International Journal of Agriculture, Food and Biodiversity IJAFAB

RESPONSE SURFACE METHODOLOGY (RSM): REVIEW OF THE PRACTICAL APPROACH IN FOOD SCIENCE AND TECHNOLOGY RESEARCH

Ishiwu, C.N.

Department of Food Science and Technology, Nnamdi Azikiwe University, Awka, Nigeria

Corresponding Author's Email: cn.ishiwu@unizik.edu.ng

Abstract

Response Surface Methodology is a statistical method for modelling and analyzing a process in which the response of interest is affected by various variables. The objective of this method is to optimize the response. RSM is widely used to conduct research in different fields especially in Food Science and Technology, Chemistry and Chemical Engineering. It is the purpose of this paper to review the literature of response surface methodology emphasizing its practical application in Food Science and Technology research. Steps for designing experiment in RSM and its analysis were highlighted. The advantages of RSM as a veritable tool in experimental design and analysis can be hinged on its ability to determine the interaction of the independent variables on chosen response variables and modelling the variables mathematically by presenting the response variables as functions of independent variables. RSM from the outset assumes second order polynomial model. It is important to note that appropriate selection of independent variables and their levels in an experiment greatly influence the successful application of RSM

Keywords: RSM, CCD, FCCCD, Optimization, Variables, Modelling

Introduction

Response Surface Methodology (RSM) is a statistical tool used to establish the optimum process variables combination for product or system development. It examines how response variables of interest are affected by some chosen independent variables factored into a research. Response Surface is a method based on surface replacement, therefore the main goal of RSM study is to understand the topography of the response surface and identify the region where the most appropriate response occurs (Vining, 2005).

According to Myers (2018), RSM is moving into areas involving the use of generalized linear models (GLM's), and optimal RS designs for these areas are either difficult or impossible to implement by the user. Example of applications of GLM's include logistic and Poisson regression. Other RSM areas that will enjoy use by practitioners in the twenty-first century include multiple responses and nonparametric and semi-parametric methods.

Response surface methodology (RSM) is a collection of statistical design and numerical optimization techniques used to optimize processes and product designs. The original work in this area dates from the 1950s and has been widely used, especially in the chemical and process industries. The last 15 years have seen the widespread application of RSM in new product developments Mayer et al. (2004).

In statistics, response surface methodology (RSM) explores the relationships between several explanatory variables and each of the response variables (William and William, 1966; Nathan and Chattopadhya, 2007) Statistical approaches such as RSM can be employed to maximize the production of a special substance by optimization of operational factors (Yong, Layun and Youhua, 2018).

Several researchers have proven that RSM is a reliable tool for optimization of process variables for development of new products or systems. Nathan and Chattopadhya (2007) used Central Composite Rotatable Design (CCRD) to study the optimization of oven toasting for improving crispiness and other quality attributes of ready- to-eat potato-soy snack. Famuwagun *et al*. (2016) used Face Centered Central Composite Design (FCCCD) to optimize production of bread enriched with leafy vegetable powder.

Udaya *et al*. (2020) employed Central Composite Rotatable Design (CCRD) to study the effect of yeast concentration and total soluble solids on the quality of wine produced from pineapple. Ifesi and Ishiwu (2020) used Face Centered Central Composite Design to study the Effect of soaking time and oven-drying temperature on Physicochemical and sensory properties of 'Ogi' powder. Optimization studies in hot air-drying and vacuum dehydration were reported (Ponciano, 2020). Optimum conditions of 68°C and 10 KPa for 1.6 mm strips were established for vacuum dehydration of carrots while a 70°C drying temperature for the hot air-drying of 2 mm garlic slices was recommended.

Response surface methodology (Box–Behnken design) was used to evaluate and model effects of three factors (sweetener, low methoxy (LM) pectin and calcium content) at three levels each, on the overall acceptability of a tropical mixed fruit (pineapple, banana and passion fruit) jelly (Acosta *et al*., 2008). Khetra *et al*. (2016) reported on Selection and optimization of salt replacer, flavour enhancer and bitter blocker for manufacturing low sodium Cheddar cheese using response surface methodology.

Advantages of Using Software to Design Experiments

Optimization of product attributes using software designs eschews trial and error experiments of the past and saves the researcher both time and cost of executing researches. Optimization of process variables has been the most reliable means of validation of product quality during product development and without using software, experiments cannot be satisfactorily optimized. The responses are usually modeled for easy prediction of results since not all possible levels of a given factor in an experiment must be tried during the execution of the work. Only statistical software that has design of experiment (DOE) such as Design Expert, Minitab, etc can be used to design and analyze RSM experiments.

Application of Response Surface Methodology in Food Science Research

There are two major designs applied for RSM by researchers in area of Food Science and Technology. These designs are Central composite design (CCD) and Box-Behnken design (BBD). Food industries and food product developers have successfully used this tool in various works. The main goal is to end up with optimization of the process variables (Wang et al., 2008; Koç and Kaymak-Ertekin, 2009; Alev, 2018). Central Composite Design (CCD) has three forms namely:

a. Face Centered Central Composite Design (FCCCD)
b. Central Composite Rotatable Design (CCRD)
c. Central Composite Inscribed Design (CCID)

Out of the three, the commonly used are FCCCD and CCRD.

Face Centered Central Composite Design (FCCCD)

When the levels of each factor to be studied in an experiment are just three, the experiment is designed in Face Centered Central Composite Design (FCCCD). The design code for FCCCD is presented as:

-1 0 +1

Where:
-1 = low value; 0 = center value; +1 = high value

N/B. When designing in FCCCD, if desired, select **small** from the drop arrow of type since **small** usually generates smaller number of experimental runs than when full is selected.

Design and Analysis of RSM Experiment

Steps	Activity
1	Open the design (RSM) of the software (Design Expert) and name the independent variables then type the low and high values of the independent variables in the spaces provided.
2	Click option and click on rotatable or face centered depending on the design you want to use. Click ok and click continue to view the design table **and save it in the software because the results will be typed into the same design table during analysis**
3	Print the design table and use it as a guide for the production of the samples and label the samples correctly
4	Analyze the samples for various responses as stated in the specific objectives of the research **at any standard Laboratory**
5	Open this same design table that you saved in the software as in step 2 and create more columns in the table. In order to do this, right click on the last empty column (R1) and click on insert response on the right. Click on each of the response columns and click edit to rename it(type the response name and unit)
6	Type in the mean values of the results in the columns created; check for mistakes and if any, do necessary corrections.
7	Under analysis (Left-hand Side) Click on the first response (R1) Click transform Click fit summary Click f (x) model Click ANOVA and check if the model is not significant stop at this, but if the model is it significant print it immediately or save in word and continue
8	Click diagnostics
9	Click model graph to see the contour plot. Copy the graph and past in word
10	Click 3D surface plot on the left-hand side. Copy the graph and paste in the word environment.

Criteria that make a model adequate

After carrying out the regression analysis of a particular response variable, its ANOVA Table from the regression output is used to check for those criteria that make a model adequate and if that response meets those conditions, the researcher can proceed to fit the regression model of that particular response variable and also present its contour and 3D-surface plots

In a nut shell, it is not just any response variable that has been analyzed in the form of regression that qualifies to be fitted into a mathematical model or subjected to optimization. Definitely, certain features must be adequate as viewed from the ANOVA table of that response variable, and these include:

1. The model as shown in the regression output (ANOVA Table) has to be significant ($P<0.05$) and the model has to be significant by its p-value being less than 0.05. That is the P-value for a model has to be significant ($p<0.05$) since only significant models should be presented in a work.

2. The adjusted coefficient of determination (R2adj) should be high. Its R2adj should be at least 60%. R2adj is used instead of R2 because adding a variable to the experimental design will always increase R2regardless of whether the additional variable is statistically significant or not. On other hand it, was observed that R2adjwill not always increase as variables are added to the experimental design. In fact, if unnecessary independent variables are added to the design, the value of R2adj will often decrease. R2adj will not always increase as irrelevant variables are added to the experiment. The R2adj has to be high. The value for R2adj ranges from 0 to 1 (or 0% to 100%). The higher the R2adj, the better the model, which is an indication that the independent variables chosen for the experiment adequately explained the observed increase or decrease in the response variables.

3. Lack of fit should have $P > 0.05$, meaning not significant. But if lack of fit has $p<0.05$ meaning significant, it is not good. A non-significant lack of fit ($p>0.05$) in the model makes the model qualifies as predictive model. That means that it is a good predictor of that very response.

4. Coefficient of variation (CV): Generally, the rule is that the coefficient of variation should not be greater than 10%. CV indicates the relative dispersion of experimental point from the prediction of the model.

5. Variance inflation factor (VIF): The low value for VIF in regression coefficient table suggests that the regression coefficients are adequately estimated.

ANOVA

The analysis of variance (ANOVA) should be used to appraise the significance of the quadratic polynomial models. For each term in the models, a small *p*- value and a large *F*- value shows a more significant effect on the response (Yolmeh *et al*., 2017).

The necessary criteria are summarized as follows:

1. The model must be significant ($p \leq 0.05$)
2. Lack of fit has to be insignificant ($p > 0.05$)
3. R^2adj (coefficient of determination) should be high (close to 1)
4. Coefficient of variation (CV) should not be greater than 10

The first three criteria must be satisfactory before model equation is formed for the analyzed response variable. In addition to the four conditions stated above, only those terms whose p-values are ≤ 0.05 in the ANOVA table are selected and used in fitting the model.

Ideal Regression model for RSM

According to Myer *et al*. (2012), a second order quadratic polynomial regression will be assumed for every RSM experiment. However, the regression model to be finally used is the one suggested in the analysis of variance (ANOVA) table of that response variable and only those terms that have met the conditions for model adequacy are included in the model

Second order polynomial response surface model showing the response variable (Y_k) with the independent variable (X) is presented in eq. 1:

$$\boldsymbol{Y_k = b_{k0} + \sum_{i=1}^{2} b_{ki}X_i + \sum_{i=1}^{2} b_{kii}X_i^2 + \sum_{i\neq j=1}^{2} b_{kij}X_iX_j} \text{ …….. } \textbf{Eq.1}$$

Where b_{k0}, b_{ki}, b_{kii} *and* b_{kij} are the constant, linear, quadratic and cross-product regression coefficients respectively

$\boldsymbol{X_i}$***'s*** are the coded independent variables of X_1 and X_2

Equation 1 can be rewritten as

$Y = b_0 + b_1x_1 + b_2x_2 + b_{11}X^2{}_1 + b_{22}X^2{}_2 + b_{12}X_1X_2 + e$ …….. Eq.1

N/B: from the above equation

$\boldsymbol{b_0}$ = intercept or constant term regression coefficient

$\boldsymbol{b_1}$**,** b_2 are linear regression coefficient terms

$\boldsymbol{b_{11}}$**,** b_{22} are quadratic being squared coefficient terms

$\boldsymbol{b_{12}}$ are interaction regression coefficient terms.

e is the error of deterministic which is often neglected

Regression analysis and Analysis of Variance (ANOVA) will be conducted for fitting the models represented in Eq.1 and to determine the statistical significance of the model terms.

Exercise 1:

A study on Effect of soaking time (h)and oven drying temperature (°C) on protein, iron and pasting time of powdered 'Ogi' using FCCCD has been carried out (Ifesi and Ishiwu, 2020)

Firstly, to generate the experimental runs the design key is formed with the feasible range of values for each independent variable; and this is to be chosen by the researcher. However, values chosen may be based on information from literature review or from preliminary investigation on the same experiment. Example, **A**: soaking time (30 – 70 h) **B**: Drying temperature (40 - 100°C).

Design key

Independent variable	Code and actual values		
	-1	0	+1
A: Soaking time	30	50	70
B: Drying temperature	40	70	100

Table 1: Experimental design generated from exercise 1 showing both codes and actual values

Run	A: Soaking time (h)	B: Drying temp (°C)	Response 1 protein %	Response 2 iron mg/100g	Response 3 peak pasting time (min)
1	50.00 (0)	70.00 (0)	12.17	20.61	5.7
2	30.00 (-1)	70.00 (0)	10.27	15.82	6
3	70.00 (+1)	40.00 (-1)	12.1	20.46	5.13
4	70.00 (+1)	70.00 (0)	13.34	22.74	5.07
5	50.00 (0)	100.00 (+1)	12.52	20.68	5.1
6	50.00 (0)	70.00 (0)	12.17	20.6	5.2
7	50.00 (0)	40.00 (-1)	11.71	20.59	4.9
8	70.00 (+1)	100.00 (+1)	13.65	22.58	3.63
9	50.00 (0)	70.00 (0)	12.2	20.62	5.3
10	30.00 (-1)	100.00 (+1)	10.84	15.4	5.4
11	30.00 (-1)	40.00 (-1)	9.58	15.98	5.2
12	50.00 (0)	70.00 (0)	11.73	20.6	5.7
13	50.00 (0)	70.00 (0)	12.19	21.06	5.3

Source: Ifesi and Ishiwu (2020); Values in brackets represent the codes

Figure 1a shows that as the soaking time increased from 34 to 60 h and the drying temperature increased from 40 to 53°C, the protein content of the "ogi" increased from 10 to 12 %. Figure 1b shows that increasing both soaking time and drying temperature had marginal increase in the protein content of the "ogi".

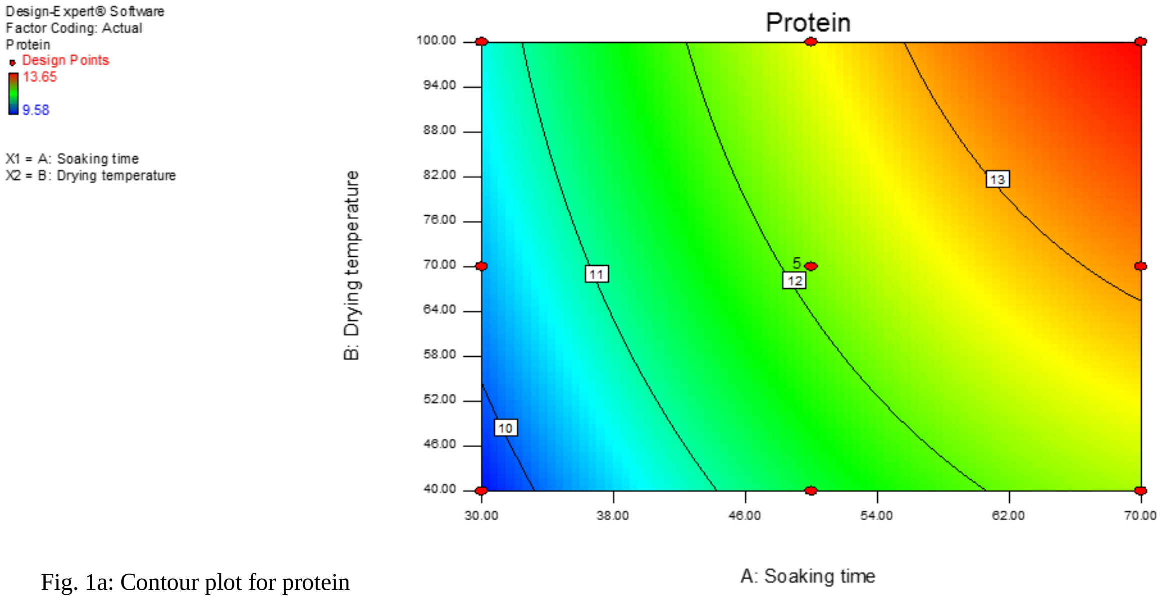

Fig. 1a: Contour plot for protein

Central Composite Rotatable Design (CCRD)

For an experiment where the levels of each independent variable or factor to be considered should exceed 3, the experiment is designed in CCRD. However, the no of levels should be exactly 5. In RCCD, it is better to limit the levels of the factors to 5 to fit into the coded form below:

-α -1 0 +1 +α

N/B: When choosing the actual vales for the design, the interval between

-1 and 0, 0 and +1 should be exactly the same value. Also the interval between –α and -1, +1 and +α should be the same value but should be smaller than that of -1 and 0, 0 and +1. Example is given below:

-α	-1	0	+1	-α
2	-6	15	24	28

Exercise 2
Optimization of oven toasting for improving crispness and other quality attributes of ready to eat potato-soy snack using response surface methodology (Nath and Chattopadhya, 2007).

Experimental design
The experiment was designed in central composite rotatable design (CCRD) that generated 13 runs.

Design key

	-α	-1	0	+1	+α
Toasting temp (°C)	85.86	90	100	110	114.14
Toasting time (min)	12.69	16	24	32	35.31

Table 2: Experimental design showing the code and actual values

Sample	Process variable		Response variable	
	Temp (°C)	Time (min)	Cr	MC%
1	110 (1)	32 (1)	45	3.39
2	110 (1)	16 (-1)	33	4.47
3	90 (-1)	32 (1)	39	5.06
4	90 (-1)	16 (1)	29	7.18
5	114.14 (1.414)	24 (0)	41	3.70
6	85.86 (-1.414)	24 (0)	32	6.56
7	100 (0)	35.31 (1.414)	43	3.17
8	100 (0)	12.69 (-1.414)	30	7.24
9	100 (0)	24 (0)	35	4.16
10	100 (0)	24 (0)	34	4.37
11	100 (0)	24 (0)	35	3.43
12	100 (0)	24 (0)	34	4.19
13	100 (0)	24 (0)	34	4.28

Source (Nath and Chattopadhya, 2007)

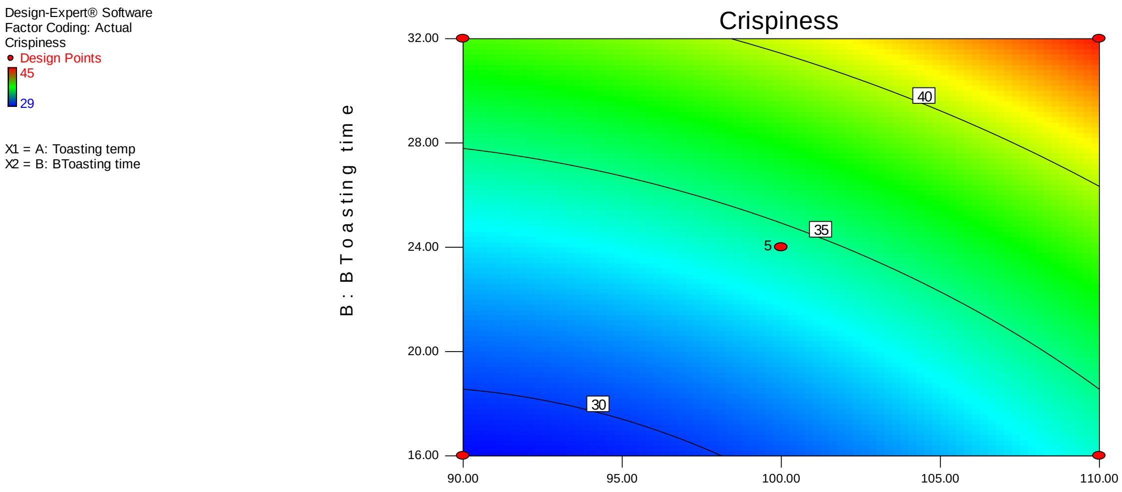

Figure 2: Contour plot of crispness
Source (Nath and Chattopadhya, 2007)

The contour plot shows that as the toasting temperature and toasting time increased from 98°C for 18.5 min to 100°C for 27 min, the crispness increased from 30 to 40. The surface plot shows that increasing both Toasting temperature and Toasting time increased the crispness. The plot shows that the snack processed at temperature of 90°C for 32 min will exhibit crispness of 38.207 and moisture content of 4.77 %.

Desirability during Numeric and Graphical Optimization
Optimization involves the detection of the best combination of factors that are able to produce the expected characteristics of the final product. In order to search for the critical values of factors that will maximize multiple responses, the desired target for each response is selected and the software combines them into an overall composite function called the desirability function.

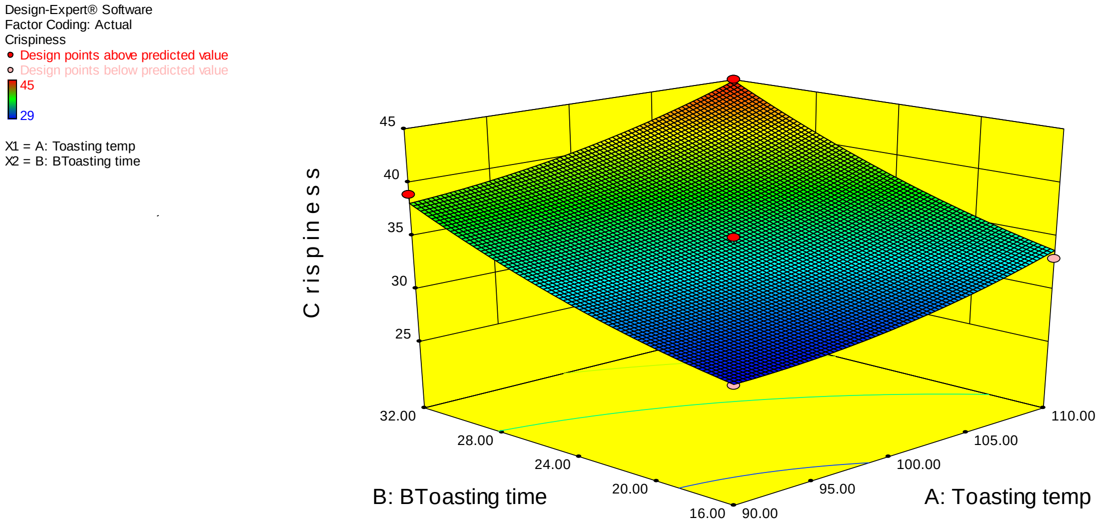

Figure 3: 3D - Surface plot of crispness; Source Nath and Chattopadhya (2007)

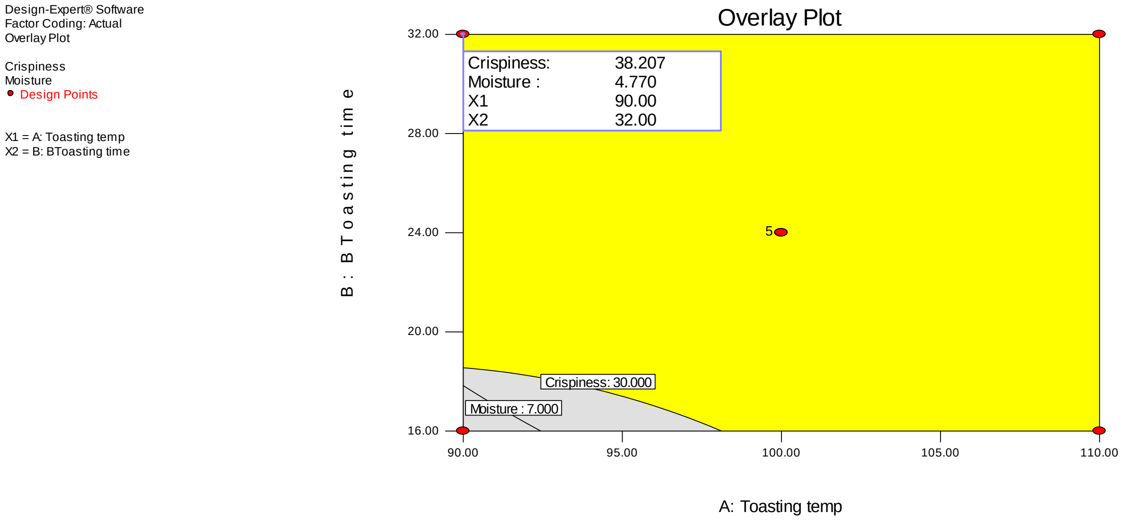

Fig. 4: Graphical Optimization for crispness and moisture content of the snack

Desirability ranges from zero to one for any given response and the higher the value the better. Design expert combines the desirability of many response variables into one value. A value of one (1.00) represents an ideal case, while zero (0.00) indicates that one or more of the responses fall outside desirable limit. Desirability that is 1 or very close to 1 is very good, since it shows that the selected components or process variables will actually produce the targeted response value when tried.

Steps for Optimization

Two types of optimization exist (numerical and graphical). Numerical heralds graphical since the graph is generated from any of the possible solutions suggested in the numerical table:

Step	Activity
1	Click numerical
2	Select the first response variable from the list
3	From the drop arrow of **goal**, select either in-range or minimize or maximize depending on what you want for that very response variable in the product being developed
4	Click solution
5	Repeat steps 2,3 and 4 by selecting each of the response variables one after the other N/B: When the last response variable from the list has been processed, the combined numerical values for both the process (independent) variables and the response (dependent) variables will appear,
6	Select the row that has exhibited desirability of 1 or close to 1and which also gives the desired response variable value(s). Also click
7	Click on graphical (LHS) under optimization
8	The first response will be highlighted, type both the lower and upper limit values, but the values chosen must fall within the range suggested/given within the dialogue box
8	Repeat step 7 for all the response variables
9	Click on graph to view the optimization (overlay) plot

Conclusion

It has been observed that not all systems with curvature are compatible with second order polynomial model normally assumed in RSM especially when not all terms meet up with the qualifying conditions for their inclusion into a model. Whenever a term lacks the conditions for being fitted into the model, such term is eschewed, thereby making the second order polynomial equation impossible. Some models of RSM can just be linear or linear with interaction as the case maybe

According to Myers (2018), RSM is moving into areas involving the use of generalized linear models (GLM's), and optimal RS designs for these areas are either difficult or impossible to implement by the user. Example of applications of GLM's include logistic and Poisson regression. Other RSM areas that will enjoy use by practitioners in the twenty-first century include multiple responses and nonparametric and semi-parametric methods. In addition, design and analysis techniques for cases where natural restrictions in randomization occur need to be addressed further and communicated to users using RSM

References

Acosta, O, Víquez, F, Cubero, E. (2008). Optimization of low calorie mixed fruit jelly by response surface methodology Food Quality and Preference, 19 (1): 79-85 https://doi.org/10.1016/j.foodqual.2007.06.010

Alev Yuksel Aydar (2018) Utilization of Response, Surface Methodology in Optimization of Extraction of Pant Materials, Licenses InTech. (http://creativecommons.org/licenses/by /3.0.

Famuwagun, A. A, Taiwo, K. D, Gbadamosi, S. O and Oyedele, D. J. (2016) Optimize of production of bread enriched with leafy vegetable powder. *Journal of Food Processing & Technology*, 7 (7): 1-7.

Ifesi J and Ishiwu, C. N. (2020). Effect of soaking time and oven-drying temperature on Physico-chemical and sensory properties of 'ogi' powder, PGD Project, Department of Food Science and Technology, Nnamdi Azikiwe University, Awka.

Khetra, Y., Kanawjia, S. K., & Puri, R. (2016). Selection and optimization of salt replacer, flavour enhancer and bitter blocker for manufacturing low sodium Cheddar cheese using response surface methodology. *LWT- Food Science and Technology*, 72, 99– 106. https://doi.org/10.1016/ j.lwt.2016.04.035.

Koç, B. and Kaymak-Ertekin, F. (2009). Response Surface Methodology and Food Processing Applications, *Gida*; 7:1-8.

Myers R. H,Khuri A. I. & Carter, W. H. (2012). Response Surface Methodology: 1966–l988. Pages 137-157 https://doi.org/10.1080/00401706.1989.10488509.

Myers, R. H. (2018). Response Surface Methodology—Current Status and Future Directions, *https://doi.org/10.1080/00224065.1999.11979891*.

Myers, R. H., Montgomery, Geoffrey Vining, D, G, Borror, C. M and. Kowalski, S. M. (2004). Response Surface Methodology: A Retrospective and Literature Survey, *J. Qual. Technology*: 36-77

Nath, A and Chattopadhya, P. K. (2007). Optimization of oven toasting for improving crispness and other quality attributes of ready-to-eat potato-soy snack. *Journal of Food Engineering* 80: 1282 – 1292.

Ponciano, S.M. (2020).The Response Surface Methodology: An Application to Optimize Dehydration Operations of Selected Agricultural Crops, *LWT - Food Science and Technology,* Volume 35 (7): 584-592.

Udaya, E. C, Ishiwu, C. N, Obiora, C. U and Iwouno, J. O. (2020) Effect of yeast concentration and total soluble solids on the quality of wine produced from pineapple, Current Journal of Applied Science and Technology, 39 (30): 28-42.

Vining, G, Kowalski, S. M and Montgomery, D. C. (2005). Response Surface Designs within a Split-Plot Structure, *Journal of Quality Technology*, 37 (2):115-129.

Wang, J, Sun, B, Cao, Y, Tian, Y and Li; X. (2008). Optimization of Ultrasound-Assisted Extraction of Phenolic Compounds from Wheat Bran. *Food Chemistry*, 106: 804-810.

William, J. H and William, G. H. (1966). A Review of Response Surface Methodology: A Literature Survey, Technometrics, Vol. 8, No. 4 (Nov., 1966), pp. 571-590 (20 pages), Published By: Taylor & Francis, Ltd, https://doi.org/10.2307/1266632, https://www.jstor.org/stable/1266632.

Yolmeh, M., Khomeiri, M., Ghorbani, M., Ghaemi, E., & Ramezanpour, S. S. (2017). High efficiency pigment production from *Micrococcus roseus* (PTCC 1411) under ultraviolet irradiation. *Biocatalysis and Agricultural Biotechnology*, 9, 156– 161. https://doi.org/10.1016/j.bcab.2016.12.010.

Yong, W, Layun, D and Youhua, F, (2018). Preparation of Soy-Based Adhesive Enhanced by Waterborne Polyurethane: Optimization by Response Surface Methodology, Advances in Materials Science and Engineering Volume 2018 |Article ID 9253670 | https://doi.org/10.1155/2018/9253670