TITLE: Optimization Of Multi Responses Using Data Envelopment Analysis: The Application in Food Industry

AUTHORS: Duygu KILIC,Meral EBEGIL,Hulya BAYRAK,Berrin OZKAYA,Basak APAYDIN AVSAR

PAGES: 1083-1090

ORIGINAL PDF URL: https://dergipark.org.tr/tr/download/article-file/792558

Optimization Of Multi Responses Using Data Envelopment Analysis: The Application in Food Industry

Meral EBEGIL[1]* (iD) , Hulya BAYRAK[1] (iD) , Berrin OZKAYA[2] (iD) , Duygu KILIC[1] (iD) , Başak APAYDIN AVSAR[3]
(iD)

[1]*Gazi University, Faculty of Sciences, Department of Statistics, Ankara, Turkey*

[2]*Ankara University, Faculty of Engineering, Department of Food Engineering, Ankara, Turkey*

[3] *The Ministry of Science, Industry and Technology, Ankara, Turkey*

Highlights
• There are multiple response problems.
• Therefore DEA has been used.
• The proposed combination provides better results with fewer substances in the field of food.
• In the field of food, the number of analyses performed using this combination is small.

Abstract

Multiple response is a widely used method of increasing product quality, optimizing cost and time in industry. However, technological developments and processes are becoming more and more complex, which means that more than one response is effective rather than a single response, in product or process optimization. The Response Surface Methodology (RSM) can be used to optimize a single response or multiple responses. It is known that when there are numerous responses, it is difficult and complex to optimize responses simultaneously. Data Envelopment Analysis (DEA) is a statistical approach where multiple inputs and multiple outputs, regardless of how many they are can simultaneously be optimized. For this reason, in this study Data Envelopment Analysis (DEA) technique was applied in combination with The Response Surface Methodology (RSM) and this enabled us to optimize more than one response concurrently.

1. INTRODUCTION

Design of experiment (DOE) is a very popular technique used in many areas, especially in industry, chemistry, food and health. It is extensively used in industries that are often concerned with improving product quality, reducing costs, or obtaining the optimal parameter-setting. Technological advancements and the ever changing customer demands both have made the experiment design process more and more complicated. When using DOE, response surface methodology (RSM) is frequently used to obtain the optimal parameter-setting following analysis of variance (ANOVA) for identifying significant factors [1]. In the past when RSM was applied, a single response was thought to be initially sufficient , however as this method has been used more extensively, it was seen that single response was not sufficient. As a result of this situation, multiple responses are optimized together. By fitting a surface to a data set via RSM, determining the optimum levels of factors, an equation representing the relationship between response variance and control variance can be obtained. In order to determine surface graphs and characterize the optimal inputs, a contour plot can be employed. Once several responses are recognized, it is possible to determine the optimal parameter by means of visualizing overlay contour plots [2]. This situation is a little tricky and complicated. A common approach for optimizing multiple responses in DOE is to formulate a multi-response problem as a constrained optimization problem [3]. The response surface of any relevant

response is selected as an objective function. Other response surfaces have limit functions. In this respect, RSM is an optimization problem in which mathematical and statistical techniques are used together. There are other approaches to multiple responses in design of experiment. One such method is the one proposed by Derringer-Suich and known as the desirability approach. [4]. On the other hand, Khuri and Conlon (1981), used a method based on a multivariate regression model [5]. In addition, Su and Tong (1997), have tried to solve multiple response problem with the help of Principal Component Analysis (PCA) [6]. In another study worth mentioning, the Derringer-Suich method was employed by Köksoy (2005) in order to optimize dual responses adopting a different approach. [7]. Besides, Tsai et al.(2010), have developed a new method involving RSM and DEA to solve the multiple response problem [1]. In a different study about electrical discharge machining (EDM) , which is a non-conventional machining process mainly used for the machining of complex shapes , Satu et. al. (2013) proposed a DEA approach where a RSM was adopted for the assessment of the effects of various process parameters on the outcome of important responses. In this study they showed that DEA methodology along with ARV (average ranked value) approach gives satisfactory results in determining suitable conditions among a large number of alternative processes for obtaining the desired quality and productivity in EDM process [8]. Diaz-Garciaand and Bashiri (2014), modelled the multiple response problem as stochastic optimization and diverse solution [9]. Shadkam and Bijari(2015), used response surface methodology, DEA and Cuckoo algorithm for multiple response problem [2]. In addition to DEA, multiple responses can be made into a single response using the Technique for Order-Preference by Similarity to Ideal Solution (TOPSIS), PCA or fuzzy logic technique.

Response surface equations and surface numbers will increase as the number of responses increases. This situation will lead to complexity. Defining optimal inputs in multiple response models is difficult in situations where more than one output needs to be optimized at the same time. In this study, a combination of DEA and RSM was used to obtain a single response. DEA is an analysis method used to calculate the relative efficiency values of units using various inputs and outputs.

In our country Turkey, as in many other countries, the role of bread is crucial in meeting people's calorie, protein and B group vitamin requirements. Therefore, all the factors that influence the nutritional value of the bread are also important for the health of the society. Apart from wheat, bread is made from cereals such as corn, oats and rye. Phytic acid, which is abundant in corn bread,is known to cause some nutritional disorders. Therefore, the amount of phytic acid in the bread and the factors affecting it are gaining importance. In this method, experimental data are analyzed and responses expressing relative efficiency for each experiment are determined. The response surface methodology is used to perform the optimal parameter-setting based on relative activity. With a combination of multiple response problem, RSM and DEA, probing with a single response can be easily converted.

2. METARIALS AND METHODS

2.1. Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a non-parametric method based on linear programming which finds a wide range of application in both public and private sector establishments. By the use of this method it is possible to measure relative efficiencies of decision making units (DMU). DMU as they are named in literature refer to public or private sector establishments, companies, non-profit organisations, government programmes etc. where they are expected as part of their duties to extract some sort of outputs from a set of inputs. DEA was first used in the literature by Charnes et al. [10] in a study conducted in 1978. In this study, Charnes et al. [10] developed the CCR (Charnes Cooper Rhodes) model using the assumption of Constant Return to Scale (CRS), then Banker et al. [11] used the BCC (Banker Charnes Cooper) model using the Variable Return to Scale (VRS) they have developed. The results show how effective each DMU is in comparison to other DMU's [12].

Until now, DEA has been applied by various establishments active in different fields and for various purposes, such as for the purposes of making comparisons, evaluating management performances and measuring the efficiencies of relative resource use by public institutions active in the fields of health and education, by banking and related sectors, by manufacturing industry and its sub-sectors, by agriculture ,

transport and others sectors, by city administrations and regional development offices. Contrary to parametric methods, one important feature of DEA is that in production environments where many outputs are obtained using many inputs it enables us to make assessments without requiring the presence of any predetermined analytical production function.

In this study, the efficiency of DMUs was measured by using the input-direction CCR model based on fixed return assumption by scale. Here, *n* number of DMU$_j$,($j=1,2,...,k,...,n$) X_{ij}, ($i=1,...,m$) ith input and Y_{rj}, ($r=1,...,s$) rth output. Accordingly, kth under the assumption of constant fixed return for DMU, the input-scale CCR model is given in Equation (1)

$$E_k = \max \sum_{r=1}^{s} u_r Y_{rk}$$

$$\sum_{i=1}^{m} v_i X_{ik} = 1$$

$$\sum_{r=1}^{s} u_r Y_{rj} - \sum_{i=1}^{m} v_i X_{ij} \leq 0, j = 1,...,n,$$

$$v_i, u_r \geq \varepsilon, i = 1,...,m, r = 1,...,s. \tag{1}$$

Here, m number of inputs were used to generate s number of output kth DMU's relative activity: If $E_k=1$, kth DMU is active, or else $E_k<1$, kth DMU is not active comment is made.

2.2. Response Surface Methodology

RSM is a statistical technique that also includes optimization. As it was defined by Myers and Montgomery, the RSM as a method where both statistical and mathematical approaches and techniques are used together for the development and optimization of processes [13]. It is an approach used to model problems where various variables influence the response [14]. In RSM, a model is created with the help of regression analysis. The degree to which a factor's main effect or interaction effect has a significant effect on the value of the response variable is determined by means of regression coefficients. The first step in RSM is to determine the factors and levels that are thought to have an effect on the response variable. After this step, experiment design, regression and optimization techniques are used together.

The first order response surface equation for a linear model is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon, \tag{2}$$

where Y is response variable, X_i 's are independent variable and β_i is the regression coefficients, i=1,2,…,n.

A high-order polynomial approach is used for the response surface in determining the curvature of a response surface. Below is shown for a second-order model

$$\hat{y} = \hat{\beta}_0 + \sum_{g=1}^{n} \hat{\beta}_g x_g + \sum_{g=1}^{n} \hat{\beta}_{gg} x_g^2 + \sum_{g<f} \sum \hat{\beta}_{gf} x_g x_f . \tag{3}$$

Response surface graphs are graphs used in the three-dimensional representation of the effect of the over-all factor levels on the response variable. If there are too many factors, it is also possible to see the values of the main factors of interest by fixing the levels of other factors. When the model for the factors and response variables is a linear model, the response surface graph takes a linear shape, while a model with quadratic terms will have a curve shape.

Generally, in response surface problems, responses occur in three ways, that is, the nominal is the best, the bigger the better and the smaller the better. In "the smaller the better method", the goal is to reach the minimum response variable. In "the bigger the better method", the goal is to reach the maximum response

variable. In "the nominal is the best method", the response variable is the χ^2 (chi-square) value that gives the levels of the control variables making the target value. This method is slightly more complicated than the other two methods.

The aims of response surface studies are to predict the future values of the response variable accurately, to determine an appropriate function (or model), to identify the relationship between the response variable and the input variable, to search for the largest or smallest response value depending on the genre of the problem ,to determine the values of the input variables that this value can provide, and finally to contribute to expectations for understanding the mechanism underlying the system [15].

2.3. Optimization Procedure

In this study, a procedure that optimizes multiple responses was applied. First, a suitable experimental design was chosen. DEA and RSM were used for the analysis of the data. These process steps are illustrated in Fig. (1) where some details about the applied steps are also given.
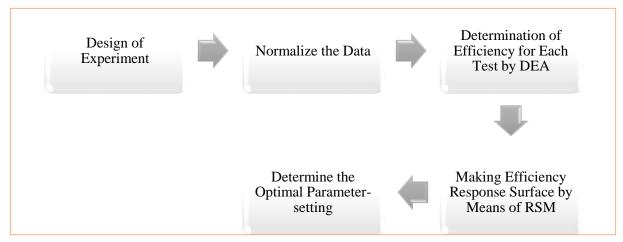

Figure 1. Steps to optimize data

Step 1. Design of Experiment

The researcher may choose control factors, response variables and quality features of response variables by his experience also taking into account the screening results. There are many different methods such as Central Composite Design (CCD), Face Centered Cube Design, Computer-Generated Design or Box-Behnken Design to design experiments. The most appropriate one was used to construct the experiment in this research. Response values were then calculated for each experiment.

Step 2. Normalizing the Data

In DEA, the experimental data is used to reduce the effect of the data in different measurement units. Euclidean norm was used in this study [2]. This calculation was shown in Equation (4).

$$y'_{ij} = \frac{y_{ij}}{\sqrt{\sum_{j=1}^{n} y_{ij}^2}} . \tag{4}$$

Step 3. Assessment of Efficiency for Each Test by DEA

Using the DEA, the relative efficiency values of the respective decision-making units are obtained using the inputs and outputs determined for the normalized data. In this paper, the relative efficiency values of the decision-making units were obtained using the CCR model under the assumption of fixed return on a scale basis.

Step 4. Establishment of Efficiency Response Surface by Means of RSM

The response surface of the activity was determined via the response surface methodology. A model fit for the lack of fit and p values was determined, a quadratic model was fitted and the response surface is given below:

$$\hat{e}_s = f(x) = \hat{y} = \hat{\beta}_0 + \sum_{g=1}^{n} \hat{\beta}_g x_g + \sum_{g=1}^{n} \hat{\beta}_{gg} x_g^2 + \sum_{g<f} \sum \hat{\beta}_{gf} x_g x_f. \tag{5}$$

RSM is an optimization algorithm which is an optimization response having two stages . In the first stage the optimization area is determined and the determination of the optimization response accomplished in the second stage using semi regression models [2].

Step 5. Determine the Optimal Parameter-setting

The optimum parameter is obtained by using the following equation with respect to the response surface of the action given in step 4

$$x_{\epsilon\Omega}^{max} \; \hat{e}_s = f(x), \tag{6}$$

where Ω is the feasible region.

3. RESULTS

In this study, wheat flour was mixed at various ratios of corn flour and corn bread was prepared by adding different proportions of yeast to these mixtures and by applying different fermentation and cooking times. The statistical analysis using the RSM was performed using the MINITAB 17 package program. The trial model was randomly generated. For DEA part, statistical analysis was done using GAMS 23.3 program.

The data set used in this study was obtained from the experimental study performed in Ankara University Food Engineering Laboratories. 81 loaves of corn bread has been used. The dataset consists of 4 inputs and 2 outputs. All these inputs and outputs were selected according to expert opinion The following four inputs were used for the analysis; wheat flour addition ratio (%) (x_1), yeast amount (x_2), (^{0}C) oven temperature (x_3) and fermentation time (x_4) (min). The variables, which were used as outputs, were the amount of phytic acid (mg / 100g) and loaf volume. The aim of the parameter optimization by the proper selection of input parameters was to achieve a uniformity that reduces the amount of phytic acid and at the same time increases the volume of the bread. Experimental responses were determined according to these measures. A Central Composite Design (CCD) was used to create the design of the experiment and the observations were then normalized according to the formula given in step 2. The data were then analyzed by means of the Charnes-Cooper-Rhodes (CCR) model as described in step 3 in data envelopment analysis. The efficiency values obtained in Table 1 were given.

Table 1. *Efficiency score values*

Bread	e_s	Bread	e_s	Bread	es	Bread	e_s	Bread	e_s	Bread	e_s
1	0.9964	16	1.0000	31	1.0000	46	1.0000	61	1.0000	76	0.9996
2	1.0000	17	0.9358	32	1.0000	47	1.0000	62	0.9991	77	0.9542
3	1.0000	18	0.9936	33	1.0000	48	1.0000	63	1.0000	78	0.9152
4	1.0000	19	0.9358	34	1.0000	49	0.9996	64	0.9965	79	1.0000
5	1.0000	20	0.9248	35	0.9998	50	0.9418	65	0.9939	80	0.9561

6	1.0000	21	0.8575	36	1.0000	51	0.8987	66	0.9964	81	0.9183
7	0.9936	22	1.0000	37	1.0000	52	1.0000	67	0.9991		
8	0.9358	23	0.9358	38	1.0000	53	0.9468	68	0.9459		
9	0.9358	24	0.9358	39	0.9963	54	0.8981	69	0.9085		
10	0.9358	25	1.0000	40	0.9936	55	1.0000	70	0.9994		
11	1.0000	26	1.0000	41	0.9358	56	1.0000	71	0.9483		
12	0.9418	27	0.9985	42	0.8859	57	1.0000	72	0.9083		
13	1.0000	28	1.0000	43	0.9977	58	1.0000	73	1.0000		
14	0.9183	29	0.9358	44	0.9412	59	0.9987	74	1.0000		
15	1.0000	30	0.9358	45	0.8936	60	1.0000	75	1.0000		

Following these steps, the response surface of the efficiency was created using the response surface methodology. Instead of obtaining two surfaces for each output, it is possible to obtain a single surface with the help of efficiency. Table 2 depicts the results obtained for the analysis of variance. When the values of F and p are considered, it is clearly seen that the application of linear and square models gives more meaningful results. However, since the objective was to select the most suitable model for the problem, the square model was chosen considering the mean square error (MSE) value.

Table 2. *Analysis of Variance for Efficiency*

Source	DF	Adj SS	Adj MS	F	P
Regression	14	0,036570	0,002612	5,70	0,001
Linear	4	0,016258	0,004065	8,87	0,001
Square	4	0,010224	0,002556	5,58	0,005
Interaction	6	0,010088	0,001681	3,67	0,017
Residual Error	16	0,007329	0,000458		
Total	30	0,043900			

Surface obtained by CCR efficiency:

$$e_s = 2{,}112 - 0{,}911x_1 + 2{,}94x_2 + 3{,}37x_3 - 24{,}3x_4$$
$$+ 1{,}303x_1^2 + 4{,}62x_2^2 + 8{,}37x_3^2 + 136{,}2x_4^2$$
$$+ 0{,}88x_1x_2 + 1{,}21x_1x_3 + 5{,}11x_1x_4$$
$$- 7{,}58x_2x_3 - 31{,}0x_2x_4$$
$$- 44{,}1x_3x_4. \tag{6}$$

Lastly, the optimal parameter-setting (30,6,120,230) was proposed when considering the equation created in step 5 for the parameter optimum. The response surfaces to be generated for each response using the conventional response surface methodology and based on the original experimental observations are shown as below:

$$Y_1 = 755 - 8{,}3x_1 + 4x_2 + 0{,}0x_3 + 3{,}8x_4$$
$$- 0{,}0181x_1^2 + 0{,}04x_2^2 - 0{,}0008x_3^2 - 0{,}0086x_4^2$$
$$- 0{,}052x_1x_2 - 0{,}0031x_1x_3 + 0{,}0007x_1x_4$$
$$- 0{,}014x_2x_3 - 0{,}025x_2x_4$$
$$+ 0{,}0002x_3x_4, \tag{7}$$

$$Y_2 = -44 + 2{,}52\,x_1 - 0{,}1\,x_2 - 0{,}02\,x_3 + 1{,}8\,x_4$$
$$- 0{,}0039x_1^2 - 0{,}032\,x_2^2 - 0{,}00003\,x_3^2 - 0{,}0037\,x_4^2$$
$$+ 0{,}075\,x_1{*}\,x_2 + 0{,}00358\,x_1{*}\,x_3 - 0{,}0009\,x_1{*}\,x_4$$
$$+ 0{,}0046\,x_2{*}\,x_3 + 0{,}003\,x_2{*}\,x_4$$

$$+ \ 0,0001 \ x_3 {}^* \ x_4. \tag{8}$$

Responses were obtained for the recommended the optimal parameter-setting. The methodology is very effective when multiple responses are optimized simultaneously.

4. DISCUSSION

When the place and the importance of bread in human life is considered, it is important to perform statistical analyzes in making healthy breads. In this study, a total of 81 loaves of corn bread were prepared using different ratios of corn and wheat flour, with different amounts of yeast and fermentation times, baked at different temperature conditions. Then the phytic acid contents and bread volume yields of these bread loaves were determined. Each loaf of bread with its specific corn flour/wheat flour ratio, amount of yeast used, fermentation time and baking time as the four input variables; the phytic acid content and bread volume yield as the two output values constituted one experimental data set. Since the amount of phytic acid in wheat flour is lower than that in corn flour, the proportion of phytic acid decreases as the amount of wheat flour increases in corn bread prepared from corn - wheat flour mixtures. When wheat flour is added to corn flour, bread volume increases as the amount of wheat flour in the mixture increases. Therefore, high-quality and healthy breads with high nutritional value and low levels of phytic acid can only be obtained by adding wheat flour to corn flour and by maintaining optimum process conditions during dough preparation and baking.

In the present study, the RSM was used in combination with DEA for the optimization of the process with multiple responses. The RSM enables one to obtain a process parameter setting through the regression equation without necessarily having to know the relation model between inputs and outputs. There are as many response equations as the number of responses, and therefore so many surfaces and contours can be drawn. For this reason, the solution of the problem can become complex by increasing the number of responses. There is more than one method to solve the multiple response problem. The appropriate method can be chosen by considering the relationship between the variables. In this study, DEA method was preferred .The DEA method has the ability to hold the multiplicity of inputs and outputs and an easy optimization technique to find the best alternatives [8]. Thus, the method employed in this study, which is a combination of DEA and RSM, is quite advantageous in that it saves time by eliminating the difficulty of calculating each response individually when compared to the conventional response surface methodology. The published studies adopting this methodology are quite rare in literature. Therefore, it is hoped that the present study makes a significant contribution to existing knowledge in the food industry.

CONFLICT OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1]   Tsai, C.W., Tong, L. and Wang, C., "Optimization of Multiple Responses Using Data Envelopment Analysis and Response Surface Methodology", Tamkang Journal of Science and Engineering, 13(2): 197–203, (2010).

[2]   Shadkam, E. and Bijari, M., "The Optimization of Bank Branches Efficiency by Means of Response Surface Method and Data Envelopment Analysis: A Case of Iran", Journal of Asian Finance, Economics and Business, 2(2): 13–18, (2015).

[3]   Myers, R.H., Montgomery, D.C., Vining, G.G., Borror, C.M. and Kowalski, S.M., "Response Surface Methodology: A Retrospective and Literature Survey", Journal of Quality Technology, 36: 53 –77, (2004).

[4]   Derringerand G.C. and Suich, R., "Simultaneous Optimization of Several Response Variables", Journal of Quality Technology, 12(1): 214–219, (1980).

[5]   Khuri, A.I. and Conlon, M., "Simultaneous Optimization of Multiple Responses Represented by Polynomial Regression Functions", Techno Metrics, 23: 363–375, (1981).

[6]   Tong, L.I. and Su, C. T., "Optimization Multi-Response Problems in Taguchi Method by Fuzzy Multiple Attribute Decision Making", Quality and Reliability Engineering International, 13: 25–34, (1997).

[7]   Köksoy, O., "Dual Response Optimization: The Desirability Approach", International Journal of Industrial Engineering: Theory, Applications and Practice, 12(4): 335–342, (2005).

[8]   Sahu, J., Mohanty, C.C. and Mahapatra, S.S., "A DEA Approach for Optimization of Multiple Responses in Electrical Discharge Machining of AISI D2 steel", Procedia Engineering, 51: 585-591, (2013).

[9]   Díaz-García, A.J. and Bashiri, M., "Multiple Response Optimization: An Approach from Multi Objective Stochastic Programming", Applied Mathematical Modelling, 38(7-8): 2015–2027, (2014).

[10] Charnes, A., Cooper, W.W. and Rhodes, E., "Measuring the Efficiency of Decision Making Units", European Journal of Operational Research, 2: 429–444, (1978).

[11] Banker, R.D., Charnes, A. and Cooper, W.W., "Some Models for Estimating Technical and Scale Efficiencies in Data Envelopment Analysis", Management Science, 30: 1078–1092, (1984).

[12] Kao, C. and Hwang, S.N., "Efficiency Decomposition in Two-stage Data Envelopment Analysis: An Application to Non-life Insurance Companies in Taiwan", European Journal of Operational Research, 185 (1): 418–429, (2008).

[13] Myers, R.H. and Montgomery, D.C., Response Surface Methodology, Process and Product Optimization Using Designed Experiments, USA: Second edition Wiley&Sons Publications, New York, (1995).

[14] Ryan, P., Modern Experimental Design. Canada: Wiley&Sons Publications, New Jersey, (2007).

[15] Bayrak, H., Özkaya, B. ve Tekindal, M.A., "Productivity in the First Degree for the Optimum Point Determination of Factorial Trials: An Application", Turkey Clinics Journal of Biostatistics, 2(1): 18–27, (2010).