Split-plot Central Composite Experimental Design Method for Optimization of Cake Height to Achieve desired Texture

*1Y. Yakubu, 1Z. Q. Aliyu, 1A. Usman and 2P. O. Evans
1Department of Statistics, Federal University of Technology, Minna, Nigeria
2Department of Mathematics and Statistics, The Federal Polytechnic, Bida, Nigeria
[*Corresponding Author: E-mail: yisa.yakubu@futminna.edu.ng]

ABSTRACT

In many industrial experimental situations, the levels of certain factors under investigation are much harder to change than others due to time and/or cost constraints. An appropriate approach to such situations is to restrict the randomization of the hard-to-change (HTC) factors, which leads to a split-plot structure. This work designs and conducts a split-plot central composite experiment for optimizing cake height using oven temperature(Factor *A*) as the HTC factor, amount of flour *(B)*, baking powder *(C)*, and amount of milk *(D)* as the easy-to-change (ETC) factors. A second-order split-plot central composite design (CCD) model was fit to the generated data and analyzed using generalized least squares (GLS). A stationary point, which gives optimum cake height, was then determined. The results show that main effects of oven temperature, flour, baking powder, and milk were highly significant on the cake height $(P < 0.01)$. Their quadratic effects were also significant except that of the flour. The flower/baking powder interaction effect was significant. The fitted model accounted for about 95% of the total variability in the cake height data. The observed optimum cake height was $\hat{y} = 11.047 cm$ at a stationary point: $A = 250^o C, \; B = 180.75 gm, C = 6.75 gm, and \; D = 194.06 gm$. This study has established the potentials of response surface experiments in optimizing products in food industries.

Keywords: Experiment, split-plot CCD, Cake height, Design, Stationary point.

INTRODUCTION

Many quality improvement projects require some form of experimentation on a process. Experimentation allows an investigator to examine the effect of purposeful changes of the settings of the input variables in a system, on the output or response. It covers a wide range of applications from household activities like food preparation to technological innovation in material science, life science, semiconductors, etc (Jeff and Hamada, 2009).

The main task of every food industry is the production of health - supporting food of high quality on modern and effective technological and economic conditions, which have to consider also demands of ecology and food security (Romisch, 2005). Optimization of a product is an effective strategy of accomplishing its successful development and, for food processes, and especially for food development, it is a way to obtain ideal conditions to achieve a desired quality (physicochemical, chemical and sensory, for example) (Myers *et al*., 2009). If a food cannot

be re-engineered or modified to fulfill consumer specifications, it will not succeed in the market. Thus a bakery scientist, for example, will always be interested in determining the settings for certain process variables to optimize a critical quality characteristic, such as nutritional qualities (like 'fluffy' texture of a cake, which depends strongly on its height), anti-nutritional or toxic levels, and shelf life etc., of the resulting product.

In developing or optimizing processes, many companies use statistical approaches, such as response surface methodology (RSM) in their research departments in order to achieve the best settings of the factors that will yield the best characteristic of a product and/or process response. In food development or unit operations, RSM has important applications in the design, analysis and determination of optimum settings of the control variables that result in a maximum (or a minimum) response over a certain region of interest (Nwabueze, 2010).Such determination requires having a 'good' fitting model that provides an adequate

representation of the mean response, because such a model is used to determine the value of the optimum (Khuri and Mukhopadhyay, 2010).

A lot of studies have been conducted that used RSM for process and product optimization (Elbert *et al.*, 2001; Iwe *et al.,* 2004; Wang *et al.,* 2008;Kliemann*et al.*, 2009; Ellendersen *et al.* 2012 and Granato *et al.* 2014). However, much of these studies focused on complete randomization of the levels of the experimental factors. For instance, Granato *et al.* (2014) did not mention what to do in instances where restriction on randomization is required. Therefore, even though the field of food science and engineering has seen a growth in the use of statistical tools, the applications often are misused or misrepresented (Nunes *et al.,* 2015).

Most industrial experiments involve two types of factors, some with levels that are easy to change (which can be completely randomized) and one or more with levels that are difficult or costly to change (which cannot be completely randomized). The hard-to-change factors place restriction on randomization of its levels and this result in split-plot experiment where the experimental runs are performed in groups, and, in a group, the levels of the HTC factors are not reset from run to run. Research works on impact of a split-plot structure on response surface designs began in the 1990s including pioneer author like Letsinger *et al.* (1996) and others (Box,1996; Trinca and Gilmour, 2000; Vining *et al*, 2005; Kowalski *et al.*, 2006 and Bradley and Christopher, 2009).

This study seeks to model and optimize cake height for the consumers to achieve desired texture using a statistical experiment that involves key cake-baking materials (oven temperature, amount of flour, baking powder, and amount of milk) as factors. The oven temperature is a hard-to-change factor. Therefore, central composite design (CCD) experiment with a split-plot structure was conducted whereby the experimental runs are grouped by temperature, rather than waiting for the oven to heat up and cool down all of the time.

Statistical Model and Notations
As in all experiments with a split-plot structure, the observations from this experiment are correlated and this violates the assumption of independence of the ordinary least squares (OLS) estimator. Therefore the generalized least squares (GLS) model below, which accounts for the correlation between the observations, was fitted (Vining *et al.,* 2005).
$$y = X\beta + Z\gamma + \varepsilon \qquad (1)$$

where **y** is the N x 1 vector of responses, **X** is the N x p overall model matrix, **β** is the p x 1 vector of regression coefficients, **Z** is an N x b incidence matrix assigning observations to each of the b whole plots; γ is the N x 1 vector of whole-plot error terms, ε is the N x 1 vector of subplot error terms. It is assumed that $\gamma_i \sim N(0, \sigma_\gamma^2)$, $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$, $\text{cov}(\gamma_i, \varepsilon_{ij}) = 0$.
The variance - covariance matrix for the observation vector **y** is:
$$\text{Var}(y) = V = \sigma_\varepsilon^2 I_n + \sigma_\gamma^2 ZZ'$$

$$= \sigma_\varepsilon^2(I_n + dZZ') \qquad (2)$$

Where $d = \frac{\sigma_\gamma^2}{\sigma_\varepsilon^2}$ gives the relative magnitude of the two variance components. The matrix $\mathbf{ZZ'}$ is a block diagonal matrix with diagonal matrices of $J_{n1}, J_{n2}, ..., J_{nz}$, where J_{ni} is an n_i x n_i matrix of 1's and $n_i$is the number of observations in the *i*th whole-plot.

The diagonal elements of V are the variances of the responses and the off- diagonal elements are the covariance between pairs of responses. The non-zero off-diagonal elements correspond to pairs of responses from within a given whole plot while the zero off-diagonal elements correspond to pairs of runs from two different whole plots.

The generalized least squares (GLS) estimators of the model coefficients (Vining *et al,* 2005) are

$$\hat{\beta}_{GLS} = (X'V^{-1}X)^{-1}X'V^{-1}y \qquad (3)$$

Thus,
$$\text{Var}(\hat{\beta}_{GLS}) = (X'V^{-1}X)^{-1} \qquad (4)$$

and
$$\hat{y} = X(X'V^{-1}X)^{-1}X'V^{-1}y = Hy \qquad (5)$$

Where H is the 'hat' matrix. Equation (1.2) accounts for covariance among the responses, such as might be present in a mixed effects model.

MATERIALS AND METHODS
The key materials used for cake-baking in this work include flour (factor B), baking powder (factor C), milk (factor D), which were measured in grams (g), and oven temperature (factor A), measured in degree centigrade (°C), with fixed amount of other necessary ingredients (butter, eggs, sugar and salt, etc.) present. Oven temperature was included because the temperature at which the batches are baked strongly impacts cake height. It is a hard-to-change (HTC) factor since its levels cannot be easily randomized and thus the whole-plot factor, while the other three are easy-to-change (ETC) and thus were the subplot factors. A four-factor central composite experiment with a split-plot structure in the given variables was designed and conducted. The experimental runs were grouped by temperature (rather than waiting for the oven to heat up and cool down all of the time) and all runs with the same temperature level were executed simultaneously in one furnace. The factorial and axial parts were replicated only once and each of the factors consists of five (5) levels, which were coded as ($\pm 1, 0, \pm \alpha$) using:

$$x_i = \frac{2X_i - (X_{il} + X_{ih})}{X_{ih} - X_{il}} \qquad (6)$$

Where x_i is the coded level of the *i*th factor , $i = 1, 2, \dots, k$; X_i is the actual (measured) value of the *i*th factor; X_{il} and X_{ih} are, respectively, the actual low and high levels of the *i*th factor. The choice of factors, levels and factor ranges are listed in Table1.

Performing the Experiment
This experimental design consists of four categories of points, which include the factorial portion (f), the whole-plot axial point with a distance, α, from the design center, the subplot axial point with distance, β, from the design center, and the center (c) points, as given in the design matrix in Table 2.For every cake, the butter and sugar were first mixed together in the same Kitchen-aid mixer on the same speed setting. The flour, baking powder and salt were sifted together and added with the specified amount of milk and mixed at the same speed. This was then poured into the cake pan and a spatula was used to scrape all into the pan. All cakes were baked from the same batch of ingredients and the same cake pans were used for each run.

At each of the factorial points involving the HTC factor (oven temperature (*A*)), the subplot runs are the factorial points in the ETC factors (*B, C* and *D*). At each of the two whole-plot axial points (i.e., $z_1 = 150°C, 250°C$), eight replicates of the center of the subplot factors (i.e., $x_1 = 120.5g, x_2 = 4.5g, x_3 = 129.37g$) were run. Two replicates of the whole-plot center points ($z_1 = 200 °C$) were run, one consisting of the six (6) axial points in the ETC factors [$x_1 = (60.25g, 180.75g), x_2 = (2.25g, 6.75g), x_3 = (64.68g, 194.06g)$] while the other contains four(4) replicates of the center of the subplot factors ($x_1 = 120.5g, x_2 = 4.5g, x_3 = 129.37g$).Thus each of the whole-plot factorial and axial portions of the design is of size eight while the first and second replicates of the whole-plot center portion are, respectively, of size six and four. Thus the resulting split-plot central composite design (CCD) has $n_f = 16$ factorial points, $n_\alpha = 16$ whole-plot axial points, $n_\beta = 6$ subplot axial points and $n_c = 4$ center points. That is, there were $N = 42$ cake mix batches but only 6 operations of the oven as given in Table 2.

Table1: Original and Coded Levels of Factors used in the Experiment

Factor	coded levels				
	-1	-1	0	1	1
Amount of Flour(X_1)	60.25g	60.25g	120.5g	180.75g	180.75g
Amt of Baking Powder(X_2)	2.25g	2.25g	4.5g	6.75g	6.75g
Amount of Milk(X_3)	64.68g	64.68g	129.37g	194.06g	194.06g
Temperature of Oven(z_1)	150°C	150°C	200°C	250°C	250°C

Table 2: Design Matrix for the D (1,3) split-plot Central Composite Experiment

Run order	Whole plot	Oven Temperature (°C) Z_1	Amount of Flour (g) X_1	Baking Powder (g) X_2	Amount of Milk(g) X_3	category
1	1	150	60.25	2.25	64.68	factorial
2		150	180.75	2.25	64.68	(f)
3		150	60.25	6.75	64.68	
4		150	180.75	6.75	64.68	
5		150	60.25	2.25	194.06	
6		150	180.75	2.25	194.06	
7		150	60.25	6.75	194.06	
8		150	180.75	6.75	194.06	
9	2	250	60.25	2.25	64.68	factorial
10		250	180.75	2.25	64.68	(f)
11		250	60.25	6.75	64.68	
12		250	180.75	6.75	64.68	
13		250	60.25	2.25	194.06	
14		250	180.75	2.25	194.06	
15		250	60.25	6.75	194.06	
16		250	180.75	6.75	194.06	
17	3	150	120.5	4.5	129.37	wp axial
18		150	120.5	4.5	129.37	$(-\alpha)$
19		150	120.5	4.5	129.37	
20		150	120.5	4.5	129.37	
21		150	120.5	4.5	129.37	
22		150	120.5	4.5	129.37	
23		150	120.5	4.5	129.37	
24		150	120.5	4.5	129.37	
25	4	250	120.5	4.5	129.37	wp axial
26		250	120.5	4.5	129.37	$(+\alpha)$
27		250	120.5	4.5	129.37	
28		250	120.5	4.5	129.37	
29		250	120.5	4.5	129.37	
30		250	120.5	4.5	129.37	
31		250	120.5	4.5	129.37	
32		250	120.5	4.5	129.37	
33	5	200	60.25	4.5	129.37	sp axial
34		200	180.75	4.5	129.37	$(\pm\beta)$
35		200	120.5	2.25	129.37	
36		200	120.5	6.75	129.37	

37		200	120.5	4.5	64.68	
38		200	120.5	4.5	194.06	
39	6	200	120.5	4.5	129.37	Center
40		200	120.5	4.5	129.37	(c)
41		200	120.5	4.5	129.37	
42		200	120.5	4.5	129.37	

The oven was big enough to accommodate eight (8) pans of the cake at a time and thus at each randomly-selected setting of the oven temperature factor, all possible combinations of the levels of the other (ETC) factors (flour, baking powder and milk) were baked at the same time. One run of the oven (i.e., one oven temperature setting) was performed in a day and the experiment took six (6) days to complete. This enables the oven temperature to stabilize (i.e., cool down to 0°C) after each run before another level was randomly selected for another run in another day. All cakes were baked for 28 minutes to minimize baking time as a nuisance factor. All cakes measurements were taken in centimeters (cm).

The design consists of two different randomization structure:- temperature (HTC) factor-levels were randomly and independently assigned to the whole-plots; within each whole plot, the ETC factor level combinations were randomly and independently assigned to the subplots using a different randomization technique. This was achieved with the aid of the Design-Expert (version 12) statistical package, which was used to create the design layout in a fully-randomized order of runs of the experiment. Thus levels of the whole-plot factor were not reset for each run of the subplot factors and this leads to two error terms for effects comparison, one for the whole-plot treatments (σ_γ^2), and one for the subplot treatments (σ_ε^2), as well as the interaction between whole-plot treatments and subplot treatments.

Data Analysis
The data analysis here was a form of maximum likelihood (ML) estimation, more specifically, restricted maximum likelihood (REML) estimation. The REML is used here to estimate whole-plot and subplot variance components (σ_γ^2 and σ_ε^2, respectively) based on maximum likelihood (ML) estimation of the residual error distributions. Then the estimated error variances are plugged into the Generalized Least Squares (GLS) estimator (equation 1.2), which was used to estimate the factor effects and their corresponding p-values. This was achieved with the aid of Design Expert (version 12) statistical package by maximizing the restricted log likelihood function:

$$l_g = -\frac{1}{2}\log|V| - \frac{1}{2}\log|X'V^{-1}X| - \frac{1}{2}r'V^{-1}r - \frac{n-p}{2}\log(2\pi) \qquad (7)$$

Where X and V are as defined in section 1.1 above; $r = y - X(X'V^{-1}X)^{-1}X'V^{-1}y$ and p is the number of parameters in $\boldsymbol{\beta}$. The standard errors for the factor effects were then computed as the square root of the diagonal elements of the covariance matrix $(X'V^{-1}X)^{-1}$. Optimization facility of the statistical package was then used to obtain the set of levels of the factors that yields optimum value of the cake height (the stationary point).

RESULTS AND DISCUSSION
The REstricted Maximum Likelihood (REML) analysis results are given in Table3 with Kenward-Roger p-values. We observed from this table that variance terms are grouped into two sections: a Whole-plot section for the HTC factors and a subplot section for ETC factors. This result is consistent with the findings of Letsinger *et al.* (1996), Hand & Crowder (1996) and Little *et al* (1996) in their experiments. The p-value (Prob>F) for the whole-plot was observed to be highly significant with a value much less than the generally acceptable 0.05 alpha level. That is, for all the terms making up the whole-plot (HTC) portion of the model. It was also observed that the p-values for the individual whole-plot terms were each

significant with values much smaller than 0.05.This indicates that the oven temperature (factor A) has significant effect on the height of the cake. Also, the quadratic effect of this factor on the cake height is highly significant. Table 3 also showed the subplot terms as a whole were observed to be highly significant, though some individual subplot terms (AB, AC, AD, BD, CD, and B^2) were not. The main effects of factors B, C, and D (amount of flour, amount of baking powder and amount of milk), the interaction effect of the quantity of flour and amount of baking powder (BC) and the quadratic effects of C and D (C^2 and D^2) were each observed to be highly significant on cake height (with p-values far less than 0.05).

The group (whole-plot) and residual (subplot) variance components were considered next. These are variations that were not explained by the model in terms of the factors, as given in Table 4.

The group (or whole-plot error) variance is due to resetting of a hard-to-change factor level (oven temperature). The computed group variance was zero, and this signifies that the whole-plot model explained all of the variation between the whole plots. The residual (or subplot error) variance is due to each of the subplot runs. In this work, the residual variance was 0.38, indicating that the subplot model accounted for about 62% of the variation between the subplots. Then we looked at the computed regression coefficients and variance inflation factors (VIF) as given in Table5. The VIF measures how much the variance of the model is inflated by lack of orthogonality in the design and indicates the extent to which multi-collinearity is present in a regression analysis. Similar results were obtained by Vining *et al.* (2005), in their experiment on strength of ceramic pipe.

Table 3: REstricted Maximum Likelihood (REML) analysis with Kenward-Roger p-values

Source	Term df	Error df	F-value	p-value	
Whole-plot	2	27	11	0.0003	significant
a-A Oven Temperature	1	27	8.14	0.0082	
a^2	1	27	13.87	0.0009	
Subplot	12	27	42.19	< 0.0001	significant
B-B Quantity of flour	1	27	355.82	< 0.0001	
C-C Amount of baking powder	1	27	10.92	0.0027	
D-D Amount of milk	1	27	17.86	0.0002	
aB	1	27	2.27	0.1432	
aC	1	27	2.03	0.1653	
aD	1	27	0.0149	0.9036	
BC	1	27	8.37	0.0075	
BD	1	27	0.0814	0.7776	
CD	1	27	0.0415	0.8401	
B^2	1	27	0.7063	0.4081	
C^2	1	27	6.97	0.0136	
D^2	1	27	6.97	0.0136	

Table 4: Whole-plot and Subplot Variance Components

Source	Variance	StdErr	95% CI Low	95% CI High
Group	0	0	0	0
Residual	0.38	0.1	0.24	0.7
Total	0.38			

Table 5: Estimates, Standard Errors and Variance Inflation Factors (VIF) of Parameters of the Fitted Split-plot CCD Model

Source	Coefficient Estimate	Standard Error	VIF
Intercept	5.73	0.1987	
Whole-plot Terms:			
a-A Oven Temperature	-0.3094	0.1085	1
a²	-0.8614	0.2313	1.08
Subplot Terms:			
B-B Quantity of flour	2.73	0.1446	1
C-C Amount of baking powder	0.4778	0.1446	1
D-D Amount of milk	0.6111	0.1446	1
aB	0.2313	0.1534	1
aC	-0.2188	0.1534	1
aD	0.0188	0.1534	1
BC	0.4438	0.1534	1
BD	-0.0437	0.1534	1
CD	0.0312	0.1534	1
B²	0.3036	0.3613	3.57
C²	0.9536	0.3613	3.57
D²	0.9536	0.3613	3.57

The VIFs values (Table 5) are equal to one except for the last three terms of the model (B², C² and D²) which indicates the predictors had non correlated orthogonal relationship with all the other predictors in the model. Each of the last three predictors had VIF values of 3.57, which indicates that it was moderately correlated with all the other predictors. Thus there was no case of multicollinearity in the data, and we have the fitted second-order model as:

$$Height = 5.73 - 0.31z_1 + 2.73x_1 + 0.48x_2 + 0.61x_3 + 0.23z_1x_1 - 0.22z_1x_2 + 0.019z_1x_3 + 0.44x_1x_2 - 0.044x_1x_3 + 0.031x_2x_3 - 0.86z_1^2 + 0.30x_1^2 + 0.95x_2^2 + 0.95x_3^2 \qquad (3.1)$$

Each of the terms in this model has the same p-value as in Table 3. Thus the effects of the terms: a, B, C, D, BC, a², C², and D² were each significant while that of the remaining terms were not significant. The model summary is given in Table6, this fitted model explains 95% of the total variability in the cake height ($R^2 = 0.95$). Thus only 5% variability was not accounted for by the model, also, the adjusted R^2 was 0.92 ($Adj. R^2 = 0.92$). These statistics signify that the fitted model was good, that is, the model captures most of the variation in the data.

Diagnostic plots
Here we diagnose the statistical properties of the above fitted model. Figure 1 gives the normal probability plot of the residuals, which indicates whether the residuals follow a normal distribution. From this plot, nearly all the points fall on the straight line. This indicates that, to some extent, the residuals were normally distributed.

Table 6: Explanatory Power of the Fitted Model (3.1)

Std. Dev.	0.6135	R²	0.9502
Mean	6.02	Adjusted R²	0.9183
C.V. %	10.19		

Thus there are no problems with our data. Next the plot of the residuals versus the predicted values was considered, as given in Figure 2. This plot is a visual check for the assumption of constant variance. As can be directly seen, this plot is a random scatter with a consistent top to bottom range of residuals across the predictions on the X axis. Thus we can conclude here that the model satisfied the constant variance assumption. Next we looked at the plot of the residuals versus the experimental run order, as presented in Figure 3. This plot provides a check for lurking variables that may have influenced the response during the experiment. As can be directly observed, the plot showed a random scatter without any trend. Thus there were no any lurking variable in the background. The plot of the predicted values versus the actual values was examined, as given in Figure 4. The plot help us to detect observations that were not well predicted by the model and shows that the data points were evenly split by the 45 degree line.
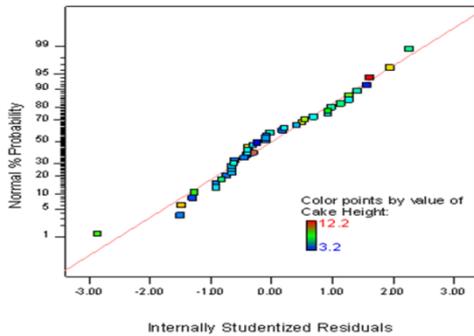

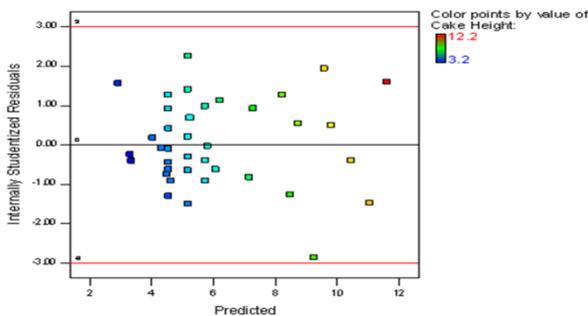**Figure 1**: Normal probability plot of the residuals for the cake height data

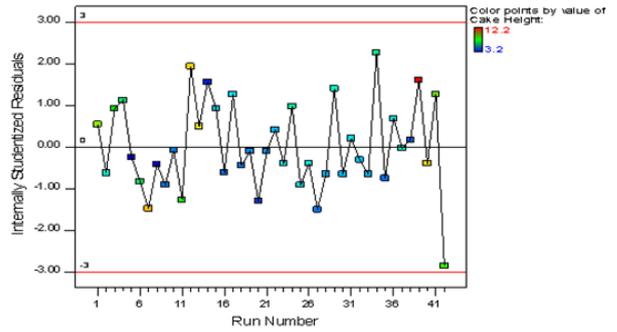**Figure 2**: Plot of residuals versus fitted values for the cake height data

Figure 3: Plot of residuals versus experimental run order for the cake height data
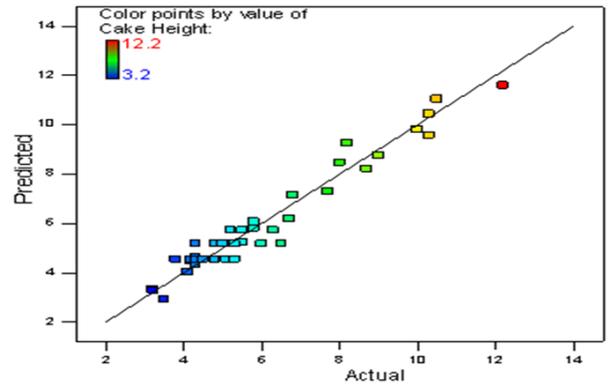

Figure 4: Plot of fitted values versus actual values for the cake height data

Lastly numerical optimization facility of the Design Expert package was used to see what combination of levels of the factors A, B, C, and D yields optimum cake height. This gives the results in Figure 5 from where we observed that the optimum cake height occurs at a factor setting, which consists of high level of each of the four factors. That is, an optimum cake height of 11.047 occurred at 250°C of oven temperature, 180.75 g of flour, 6.75 g of baking powder, and 194.06 g of milk. The package conducted the numerical optimization with these high factor levels expressed in their coded (-1, +1) values (equation 6) as given in Figure 5 below. This is the stationary point since all of the other possible solutions that can be explored through this package gave less cake height.
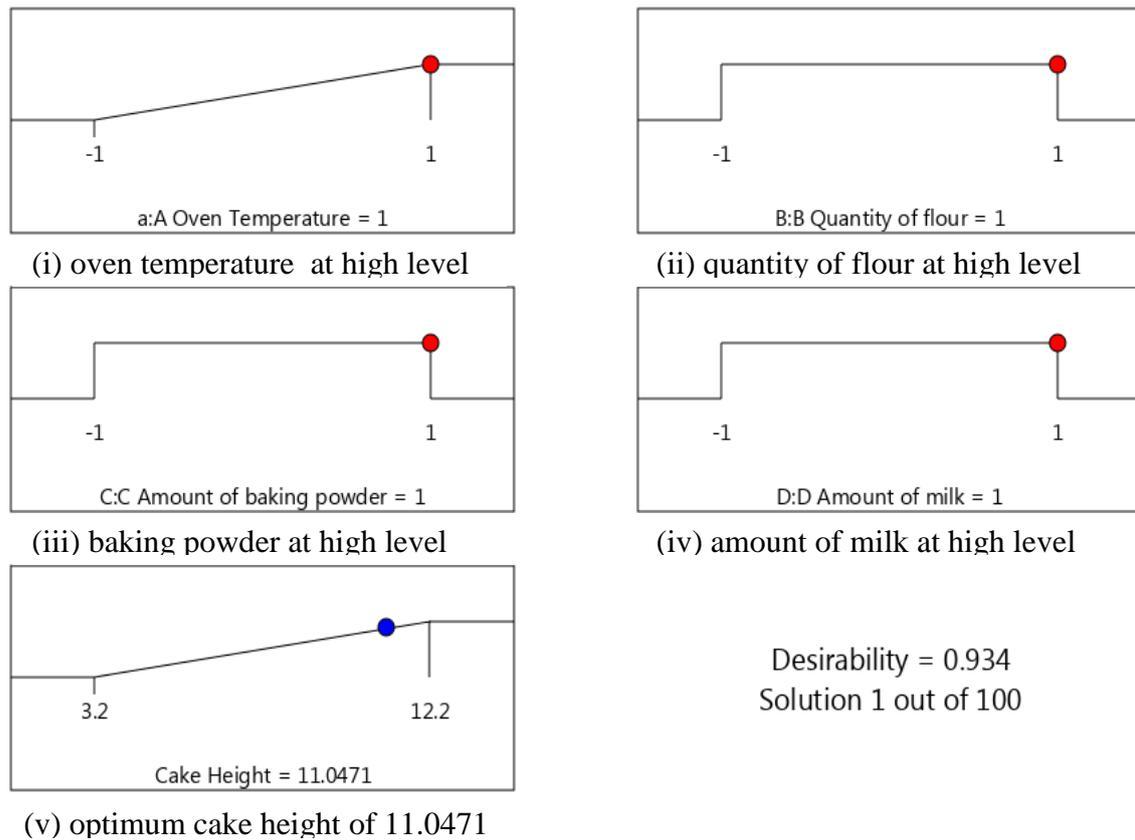
(i) oven temperature at high level

(ii) quantity of flour at high level

(iii) baking powder at high level

(iv) amount of milk at high level

(v) optimum cake height of 11.0471

Desirability = 0.934
Solution 1 out of 100

Figure 5: Numerical optimization results showing optimum cake height occuring at a factor setting, which consists of high level of each of the factors.

CONCLUSION

This study revealed that cake height and its texture is positively and significantly affected by each of the whole-plot and subplot factors investigated. The fitted second-order split-plot CCD model captured almost all the variability (95%) in the generated data, and thus provides an adequate representation of the mean cake height response. The model provides an optimum cake height of $\hat{y} = 11.047cm$ at a factor setting, which consists of high level of each of the factors. This is the stationary point since all of the other possible solutions that can be explored gave less cake height. The study has established the potentials of split-plot response surface experiments in the design, analysis and determination of optimum settings of the control variables that result in a maximum (or a minimum) response for nutritional qualities of products in food industries. The proposed design will help food scientists to use their specialized knowledge to better understand the basic mechanism of the processes that go on in food industry.

ACKNOWLEDGMENT

The authors are grateful to the entire management of the Step-B Project-building in the University for allowing access to Laboratory facilities.

REFERENCES

Box, G. E. P. (1996). Split-Plot Experiments. *Quality Engineering*, *8*:515-520.

Bradley J. and Christopher J. N. (2009). Split-plot Designs: What, Why, and How. *Journal of Quality Technology,* **41**(4):340-361.

Elbert, G., Tolaba, M. P. and Suarez, C. (2001). Effects of drying conditions on head rice yield and browning index of parboiled rice. *Journal of Food Engineering*, **47**(1):37-41

Ellendersen, L. S. N., Granato, D., Guergoletto, K. B. and Wosiacki, G. (2012). Development and sensory profile of a

probiotic beverage from apple fermented with *Lactobacillus casei*. *Engineering in Life Sciences*, **12**:475–485.

Granato, D., Araujo, D., Calado, V. M. and Jarvis, B. (2014). Observations on the use of statistical methods in food science and technology. *Food Research International*, **55**:137-149.

Hand, D. and Crowder, M. (1996). Practical Longitudinal Data Analysis. 1st edition, Chapman and Hall/CRC. New York, ISBN 0-412-59940-6, pp 1-232.

Iwe, M. O., van Zuilichem, D. J., Stolp, W. and Ngoddy, P. O. (2004). Effect of extrusion cooking of soy–sweet potato mixtures on available lysine content and browning index of extrudates. *Journal of Food Engineering*, **62**: 143–150.

Jeff C.F.W. and Hamada M. S. (2009). Experiments: Planning, Analysis and Optimization. 2nd edition, John Wiley & sons, Inc., Publication. ISBN 978-0-471-69946-0, pp 1-378.

Khuri, A. I. and Mukhopadhyay, S. (2010) Response surface methodology. *WIREs Computational Statistics*, **2**: 128–149.

Kliemann, E., De Simas, K. N., Amante, E. R. et al. (2009) Optimization of pectin acid extraction from passion fruit peel (Passifloraedulisflavicarpa) using response surface methodology. *International Journal of Food Science and Technology*, **44**:476–483.

Kowalski, S. M., Parker, A. P., and Vining, G. G. (2006). Tutorial: Industrial Split-plot Experiments. *Quality Engineering,* **19**:1-15.

Little, R. C, Milliken, G. A, Stroup, W. W., Wolfinger, R.D. (1996). *SAS System for Mixed Models. Technometrics*, **39**:3, 339-344.

Letsinger, J. D., Myers, R. H. and Lentner, M. (1996). Response Surface Methods for Bi-Randomization Structures. *Journal of Quality Technology,* **28**:381-397.

Myers, R. H., Montgomery, D. C. and Christine M. A. (2009). Response Surface Methodology, Process And Product Optimization Using Designed Experiments, New York, Wiley. ISBN 978-0-470-17446-3, pp 1-680.

Nwabueze, T. U. (2010) Basic steps in adapting response surface methodology as mathematical modelling for bioprocess optimization in the food systems. *International Journal of Food Science and Technology*, **45**(9): 1768–1776.

Nunes C. A., Veronica, O. A., Anderson, S. S., Santos, J. S., and Daniel, G. (2015). The use of Statistical Software in Food Science and Technology: Advantages, Limitations and Misuses. *Food Research international*, **75**: 270-280.

Romisch U. (2005). Application of Statistical Experimental Design in Food Sciences. Institute of mathematics and informatics, Bulgarian Academy of Sciences pp 229-239.

Trinca, L. A. and Gilmour, S. G. (2000). Multi-Stratum Response Surface Designs. *Technometrics,* **43**:25-33.

Vining, G. G., Kowalski, S. M. and Montgomery, D. C. (2005). Response Surface Designs Within a Split-Plot Structure. *Journal of Quality Technology,* **37**:115-129*.*

Wang, Y., Wang L., Chiu Y. L., Chen X. O., Mao Z., Song C. F. (2008). Optimization of extrusion of flaxseeds for in vitro protein digestibility analysis using response surface methodology. *Journal of Food engineering,* **85**(1):59-64.