Optimization Of A Soap Production Mix Using Response Surface Modeling: A Case Of Niger Bar Soap Manufacturing Industry Onitsha, Anambra State, Nigeria

Okolie Chukwulozie Paul, Ezeliora Chukwuemeka Daniel, Iwenofu Chinwe Onyedika, Sinebe Jude Ebieladoh

Abstract: This paper improves the productivity of Soap mix using Response Surface and presents the optimization of a soap production mix using the previous production data. The data were optimized using response surface modeling tool to observe the optimum production mix of the raw material. Response surface regression analysis was used to estimate the coefficients for Y using data in coded units were the coefficient of determination (R-sq) is 100%. The response surface optimization model shows that the optimum production mix quantity of the soap should be 364.999kg.

Key word: Fatty matter in the oil, Oil, Production, Regression analysis, Response surface model, Salt and optimal point, Saponification, Silicate, Soap, Starch.

1. INTRODUCTION

In chemistry, soap is a salt of a fatty acid [1]. Soaps are mainly used as surfactants for washing, bathing, and cleaning, but they are also used in textile spinning and are important components of lubricants. Soaps for cleansing are obtained by treating vegetable or animal oils and fats with a strongly alkaline solution. Fats and oils are composed of triglycerides; three molecules of fatty acids are attached to a single molecule of glycerol [2]. The alkaline solution, which is often called lye (although the term "lye soap" refers almost exclusively to soaps made with sodium hydroxide), brings about a chemical reaction known as saponification. In this reaction, the triglyceride fats are first hydrolyzed into free fatty acids, and then these combine with the alkali to form crude soap, an amalgam of various soap salts, excess fat or alkali, water, and liberated glycerol (glycerin). The glycerin is a useful by-product, which can be left in the soap product as a softening agent, or isolated for other uses [2]. Soaps are key components of most lubricating greases, which are usually emulsions of calcium soap or lithium soaps and mineral oil. These calcium- and lithium-based greases are widely used.

- *Okolie Chukwulozie Paul, Ezeliora Chukwuemeka Daniel, Iwenofu Chinwe Onyedika, Sinebe Jude Ebieladoh*
- *Mechanical Engineering Department, Nnamdi Azikiwe University, P.M.B. 5025 Awka, Anambra State, Nigeria*
- *Industrial Production Engineering, Nnamdi Azikiwe University, P.M.B. 5025 Awka, Anambra State, Nigeria*
- *Mechanical Engineering Department, Federal Polytechnic Oko, Anambra State, Nigeria*
- *Special Adviser to the Governor on Local Government Projects Monitoring, Delta State,Nigeria*
- *Email: pc.okolie@unizik.edu.ng*

Many other metallic soaps are also useful, including those of aluminum, sodium, and mixtures of them. Such soaps are also used as thickeners to increase the viscosity of oils. In ancient times, lubricating greases were made by the addition of lime to olive oil [3]. The aim of this study is to analyze the past production mix data in other to observe the optimal production quantity of the soap mix material.

I. Saponification

This is a process that produces soap, usually from fats and lye. In technical terms, saponification involves base (usually caustic soda NaOH) hydrolysis of triglycerides, which are esters of fatty acids, to form the sodium salt of a carboxylate. In addition to soap, such traditional saponification processes produce glycerol. "Saponifiable substances" are those that can be converted into soap [4]. Knowledge of saponification is relevant to many technologies and many aspects of everyday life.

II. Soft vs hard soap

Depending on the nature of the alkali used in their production, soaps have distinct properties. Sodium hydroxide (NaOH) gives "hard soap", whereas, when potassium hydroxide (KOH) is used, a soft soap is formed [5].

III. Soap-making processes

The industrial production of soap involves continuous processes, such as continuous addition of fat and removal of product. Smaller-scale production involves the traditional batch processes. The three variations are: the 'cold process', wherein the reaction takes place substantially at room temperature, the 'semi boiled' or 'hot process', wherein the reaction takes place near the boiling point, and the 'fully boiled process', wherein the reactants are boiled at least once and the glycerol is recovered. There are two types of 'semi boiled' hot process methods. The first is the ITMHP (in the mold hot process) and the second is the CPHP (Crockpot hot process). Typically soap makers choose the hot process methods if they wish to remove the cure time to a three-day air dry process. Most soap makers, however, continue to prefer the cold process method. The cold process and hot process (semi boiled) are the simplest and typically used by small artisans and hobbyists producing handmade decorative soaps. The glycerin remains in the soap and the reaction continues for many days after the soap is poured into moulds. The glycerin is left during the hot-process method, but at the high temperature employed, the reaction is practically completed in the kettle, before the soap is poured into moulds. This simple and quick process is employed in small factories all over the world [6]. Handmade soap from the cold process also differs from industrially made soap in that an excess of fat is used, beyond that needed to consume the alkali (in a cold-pour process, this excess fat is called “superfatting”), and the glycerin left in acts as a moisturizing agent. However, the glycerin also makes the soap softer and less resistant to becoming “mushy” if left wet. Since it is better to add too much oil and have left-over fat, than to add too much lye and have left-over lye, soap produced from the hot process also contains left-over glycerin and its concomitant pros and cons. Further addition of glycerin and processing of this soap produces glycerin soap. Superfatted soap is more skin-friendly than one without extra fat. However, if too much fat is added, it can leave a “greasy” feel to the skin. Sometimes, an emollient additive, such as jojoba oil or shea butter, is added “at trace” (i.e., the point at which the saponification process is sufficiently advanced that the soap has begun to thicken in the cold process method) in the belief that nearly all the lye will be spent and it will escape saponification and remain intact. In the case of hot-process soap, an emollient may be added after the initial oils have saponified so they remain unreacted in the finished soap. Superfatting can also be accomplished through a process known as “lye discount” in which the soap maker uses less alkali than required instead of adding extra fats [7].

IV. Introduction to Response surface methodology (RSM)

Response surface methodology (RSM) is a collection of mathematical and statistical techniques for empirical model building. By careful design of experiments, the objective is to optimize a response (output variable) which is influenced by several independent variables (input variables). An experiment is a series of tests, called runs, in which changes are made in the input variables in order to identify the reasons for changes in the output response. Originally, RSM was developed to model experimental responses [8], and then migrated into the modeling of numerical experiments. The difference is in the type of error generated by the response. In physical experiments, inaccuracy can be due, for example, to measurement errors while, in computer experiments, numerical noise is a result of incomplete convergence of iterative processes, round-off errors or the discrete representation of continuous physical phenomena [9, 10, 11]. In RSM, the errors are assumed to be random. The application of RSM to design optimization is aimed at reducing the cost of expensive analysis methods (e.g. finite element method or CFD analysis) and their associated numerical noise. The problem can be approximated as described in with smooth functions that improve the convergence of the optimization process because they reduce the effects of noise and they allow for the use of derivative-based algorithms. Venter et al. (1996) [12] have discussed the advantages of using RSM for design optimization applications. For example, in the case of the optimization of the calcinations of Roman cement described in Section 6.3, the engineer wants to find the levels of temperature(x) and time (x) that maximize the early age strength (y) of the cement. The early age strength is a function of the levels of temperature and time, as follows:

$$y = f(x_1, x_2) + e \qquad (1)$$

Where; e represents the noise or error observed in the response *y*. The surface represented by $f(x_1, x_2)$ is called a *response surface*. The response can be represented graphically, either in the three-dimensional space or as contour plots that help visualize the shape of the response surface. The two basic concepts in RSM are first the choice of the approximate model and, second, the plan of experiments where the response has to be evaluated. Generally, the structure of the relationship between the response and the independent variables is unknown. The first step in RSM is to find a suitable approximation to the true relationship. The most common forms are low-order polynomials (first or second-order). The advantage is that the structure of the approximation is not assumed in advance, but is given as part of the solution, thus leading to a function structure of the best possible quality. In addition, the complexity of the function is not limited to a polynomial but can be generalized with the inclusion of any mathematical operator (e.g. trigonometric functions), depending on the engineering understanding of the problem [8].

2. RESEARCH METHOD AND DATA COLLECTION

Quantitative approach was used. The data collected were analyzed using surface response method. The method was used to observe the optimal point of the data.

Table 1: Quantity of the raw material for Soap Production Mix

Serial No	(X1) (liters)	(X2) (kg)	(X3) (kg)	(X4) (liters)	(X5) (kg)	(Y) (kg)
1	1500	26	50	150	69	360
2	1500	25	63	150	67	364
3	1500	25	64	175	69	365
4	1500	25	62	150	66	363
5	1500	28	68	175	68	366
6	1500	26	70	175	70	367
7	1500	25	63	150	65	364
8	1500	27	65	175	66	365
9	1500	26	70	175	70	367
10	1500	28	68	175	70	366
11	1500	25	62	150	67	363
12	1500	25	50	150	67	362
13	1500	28	70	175	68	367
14	1500	26	50	150	68	361
15	1500	28	68	175	67	366
16	1500	26	63	150	69	364

Where X1=Quantity of oil used (in liters), X2 = Quantity of salt used (in kg), X3 = Quantity of starch used (in kg), X4 = quantity of silicate used (in liters), X5 = Total fatty matter in oil (in %), and Y = Quantity of soap produced (in units of 11kg)

3. Data analysis and Results

I. Optimal Design: (X2), (X3), (X4), (X5)

Response surface design selected according to D-optimality

Number of candidate design points: 16

Number of design points in optimal design: 20

Model terms: A, B, C, D, AA, BB, CC, DD, AB, AC, AD, BC, BD, CD

Initial design generated by Sequential method

Initial design improved by Exchange method

Number of design points exchanged is 1

Optimal Design

Row number of selected design points: 6, 16, 10, 7, 8, 11, 12, 14, 2, 13, 5, 4,
15, 10, 8, 3, 6, 1, 16, 3

II. Response Surface Regression: (Y) versus (X2), (X3), (X4), (X5)

The analysis was done using coded units.

Estimated Regression Coefficients for (Y)

Term	Coef	SE Coef	T	P
Constant	364.534	0.23263	1567.027	0.000
(X2)	0.472	0.10815	4.368	0.049
(X3)	-0.749	0.14632	-5.118	0.036
(X4)	2.934	0.18009	16.294	0.004
(X5)	-0.795	0.12453	-6.384	0.024
(X2)*(X2)	-0.009	0.33586	-0.027	0.981
(X3)*(X3)	7.323	0.21315	34.358	0.001
(X5)*(X5)	0.023	0.09714	0.237	0.835
(X2)*(X3)	0.525	0.31255	1.681	0.235
(X2)*(X4)	0.152	0.16112	0.945	0.444
(X2)*(X5)	-0.965	0.21051	-4.583	0.044
(X3)*(X4)	-8.241	0.43445	-18.970	0.003
(X3)*(X5)	1.474	0.25770	5.721	0.029
(X4)*(X5)	0.559	0.35152	1.591	0.253

S = 0.0252701 PRESS = *

R-Sq = 100.00% R-Sq(pred) = *% R-Sq(adj) = 99.99%

Analysis of Variance for (Y)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	13	69.7487	69.7487	5.36529	8401.92	0.000
Linear	4	66.1891	15.9633	3.99082	6249.53	0.000
(X2)	1	19.4213	0.0122	0.01218	19.08	0.049
(X3)	1	44.4199	0.0167	0.01673	26.19	0.036
(X4)	1	2.3076	0.1695	0.16954	265.49	0.004
(X5)	1	0.0403	0.0260	0.02602	40.75	0.024
Square	3	1.3433	0.7734	0.25780	403.71	0.002
(X2)*(X2)	1	0.0407	0.0000	0.00000	0.00	0.981
(X3)*(X3)	1	1.2931	0.7538	0.75381	1180.45	0.001
(X5)*(X5)	1	0.0095	0.0000	0.00004	0.06	0.835
Interaction	6	2.2163	2.2163	0.36939	578.46	0.002
(X2)*(X3)	1	0.9067	0.0018	0.00181	2.83	0.235
(X2)*(X4)	1	0.1658	0.0006	0.00057	0.89	0.444
(X2)*(X5)	1	0.0462	0.0134	0.01341	21.00	0.044
(X3)*(X4)	1	0.4209	0.2298	0.22980	359.86	0.003
(X3)*(X5)	1	0.6751	0.0209	0.02090	32.72	0.029
(X4)*(X5)	1	0.0016	0.0016	0.00162	2.53	0.253
Residual Error	2	0.0013	0.0013	0.00064		
Lack-of-Fit	1	0.0013	0.0013	0.00128	*	*
Pure Error	1	0.0000	0.0000	0.00000		
Total	15	69.7500				

Obs	StdOrder	(Y)	Fit	SE Fit	Residual	St Resid
1	1	360.000	360.002	0.025	-0.002	-1.41 X
2	2	364.000	364.010	0.024	-0.010	-1.41
3	3	365.000	365.000	0.025	0.000	1.41 X
4	4	363.000	363.022	0.020	-0.022	-1.41
5	5	366.000	365.996	0.025	0.004	1.41
6	6	367.000	367.000	0.018	-0.000	-0.01
7	7	364.000	363.990	0.024	0.010	1.41
8	8	365.000	364.999	0.025	0.001	1.41 X
9	9	367.000	367.000	0.018	-0.000	-0.01
10	10	366.000	366.001	0.025	-0.001	-1.41 X
11	11	363.000	362.977	0.019	0.023	1.41
12	12	362.000	362.001	0.025	-0.001	-1.41 X
13	13	367.000	366.998	0.025	0.002	1.41 X
14	14	361.000	360.997	0.025	0.003	1.41 X
15	15	366.000	366.004	0.025	-0.004	-1.41
16	16	364.000	364.000	0.025	-0.000	-1.41 X

X denotes an observation whose X value gives it large leverage.

Estimated Regression Coefficients for (Y) using data in uncoded units

Term	Coef
Constant	44.4192
(X2)	14.4700
(X3)	-3.05753
(X4)	2.76733
(X5)	-0.444431
(X2)*(X2)	-0.00400961
(X3)*(X3)	0.0732317
(X5)*(X5)	0.00368487
(X2)*(X3)	0.0350327
(X2)*(X4)	0.00812319
(X2)*(X5)	-0.257253
(X3)*(X4)	-0.0659309
(X3)*(X5)	0.0589666
(X4)*(X5)	0.0178966

Predicted Response for New Design Points Using Model for (Y)

Point	Fit	SE Fit	95% CI	95% PI	
1	360.002	0.0252262	(359.894, 360.111)	(359.848, 360.156)	X
2	364.010	0.0242043	(363.906, 364.114)	(363.860, 364.161)	
3	365.000	0.0252698	(364.891, 365.109)	(364.846, 365.154)	X
4	363.022	0.0198380	(362.937, 363.107)	(362.884, 363.160)	
5	365.996	0.0251136	(365.888, 366.104)	(365.843, 366.149)	
6	367.000	0.0178678	(366.923, 367.077)	(366.867, 367.133)	
7	363.990	0.0243012	(363.886, 364.095)	(363.839, 364.141)	
8	364.999	0.0252676	(364.891, 365.108)	(364.846, 365.153)	X
9	367.000	0.0178678	(366.923, 367.077)	(366.867, 367.133)	
10	366.001	0.0252535	(365.893, 366.110)	(365.848, 366.155)	X
11	362.977	0.0192760	(362.894, 363.060)	(362.840, 363.114)	
12	362.001	0.0252676	(361.892, 362.109)	(361.847, 362.154)	X
13	366.998	0.0252446	(366.890, 367.107)	(366.845, 367.152)	X
14	360.997	0.0252065	(360.889, 361.106)	(360.844, 361.151)	X
15	366.004	0.0250734	(365.897, 366.112)	(365.851, 366.158)	
16	364.000	0.0252683	(363.892, 364.109)	(363.847, 364.154)	X

X denotes a point that is an outlier in the predictors.

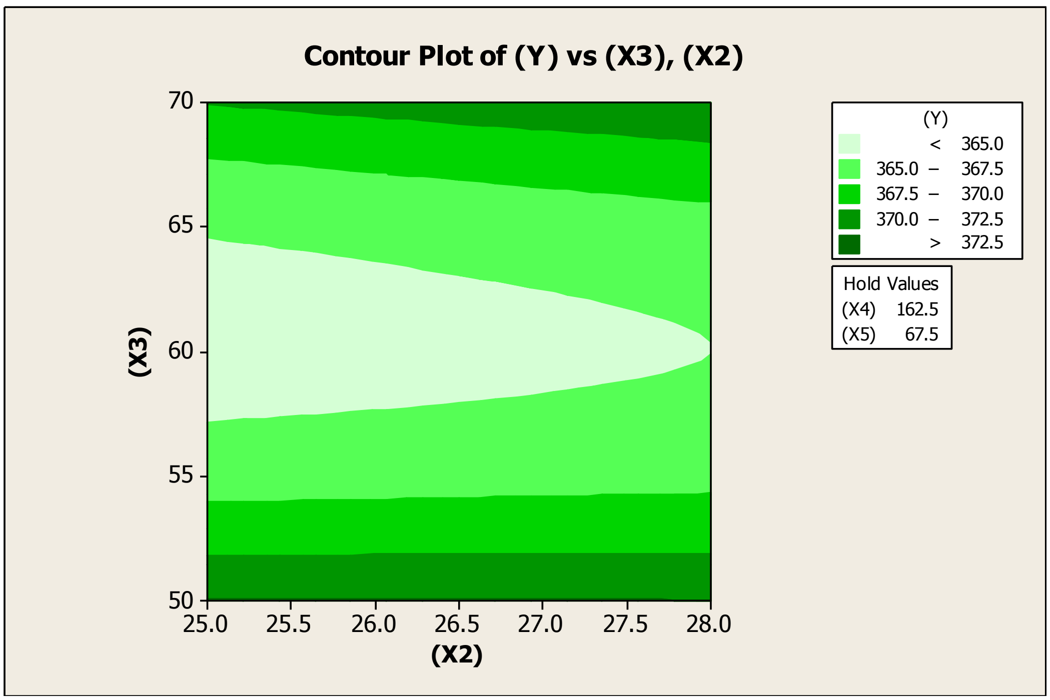

Fig.1: Contour Plot of (Y) vs (X3), (X2)

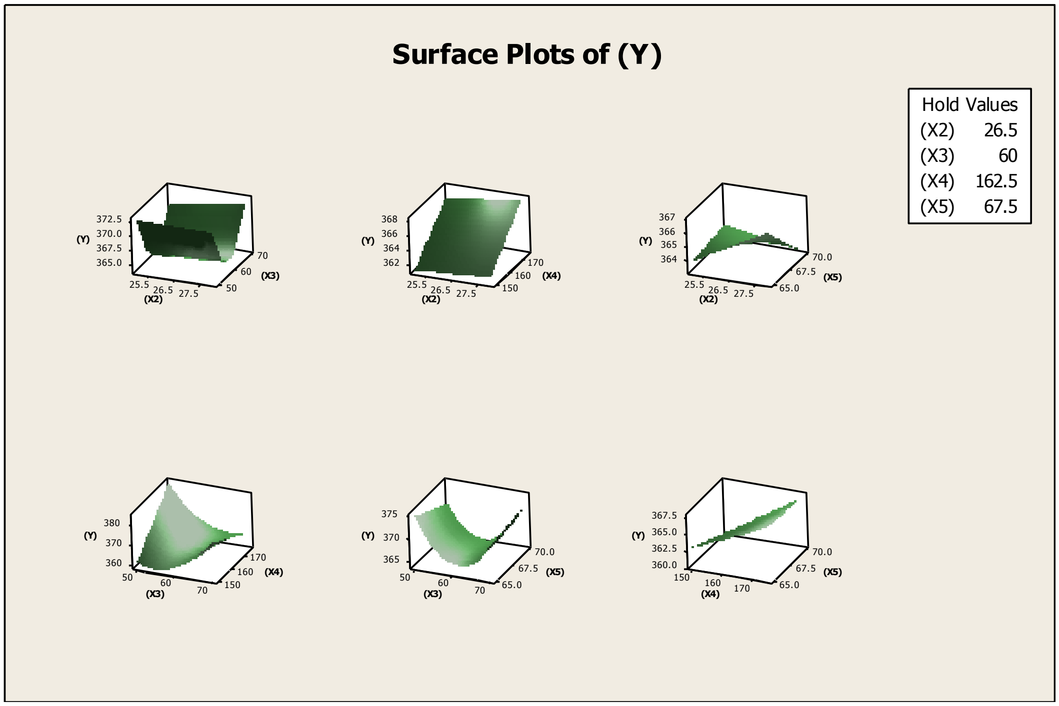

Fig.2: Surface Plots of (Y)

III. Response Optimization

Parameters

Goal	Lower	Target	Upper	Weight	Import
(Y) Target	0	365	400	1	1

Starting Point

(X2) = 25
(X3) = 50
(X4) = 150
(X5) = 65

Local Solution
(X2) = 27.9991
(X3) = 50
(X4) = 156.251
(X5) = 68.8436

Predicted Responses
(Y) = 364.999 , desirability = 0.999998

Composite Desirability = 0.999998

Global Solution

(X2) = 27.9991
(X3) = 50
(X4) = 156.251
(X5) = 68.8436

Predicted Responses

(Y) = 364.999 , desirability = 0.999998

Composite Desirability = 0.999998

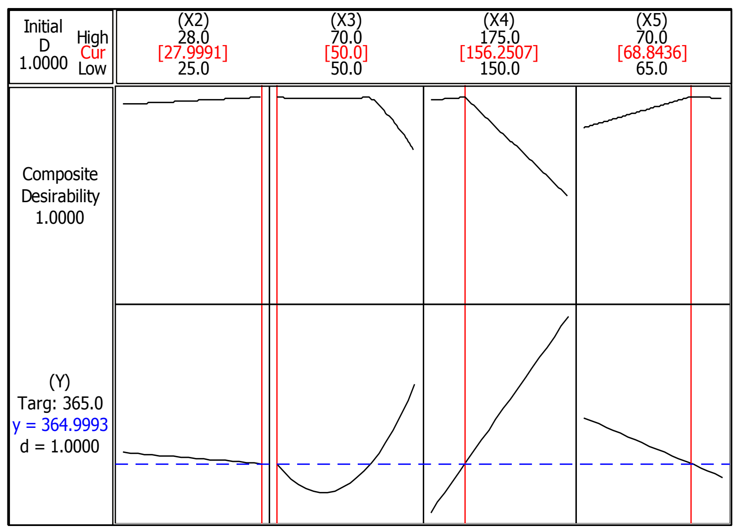

Fig.3: Optimization Plot

Table 2: Optimum production Mix of the Raw Materials

Setting	D-Value	(X1)	(X2)	(X3)	(X4)	(X5)
Initial	1.0	1500	27.999 1	50.000 0	156.250 7	68.8436
Optimal	1.0	1500	27.999 1	50.000 0	156.250 7	68.8436

Table 3: Optimization Results

StdOrder	RunOrder	Blocks	PtType	FITS1	RESI1	COEF1	QUAD1	DESIR1
1	1	1	1	360.002	-0.0021058	364.534	4.000	0.98631
2	2	1	1	364.010	-0.0102693	0.472	0.000	0.99729
3	3	1	1	365.000	0.0001690	-0.749	0.000	1.00000
4	4	1	1	363.022	-0.0221374	2.934	3.000	0.99458
5	5	1	1	365.996	0.0039712	-0.795	4.000	0.97154
6	6	1	1	367.000	-0.0002535	-0.009	5.000	0.94285
7	7	1	1	363.990	0.0098013	7.323	6.000	0.99723
8	8	1	1	364.999	0.0005070	0.023	364.534	1.00000
9	9	1	1	367.000	-0.0002535	0.525	0.472	0.94285
10	10	1	1	366.001	-0.0012956	0.152	-0.749	0.97139
11	11	1	1	362.977	0.0231091	-0.965	2.934	0.99446
12	12	1	1	362.001	-0.0005037	-8.241	-0.795	0.99178
13	13	1	1	366.998	0.0016054	1.474	-0.009	0.94290
14	14	1	1	360.997	0.0025348	0.559	7.323	0.98903
15	15	1	1	366.004	-0.0044500		0.000	0.97130

4. DISCUSSION AND CONCLUSION

The two main basic concepts in RSM, are; first the choice of the approximate model and, second, the plan of experiments where the response has to evaluated.From the results, it was observed that the P-value of all the variables were below a 2 tail significant value of 0.05. This shows that all the values were significant for the analysis. It shows that all the variables are important to the model. The response surface regression analysis was done using coded units. From the analysis it was observed that the coefficient of determination (R-sq) is 100%. This shows that the proportion of the variation of the dependent variable y was accounted for 100% by all the independent variables combined. Analysis of variance observes the interaction of the variables and it's significant. From its observation, it shows that there is zero pure error in the model. This confirms the result of the response surface regression model which shows that the coefficient of determination of the model is 100%. However, the confidence level of the results is 95% confidence The graphically representation of the response can either be in the three-dimensional space or as contour plots that help visualize the shape of the response surface. Contours are curves of constant response drawn in the xi, xj plane keeping all other variables fixed. Each contour corresponds to a particular height of the response surface. Three-dimensional response surface and the corresponding contour plot for the production mix of soap were plotted. Where X1=qty of oil used (in liters), X2 = qyt of salt used (in kg), X3 = Quantity of starch used (in kg), X4 = quantity of silicate used (in liters), X5 = Total fatty matter in oil (in %), Y = Quantity of soap produced (in units of 11kg) The optimization results show that X2 = quantity of salt (in kg) will be used at 27.9991kg, X3 = Quantity of starch (in kg) will be used at 50.0000kg, X4 = quantity of silicate (in liters), will be used at 156.2507 litters while X5 = Total fatty matter in oil (in %), 68.8436%. However, Y = Quantity of soap produced (in units of 11kg) will be produced at 364.999kg. Furthermore, X1=qty of oil (in liters) is constant; it will be used at the same rate of 1500litres. It was observed that the response surface optimization model shows that the optimum production mix quantity of the soap will be 364.999kg. The result was recommended to the case study company to work with the optimum of their production mix quantity.

References

[1] IUPAC. "IUPAC Gold Book – soap" Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. doi:10.1351/goldbook. Accessed 2010-08-09

[2] Jump up to: [a] [b] Cavitch, Susan Miller. The Natural Soap Book. Storey Publishing, 1994 ISBN 0-88266-888-9.

[3] Thorsten Bartels et al. "Lubricants and Lubrication" in Ullmann's Encyclopedia of Industrial Chemistry, 2005, Weinheim. doi:10.1002/14356007.a15_423

[4] K. Schumann, K. Siekmann “Soaps” in Ullmann’s Encyclopedia of Industrial Chemistry 2005, Wiley-VCH, Weinheim

[5] Thorsten Bartels et al. "Lubricants and Lubrication" in Ullmann's Encyclopedia of Industrial Chemistry, 2005, Weinheim. doi:10.1002/14356007.a15 423

[6] McNeil, Ian (1990). An Encyclopaedia of the history of technology. Taylor & Francis. pp. 2003–205. ISBN 978-0-415-01306-2.

[7] "Colgate-Palmolive Company History: Creating Bright Smiles for 200 Years". Colgate-Palmolive Company. Retrieved 17 October 2012.

[8] Box, G.E.P. and Draper, N.R. (1987), Empirical model-building and response surfaces, New York: John Wiley & Sons.

[9] Giunta, A.A., Balabanov, V., Haim, D., Grossman, B., Mason, W.H., Watson, L.T.,Haftka, R.T. (1996), “Wing design for a high-speed civil transport using a design of experiments methodology”. AIAA paper 96-4001-CP, Proceedings of 6th AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Bellevue WA, Part 1, pp. 168-183.

[10] Van Campen, D.H., Nagtegaal, R. and Schoofs, A.J.G. (1990), “Approximation methods in structural optimization using experimental designs for multiple responses”, In: Eschenauer, H., Koski, J., Osyczka, A. (Ed.), Multicriteria Design Optimization, Springer-Verlag, Berlin, Heidelberg, New York, pp. 205-228.

[11] Toropov, V.V., van Keulen, F., Markine, V.L. and de Boer, H. (1996), “Refinements in the multi-point approximation method to reduce the effects of noisy responses”, 6th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis andOptimization, Bellevue WA, Part 2, pp. 941-951.

[12] Venter, G., Haftka, R.T. and Starnes, J.H. (1996), “Construction of response surfaces for design optimization applications”, AIAA paper 96-4040-CP, Proceedings of 6th AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization,Bellevue WA, Part 2, pp. 548-564.