**	**%TS**
2	100	16	96.49	96.60	96.55	96.45	96.56	96.15	96.54	96.48	3.52
2	100	32	96.50	96.53	96.52	96.49	96.55	96.54	96.38	96.50	3.50
2	200	16	95.70	96.28	94.11	96.48	95.78	94.81	95.39	95.54	4.46
2	200	32	96.60	95.38	95.90	96.11	95.56	94.13	94.82	95.50	4.50
4	100	16	96.51	96.38	96.49	96.40	96.55	96.56	96.49	96.48	3.52
4	100	32	96.49	96.50	96.49	96.47	96.39	96.48	96.44	96.47	3.53
4	200	16	95.80	95.11	96.01	95.21	96.00	94.90	95.54	95.51	4.49
4	200	32	94.97	93.67	95.84	95.87	96.10	96.01	96.00	95.55	4.45
5	150	24	96.33	95.72	95.79	96.01	94.89	95.70	95.51	95.72	4.29
1	150	24	95.79	95.89	95.87	95.90	94.98	95.79	95.73	95.71	4.29
3	250	24	95.25	94.33	95.81	95.75	96.07	94.80	95.69	95.38	4.61
3	50	24	96.60	96.52	96.45	96.50	96.50	96.43	96.44	96.49	3.51
3	150	40	96.09	95.70	94.90	96.10	96.10	95.77	95.70	95.70	4.30
3	150	8	95.73	96.28	93.88	96.40	96.40	96.34	95.70	95.71	4.29
3	150	24	96.70	93.95	96.00	95.67	95.67	95.74	95.80	95.71	4.29

T_0 = MC of fortified soy milk samples obtained before storage. T_1, T_2, T_3, T_4, T_5 and T_6 each represents MC obtained at two weeks interval %TS represents the corresponding average of total solids obtained from the soymilk samples at two weeks intervals. X_1, X_2, and X_3 represent concentrations in (mg/100ml) of Fe, Ca and Vit. C.

TABLE 4.8

RESPONSE VALUES OF ASH CONTENT OF FORTIFIED SOYMILK SAMPLES OBTAINED AT 2 WEEKS INTERVALS FOR 12 WEEKS.

Process variables			Ash Content (g)							Row Ave.	
X1	X2	X3	T0	T1	T2	T3	T4	T5	T6		%Ash
2	100	16	0.028	0.028	0.027	0.026	0.027	0.026	0.028	0.0271	0.90
2	100	32	0.028	0.027	0.027	0.026	0.026	0.027	0.026	0.0179	0.89
2	200	16	0.043	0.043	0.041	0.044	0.042	0.044	0.043	0.0431	1.44
2	200	32	0.042	0.043	0.043	0.041	0.043	0.042	0.044	0.0426	1.42
4	100	16	0.028	0.029	0.028	0.027	0.029	0.021	0.029	0.0273	0.60
4	100	32	0.027	0.029	0.026	0.027	0.028	0.026	0.029	0.0274	0.60
4	200	16	0.043	0.044	0.042	0.044	0.042	0.043	0.044	0.0431	1.44
4	200	32	0.044	0.042	0.043	0.042	0.045	0.042	0.045	0.043	1.43
5	150	24	0.037	0.037	0.036	0.037	0.036	0.036	0.037	0.0366	1.22
1	150	24	0.037	0.036	0.038	0.037	0.038	0.036	0.037	0.037	1.23
3	250	24	0.056	0.055	0.056	0.054	0.055	0.056	0.055	0.055	1.83
3	50	24	0.018	0.018	0.020	0.018	0.017	0.015	0.017	0.017	0.59
3	150	40	0.037	0.035	0.034	0.035	0.036	0.037	0.035	0.0356	1.19
3	150	8	0.036	0.035	0.037	0.036	0.037	0.035	0.036	0.036	1.2
3	150	24	0.036	0.036	0.035	0.036	0.034	0.037	0.035	0.0356	1.19

T_0 represent ash content of fortified soy milk samples obtained before storage. T_1, T_2, T_3, T_4, T_5 and T_6 each represents ash content obtained at two weeks intervals for 12 weeks ambient storage. X_1, X_2, and X_3 represent concentrations in (mg/100ml) of Fe, Ca and Vit. C. in each trial run of fortified soymilk samples.

TABLE 4.9

RESPONSE VALUES OF COLOUR OF FORTIFIED SOYMILK SAMPLES OBTAINED AT 2 WEEKS INTERVALS FOR 12 WEEKS

Process variables			Colour (absorbance at 620nm)							
X1	X2	X3	T0	T1	T2	T3	T4	T5	T6	Row Ave.
2	100	16	0.669	0.741	0.549	0.481	0.409	0.368	0.325	0.506
2	100	32	0.665	0.739	0.551	0.479	0.410	0.369	0.322	0.505
2	200	16	0.674	0.735	0.550	0.509	0.409	0.370	0.341	0.513
2	200	32	0.673	0.734	0.552	0.479	0.410	0.378	0.347	0.510
4	100	16	0.698	0.928	0.917	0.840	0.461	0.443	0.223	0.644
4	100	32	0.703	0.835	0.821	0.685	0.524	0.512	0.452	0.647
4	200	16	0.793	0.865	0.860	0.766	0.739	0.501	0.476	0.714
4	200	32	0.793	0.866	0.860	0.780	0.560	0.551	0.545	0.708
5	150	24	0.796	0.871	0.806	0.801	0.761	0.690	0.559	0.763
1	150	24	0.749	0.784	0.738	0.682	0.406	0.383	0.355	0.585
3	250	24	0.792	0.865	0.860	0.774	0.561	0.549	0.546	0.706
3	50	24	0.798	0.881	0.718	0.601	0.394	0.390	0.353	0.591
3	150	40	0.694	0.835	0.699	0.670	0.502	0.399	0.357	0.594
3	150	8	0.664	0.827	0.761	0.681	0.590	0.487	0.323	0.619
3	150	24	0.665	0.831	0.770	0.670	0.601	0.490	0.320	0.619

T_0 represent the absorbance (colour) of fortified soy milk samples obtained before storage. T_1 to T_6

Each represents absorbance (colour) of fortified sample at 2 weeks interval for 12 weeks. Each row represents

A combination of three independent process variables.

Diagnostics

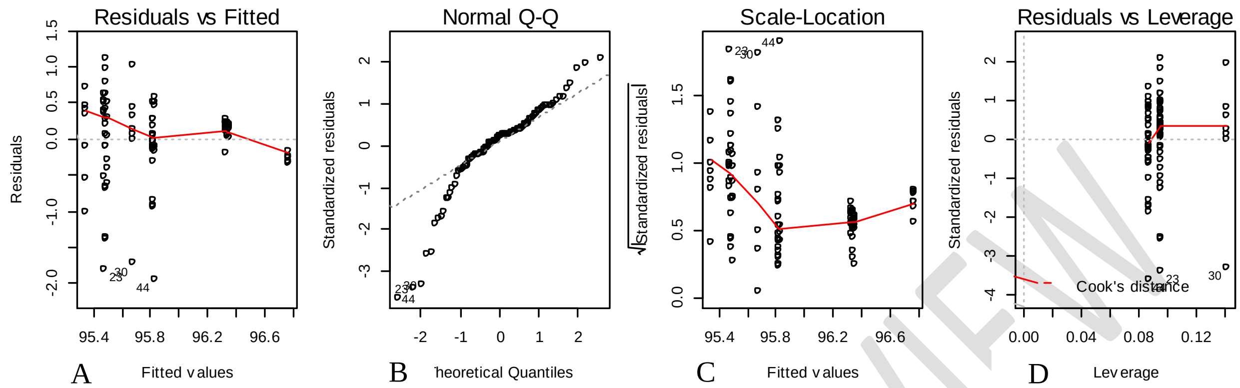


FIGURE 4.4 DIAGNOSTICS PLOTS OF THE MOISTURE CONTENT

Diagnostics

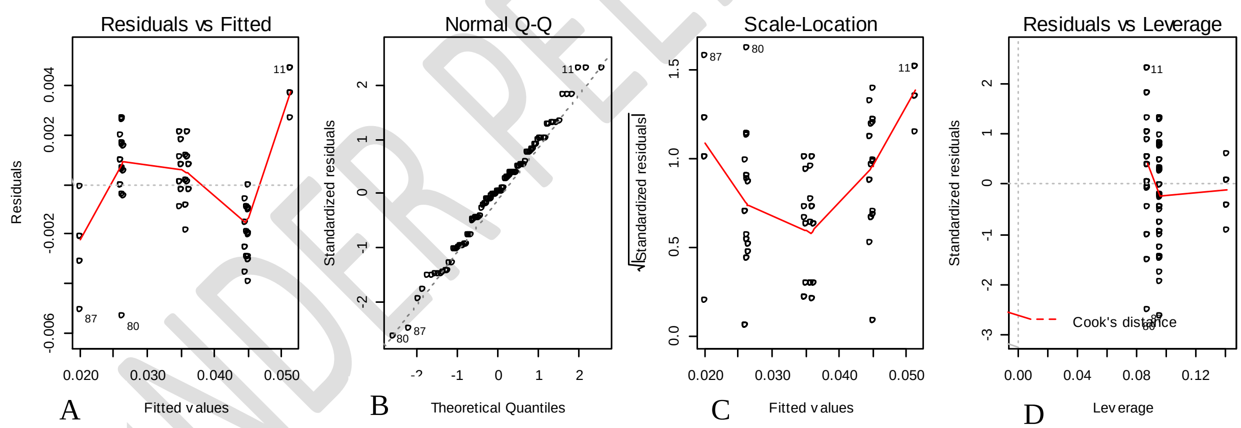


FIGURE 4.5 DIAGNOSTICS PLOTS OF THE ASH CONTENT