Development of Design of Experiment Optimization to Obtain High-Quality Sugar

Yasemin Ayaz Atalan[1], Abdulkadir Atalan[2,*]

[1]Bozok University, Faculty of Engineering and Architecture, Yozgat, Turkey, yasemin.ayaz@bozok.edu.tr
[2]Gaziantep Islam Science and Technology University, Faculty of Engineering and Natural Sciences, Gaziantep, Turkey
Orcid: 0000-0001-7767-0342[1] Orcid: 0000-0003-0924-3685[2]
*Corresponding Author, e-mail: abdulkadir.atalan@gibtu.edu.tr

Abstract: This study aimed to use the design of experiment technique to calculate the optimum values of dependent and independent variables considered in the production of quality sugar. This study consists of two stages. Firstly, four different independent variables and one dependent variable data were used for this study. The levels and limits of independent factors were determined as a result of descriptive statistical analysis. They were considered the experiment as a two-level system; a full-factorial design of experiment, including 2^4 experiments with three replications, was made. Statistical analysis was performed using the data obtained, and it was determined whether the independent variables had an effect on the dependent variable. In the second stage of the study, the optimization model developed in order to obtain optimum results for both dependent and independent variables was run. According to the optimum data obtained in the optimization model, the quality-sugar score was calculated as 13.5821 according to 95% confidence and prediction interval, while the color of sugar was determined to be brown.

1. Introduction

This study is discussed on the quality of sugar that people consume a lot in daily life. Sugar consumption occurs in many types of food produced. For this reason, the quality of the sugar used in the quality of the produced product is discussed. In this study, the design of experiment method was preferred to increase the quality of the sugar produced. With this method, an optimization model was developed, and optimum values were obtained for the high-quality sugar score. The basis of sugar production is based on sugar cane, which was grown in different parts of the world thousands of years ago and is a vital foodstuff [1]. Sugar production in today's technology consists of two stages: refining the raw sugar obtained from sugar cane or sugar beets [2]. In sugar production, sugar canes or sugar beets must go through some chemical processes in order to turn into refined sugar [3], [4]. In this study, some of the chemical processes are considered independent variables.

The design of experiment is used by researchers in many fields, especially in studies that require experimentation [5]–[8]. In some studies, not the production of raw sugar, but the experimental design method was used to increase the yield from sugar beet, which is the raw material of raw sugar [9]. The design of experiment has been used in some studies to optimize sugar-related processes [10]. However, this study is directly aimed at increasing the quality of sugar using this method. Using this method is to have more than one variable in an experiment and measure the effects of these variables easily. The

method of experimental design consists of many stages. However, there are three most important processes. These processes are defined as randomization, replication, and blocking [11]. In this study, two processes other than blocking are discussed.

Another critical stage in the design of experiment is to determine the types and levels of variables to be used in the design. There are two types of variables or factors [12]. Independent variables and dependent variables are the cornerstones of the design of experiment [13]. It is one of the best methods developed to measure the effect of independent variables on dependent variables in the design of experiment [14], [15]. In this study, four different independent variables and one dependent variable are defined to calculate the optimum value of the quality sugar score. We have investigated the factors which may have possible effects on sugar quality. This experiment is designed in a factory producing sugar. The main objective is analyzing these factors and coming up with some recommendations of these factor levels to get the high sugar quality.

There are different techniques in the method of experimental design [16]. Some of these techniques are included in the literature as full-factorial design, fractional factorial design, response surface design (central composite design, Box-Behnken design), mixture design, screening design. The researchers preferred these designs depending on the types, levels, data numbers of the dependent and independent variables used for the experiment, and most importantly, the cost and time of the experiment [5]. If less data is needed in terms of cost and time, other design types are used instead of full-factorial design. For example, if a full-factorial design is used for an experiment with two levels and five independent factors, 32 data is needed, while a fractional factorial design requires 16 or 8 data. The full-factorial experiment design was used depending on the data and the level of factors obtained in this study.

This study includes four sections. The first section deals with the literature review of studies related to the design of experiments and sugar production. The second part gives detailed information about the methodology of the study. The results obtained from the method mentioned in the methodology section were discussed in the third section. In the last section, a conclusion about the research has been provided.

2. Materials and Methods

2.1. Design of Experiment

The design of experiment approach developed by Ronald Fisher in the early 1920s is widely used in many scientific fields that require experimentation [5], [17]. The central concept of the design of experiment is that it deals with the effect of independent variables on the dependent variable in detail [18]. In the design of experiment, the impact of independent variables on the dependent variable is examined depending on the levels of the variables.

There are three different steps to be considered in the design of experiment (Besides the principles of randomization, blocking, and replication [19], [20], other steps for this study are discussed):

- determination of dependent and independent variables: dependent variables are represented by the symbol y, while independent variables are defined with x. The number of independent variables is shown as d in this study.
- determining the independent variables' levels: the levels of the independent variables are represented by the symbol l.
- Determining the number of replications: The trials to be made in the experiments are done in a random order [17]. The notation n represents replication.

The design combination for an experiment is obtained using the following formula in the design of experiment:

$$the\ number\ of\ run = n.\,l^d \tag{1}$$

The above formula is valid for full-factorial design in the design of experiment and the same number of levels for each independent variable.

2.2. Selection of input and output variables and their Levels

Four independent variables and one dependent variable were chosen for this study because they play an important role in the production of quality sugar. Experts have determined these variables as factors that can affect the sugar quality level. While the independent variables were determined as absorption, conductivity, drying loss rate during the drying process, and sugar color, the sugar quality score was considered the dependent variable. The visual of white sugar and brown sugar is given in **figure 1**.

Figure 1. The visual of white and brown sugar [21]

With the design of experiment, the correct combination of the factor levels of different independent variables is obtained, and the constraint equations of the optimization model are formed in order to obtain the highest sugar quality. The dependent variable is accepted as the objective function of the optimization model developed for this study.

2.3. Full-Factorial Design of Experiment

There are four main factors to be considered in this experiment as described above. Each of the factors has two levels. Therefore, we have decided to apply for the 2^d —factorial design (d, number of factors). Accordingly, we have to consider a total of $2^4 - 1$ (= 15) effects which concerned the input factors and their interactions –second order, third order [22], and fourth-order.

We have focused on analyzing three replicated full factorial two-level designs, typically referred to as a 2^d factorial design (d factors, all crossed, with two levels each). Because we did three replicated experiments, we have just 48 runs. The coded and actual levels notations of decision variables are shown in **Table 1**.

Table 1. The coded and actual levels notations of decision variables

Factors	Notations	Level 1	Level 2
		Coded Level -1	Coded Level 0
		Value	Value
Absorbance	x_1	x_1^l	x_1^u
Conductivity	x_2	x_2^l	x_2^u
Dryness Loss	x_3	x_3^l	x_3^u
Color	x_4	x_4^l	x_4^u

MINITAB 18.0 statistical software was used for this study to statistically analyze the data and to optimize the quality of sugar score.

2.4. Optimization Model

In this study, an optimization model has been developed to calculate the optimum values of dependent and independent variables used for quality sugar production. In the optimization model, the objective function is determined as the sugar quality, while the constraints are considered the minimum and maximum values that the independent variables will take. The objective function is the regression

equation obtained as a result of experimental design analysis. The optimization model was created in the form of a closed formula as follows [18], [23]:

Objective function:

$$y_{max} \quad \beta_0 + \sum_i^n \beta_i x_i + \sum_{i<j} \beta_{ij} x_i x_j + \sum_{i<j<k} \beta_{ijkm} x_i x_j x_k + \sum_{i<j<k<m} \beta_{ijkm} x_i x_j x_k x_m + \in_i \tag{2}$$

s.t.

$$x_i^l \le x_i \le x_i^u, i = \{1,2, \dots, n\} \tag{3}$$

where x_i^l and x_i^u represent the lower and upper limits for the decision variables (input or independent factors). x_i symbolizes the optimum or feasible values that the decision variables will take. The regression equation representing the objective function is derived for four independent variables.

3. Results

For this study, four independent variables and one dependent variable were considered. Descriptive statistical data of the variables considered are given in **Table 2**. In particular, minimum and maximum values were calculated for each variable. With the acquisition of these values, the levels and limits of the independent variables are also determined.

Table 2: The descriptive statistical data for variables

Statistics	x_1	x_2	x_3	x_4	y
Mean	0.10600	20.0000	0.07450	0.0000	13.5030
SE Mean	0.00087	0.37900	0.00955	0.1460	0.00848
StDev	0.00606	2.62800	0.06619	1.0110	0.05870
Variance	0.00004	6.90400	0.00438	1.0210	0.00345
Sum of Square	0.54106	19524.5	0.47234	48.000	8752.68
Minimum	0.10000	17.4000	0.00900	-1.0000	13.3950
Q1	0.10000	17.4000	0.00900	-1.0000	13.4600
Median	0.10600	20.0000	0.07450	0.0000	13.5000
Q3	0.11200	22.6000	0.14000	1.0000	13.5500
Maximum	0.11200	22.6000	0.14000	1.0000	13.6180
Skewness	0.00000	0.00000	0.00000	0.0000	0.06000

For this study, the statistical analysis of the data of 48 scenarios (design, run) was made by using the full factorial design method. The statistical analysis's reliability values were calculated as 0.972 for R^2, 0.958 for adjusted R^2, and 0.936 for predicted R^2, respectively. The statistical analysis of decision variables was performed not only individually but also interactively in this analysis. ANOVA data of the decision variables are shown in **Table 3**.

Table 3: ANOVA results for decision variables

Status	Variables	Adj SS	Adj MS	F-Value	P-Value
Model	Model	0.157448	0.010497	72.530	0.001
Individual	x_1	0.000727	0.000727	5.0200	0.032

	x_2	0.028940	0.028940	199.96	0.001
	x_3	0.096571	0.096571	667.27	0.001
	x_4	0.029661	0.029661	204.95	0.001
2-way interaction	x_1x_2	0.000131	0.000131	0.9100	0.348
	x_1x_3	0.000264	0.000264	1.8300	0.186
	x_1x_4	0.000186	0.000186	1.2800	0.266
	x_2x_3	0.000000	0.000000	0.0000	0.992
	x_2x_4	0.000009	0.000009	0.0600	0.810
	x_3x_4	0.000186	0.000186	1.2900	0.265
3-way interaction	$x_1x_2x_3$	0.000096	0.000096	0.6700	0.421
	$x_1x_2x_4$	0.000037	0.000037	0.2600	0.616
	$x_1x_3x_4$	0.000058	0.000058	0.4000	0.532
	$x_2x_3x_4$	0.000571	0.000571	3.9500	0.056
4-way interaction	$x_1x_2x_3x_4$	0.000011	0.000011	0.0800	0.783

According to Table 2, it was observed that four different independent variables were effective on the dependent variable. As a result of the interaction of independent variables, it is understood that the interaction in which only the second, third, and fourth independent variables (3-way interaction, $x_2x_3x_4$) are together affects the dependent variable (p-value is 0.056). Since four independent variables are significantly effective on the dependent variable, no independent variable should be removed from the statistical model. However, interactive situations that are not statistically effective from independent variables should be removed from the model.

This study aims to use the design of experiment technique to calculate the optimum values of dependent and independent variables taken into account for producing quality sugar. For this reason, as a result of the statistical analysis, it was analyzed that four different independent variables were effective on the dependent variable. Optimum results were obtained by running the optimization model created on the condition of sticking to the data used in the next step of the study. The optimum results for dependent and independent variables are given in **Table 4.**

Table 4: The results of the optimization model and Confidence/Prediction Intervals

Decision or independent variables				Response	95% Confidence Interval	95% Prediction Interval
x_1	x_2	x_3	x_4	y		
0.1076	21.8546	0.029	+1*	13.5821	12.5985; 13.6268	12,5844; 13,6410

*Sugar color: -1: white, +1 Brown

According to Table 3, the optimum result was obtained by calculating the quality sugar score as 13.5821, according to the 95% confidence and prediction interval. In addition to the optimum results of the first three independent variables, the color of the sugar produced according to the fourth independent factor's optimum result should be brown in this research.

4. Conclusions

In this study, the design of experiment technique was used to increase the quality of sugar used in daily life. Four different independent variables and one dependent variable data were used in the study. As a result of descriptive statistical analysis, the levels and limits of the independent factors were determined. This study consists of two stages. First, by using these data, statistical analysis was performed, and it was determined whether the independent variables had an effect on the dependent variable. In the second stage of the study, the optimization model developed to obtain optimum results for both dependent and independent variables was run. According to the optimum data obtained in the optimization model, while the quality sugar score was calculated as 13.5821, according to the 95% confidence and prediction interval, and it was stated that the color of the sugar was brown.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, A.A. and Y.A.A.; methodology, A.A.; software, Y.A.A; validation, AA., and Y.A.A; formal analysis, A.A.; investigation, Y.A.A; resources, Y.A.A; data curation, A.A.; writing—original draft preparation, Y.A.A; writing—review and editing, Y.A.A; visualization, A.A.; supervision, A.A. All authors have read and agreed to the published version of the manuscript." Please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

References

[1]   R. A. BALLINGER, *A History of Sugar Marketing Through 1974*. 1978.

[2]   R. Singh, "Hybrid Membrane Systems – Applications and Case Studies," in *Membrane Technology and Engineering for Water Purification*, Elsevier, 2015, pp. 179–281.

[3]   A. L. Alarcón *et al.*, "Chemical characteristics and colorimetric properties of non-centrifugal cane sugar ('panela') obtained via different processing technologies," *Food Chem.*, vol. 340, p. 128183, Mar. 2021, doi: 10.1016/j.foodchem.2020.128183.

[4]   P. Borysiuk, I. Jenczyk-Tolloczko, R. Auriga, and M. Kordzikowski, "Sugar beet pulp as raw material for particleboard production," *Ind. Crops Prod.*, vol. 141, p. 111829, Dec. 2019, doi: 10.1016/j.indcrop.2019.111829.

[5]   A. Atalan and C. C. Dönmez, "Optimizing experimental simulation design for the emergency departments," *Brazilian J. Oper. Prod. Manag.*, vol. 17, no. 4, pp. 1–13, 2020, doi: 10.14488/BJOPM.2020.026.

[6]   N. A. Anika, N. Tanzeem, and H. Sen Gupta, "Design of Experiment (DoE): Implementation in Determining Optimum Design Parameters of Portable Workstation," *Engineering*, vol. 12, no. 01, pp. 25–32, 2020, doi: 10.4236/eng.2020.121002.

[7]   Y. Ayaz Atalan, M. Tayanç, K. Erkan, and A. Atalan, "Development of Nonlinear Optimization Models for Wind Power Plants Using Box-Behnken Design of Experiment: A Case Study for Turkey," *Sustainability*, vol. 12, no. 15, p. 6017, Jul. 2020, doi: 10.3390/su12156017.

[8]   B. Durakovic, "Design of experiments application, concepts, examples: State of the art," *Period. Eng. Nat. Sci.*, vol. 5, no. 3, Dec. 2017, doi: 10.21533/pen.v5i3.145.

[9]   I. K. Kapdan, F. Kargi, and R. Oztekin, "Effects of operating parameters on acid hydrolysis of ground wheat starch: Maximization of the sugar yield by statistical experiment design," *Starch - Stärke*, vol. 63, no. 5, pp. 311–318, May 2011, doi: 10.1002/star.201000158.

[10]   S. Both, J. Eggersglüß, A. Lehnberger, T. Schulz, T. Schulze, and J. Strube, "Optimizing Established Processes like Sugar Extraction from Sugar Beets - Design of Experiments versus

Physicochemical Modeling," *Chem. Eng. Technol.*, vol. 36, no. 12, pp. 2125–2136, Dec. 2013, doi: 10.1002/ceat.201300484.

[11] K. Hinkelmann, *Design and Analysis of Experiments*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2012.

[12] C. ÖRGEV, B. TOPAL, and H. ŞAHİN, "İşgörenlerin Demografik Özelliklerinin Örgütsel İklim Algısına Etkisinin İstatistik Analizi: Süreyya Paşa Göğüs Hastalıkları Hastanesi Örneği," *Eur. J. Sci. Technol.*, vol. 19, pp. 165–175, Aug. 2020, doi: 10.31590/ejosat.709785.

[13] O. Kempthorne, *The Design and Analysis of Experiments*. Wiley, 1952.

[14] A. Atalan, "THE IMPACTS OF HEALTHCARE RESOURCES ON SERVICES OF EMERGENCY DEPARTMENT: DISCRETE EVENT SIMULATION WITH BOX-BEHNKEN DESIGN," *PONTE Int. Sci. Res. J.*, vol. 75, no. 6, pp. 12–23, 2019, doi: 10.21506/j.ponte.2019.6.10.

[15] A. Atalan and C. Donmez, "Employment of Emergency Advanced Nurses of Turkey: A Discrete-Event Simulation Application," *Processes*, vol. 7, no. 1, p. 48, Jan. 2019, doi: 10.3390/pr7010048.

[16] R. P. Niedz and T. J. Evens, "Design of experiments (DOE)—history, concepts, and relevance to in vitro culture," *Vitr. Cell. Dev. Biol. - Plant*, vol. 52, no. 6, pp. 547–562, Dec. 2016, doi: 10.1007/s11627-016-9786-1.

[17] J. Antony, *Design of Experiments for Engineers and Scientists*. Elsevier, 2003.

[18] D. C. Montgomery, *Design and Analysis of Experiments*, 8th ed. Wiley, 2012.

[19] I. A. Rauf, "Applied Statistical Design of Experiments: Applications in Natural Sciences," *J. Phys. Chem. Biophys.*, vol. 5, no. 5, p. e129, 2015.

[20] G. W. Oehlert, *A First Course in Design and Analysis of Experiments*. Library of Congress Cataloging-in-Publication Data, 2010.

[21] Escali, "Brown Sugar and White Sugar: What's the Difference?," *Escali*, 2012. https://escali.com/blog/brown-sugar-vs-white-sugar-whats-the-difference/.

[22] S. Gora, G. Mustafa, J. K. Sahni, J. Ali, and S. Baboota, "Nanosizing of valsartan by high pressure homogenization to produce dissolution enhanced nanosuspension: pharmacokinetics and pharmacodyanamic study," *Drug Deliv.*, vol. 23, no. 3, pp. 930–940, Mar. 2016, doi: 10.3109/10717544.2014.923066.

[23] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 5th ed. Wiley, 2012.