The Effect of Sugar Concentration, Citric Acid Concentration and Storage Time on Orange Syrup’s pH Using Completely Randomized Factorial Design

Aulia Ishak¹, Khawarita Siregar², Asfriyati³ and Miftah Safira⁴

¹,²,⁴Industrial Engineering Department, Faculty of Engineering, Universitas Sumatera Utara, Medan, Indonesia
³Public Health Faculty, Universitas Sumatera Utara, Medan, Indonesia

E-mail: mitusafira12@gmail.com aulia.ishak@usu.ac.id

Abstract. One of the company in industrial product is orange syrup company. The quality of orange syrup can determine the success rate of company. Orange fruits contain acidic pH that can affect syrup quality depending on affecting factors. By using randomized factorial design, which consists of 3 factors, each factor consists of 2 factor levels, sugar concentration (50 gr and 70 gr), citric acid concentration (0.5 gr and 1 gr) and storage time (1 day and 3 days), this experiment was conducted using the axbxc fixed model with factorial 2x2x2 and 3 replications. Based on ANOVA table, there is a significant effect given by sugar concentration, citric acid concentration and storage time with $F$ result = 0.532 > $F$ table = 0.476, but there are no significant effect by conducted one or two factor only. The regression equation obtained is $Y = 3.317 + 0.002 X_1 + (-0.717) X_2 + (-0.046) X_3$. For the correlations coefficient this research concluded that sugar concentration with value of 0.095 gives very low effect, citric acid concentration with value of -0.813 gives very strong opposite effect and storage time with value of -0.208 gives very low opposite effect on orange syrup’s pH.

1. Introduction

Industry is an economic activity processes raw materials, semi-finished good and finished goods until can be used by customers. In this era, the company requires to be able to increase the competitiveness with other companies industrial products. One of the company that produces foodstuff is orange syrup. Syrup is a thick processed product that can be consumed as a thirst-releasing drink, in the form of concentrated preparations in water from sugar or sugar substitutes with or without additives flavor [1], contains 50%-65% of sucroses [2].

pH or degree of acidity is a very small concentration of hydrogen ions. pH is defined as a basic logarithm of -10 from the hydrogen ion concentration [3]. The pH content is needed by plants in a certain value to obtain the optimal growth. Excessive or insufficient pH can cause stunted plant growth. The degree of acidity has a change in value over a period of time. pH can also change depending on affecting factors.

Orange fruits are acidic which means have pH smaller than 7. The pH level of oranges can affect the quality of the syrup itself. Hence, conducted an experimental design on the amount of pH on orange syrup to determine whether the factors and the level of a given factor affect the pH to be obtained.
The treatments given are sugar concentration (50 gr and 70 gr), storage time (1 day and 3 days), and citric acid concentration (0.5 gr and 1 gr). The experimental unit on this research is Citrus Sinensis. Replication is needed to get more accurate estimation results for experimental error. This research will use Completely Randomized Factorial Experimental Design. This experimental design is conducted to determine which factor that give significant effect on the amount of pH on orange syrup. On the off chance that the consequences of QC tests can't satisfy the acknowledgment models, the aftereffects of examination of the entire arrangement of the estimations on that day must be eliminated or should be re-dissected, and an incomplete or full re-approval of the strategy considered [14].

2. Method

An experimental design is a design or experiment (with each step of the action that is truly defined) in such a way so that the results can be used to collect information related to or needed for the research to be conducted [4]. Experimental design methodology is an important way in modern research and development efforts [5]. Experimental design started with determining desired responses and will be followed with determining factors and levels and what kind of experimental design that will be used in this research [6]. One of the experimental design models is a completely randomized factorial design. In this model, the treatment position was randomized to all experimental material. This means that all trial units have an equal chance of receiving treatment, whereas in the other two basic designs there are limitations to randomization [7]. This model is widely used in research because of its simple way, but to note, this model can only be used if the experimental unit and the environment are homogen.

Data was collected based on the results of research on object with work procedures that had been determined at the time of conducting experiments. Objects for this study were 1440 grams of sugar and 18 grams of citric acid. The following are the steps of the experiment:

- Squeeze orange fruit with a squeezer to get the orange juice.
- Store orange juice in a bottle and put in the refrigerator.
- Pour 100 ml of orange juice into a beaker glass.
- Add sugar and citric acid based on factor level.
- Stir the mixture using a spoon.
- Measure orange syrup's pH using a pH meter.

Determination of the number of replications in this study was calculated using the formula:

\[ t (n-1) \geq 15 \]  

(1)

In this study, the collected data will pass several test for normality data by using Anderson Darling Test. Anderson Darling test is used to test whether the sample data comes from populations with certain distributions. Anderson Darling is a modification of the Kolmogorov Smirnov (KS) test. The difference between both of them is the critical values in the KS test is not depend on the particular distribution being tested while the Anderson Darling test utilizes certain distributions in calculating the critical values [8]. The advantage of this test is it allows more specific tests, but on the other side, the critical values must be calculated for each distribution. Then, determine homogeneity of variance for each factor and factor interactions by using Fisher Test and also using ANOVA based on fixed model because the level of the factor has a certain characteristic that can distinguish with other levels. The use of ANOVA depends on the questions asked and the prime sources of error. The mean sum of squares due to any given parameter or measurement is compared to another sum of squares due to errors [9]. It can determine which factor that give significant effect on the amount of pH on orange syrup by using Minitab software. Then, measure the regression equation, single and multiple coefficients by using SPSS to determine the level of closeness from each factors and level factors.
3. Results and Discussion
In this study, the results of measurements of orange syrup’s pH can be seen in Table 1.

Table 1. Result of measurement on orange syrup’s pH

| Storage Time | Sugar Concentration | Citric Acid Concentration |
|--------------|---------------------|---------------------------|
|              | 50 gram             | 70 gram                   |
|              | 0.5 gram            | 1.0 gram                  |
| 1 day        | 3.0 2.7             | 2.9 2.6                   |
|              | 3.0 2.6             | 3.0 2.7                   |
| 3 days       | 2.7 2.5             | 2.8 2.6                   |
|              | 2.9 2.6             | 3.1 2.8                   |
|              | 3.2 2.5             | 2.9 2.6                   |

Normality test data that has been obtained is used to test whether the data comes from populations following a normal distribution by using Anderson Darling Test. The normal distribution is the most important distribution in both theory and statistical application [10]. The normality data using Anderson Darling Test can be seen at Figure 1 below.

From the Figure 1, a line which is a parameter line and blue dots are data that will be tested for normality. Based on the distribution of these points and the p-value > α (0.108 > 0.05), it can be seen that the data is normally distributed.

Homogeneity test is a statistical test procedure to prove that two or more groups of sample data come from populations that have the same variance [11]. Homogeneity test for each factor is conducted by using Fisher test with confidence level 95% by using Minitab software. If the significant value is greater than 0.005, it can be said that the variants of two or more groups are the same [12]. In the comparison factor of sugar concentration, given 50 grams and 70 grams sugar concentration. The following graph results of homogeneity test on sugar concentration factor can be seen at Figure 2.
Figure 2. Homogeneity test graph of sugar concentration factor

In the comparison factor of citric acid concentration, given 0.5 grams and 1 gram citric acid concentration. The following graph results of homogeneity test on citric acid concentration factor can be seen at Figure 3.

Figure 3. Homogeneity test graph of citric acid concentration factor

In the comparison factor of storage time, given 1 day and 3 days. The following graph results of homogeneity test on citric acid concentration factor can be seen at Figure 4.
Figure 4. Homogeneity test graph of storage time factor

In Figure 1, Figure 2 dan Figure 3 it can be seen that the variance of the two factor level of each sugar concentration, citric acid concentration and storage time are homogen. The result of homogeneity factor interaction test results using Minitab software can be seen in Figure 5.

Figure 5. Factor interaction test graph

Based on the interaction factor test graph, it can be obtained that the interaction of sugar concentration and storage time gave effect on orange syrup’s pH.

This study used a Completely Randomized Design, which consists of 3 factors, sugar concentration, storage time, and citric acid, each factor consists of 2 factor levels. This experiment was conducted using the axbxc fixed model. ANOVA calculation with this fixed model is conducted to see
whether the treatments given in the comparison have a significant effect on pH. ANOVA table with factorial 2x2x2 and 3 replications can be seen at Table 2.

Table 2. ANOVA table for 2x2x2 factorial experiment calculation results with SPSS

| Source                | Type III Sum of Squares | df | Mean Square | F    | Sig. |
|-----------------------|-------------------------|----|-------------|------|------|
| Intercept             | 189.844                 | 1  | 189.844     | 9694.149 | 0.000 |
| Sugar                 | 0.010                   | 1  | 0.010       | 0.532 | 0.476 |
| CitricAcid            | 0.770                   | 1  | 0.770       | 39.340 | 0.000 |
| Storage               | 0.050                   | 1  | 0.050       | 2.574 | 0.128 |
| Sugar * CitricAcid    | 0.004                   | 1  | 0.004       | 0.191 | 0.668 |
| Sugar * Storage       | 0.004                   | 1  | 0.004       | 0.191 | 0.668 |
| CitricAcid * Storage  | 0.004                   | 1  | 0.004       | 0.191 | 0.668 |
| Sugar * CitricAcid * Storage | 0.010 | 1  | 0.010 | 0.532 | 0.476 |
| Error                 | 0.313                   | 16 | 0.020       |       |      |
| Total                 | 191.010                 | 24 |             |       |      |

Hypothesis that will be tested in this experiments are:

- $F$ result $< F$ table, $H_0$ is accepted, which means there is no significant effect given by the factor on orange syrup’s pH.
- $F$ result $> F$ table, $H_0$ is rejected, which means there is significant effect given by the factor on orange syrup’s pH.

The results obtained in this research are:

- $F$ sugar and citric acid $(0.191) < F$ table $(0.668)$, which means there is no significant effect given by sugar concentration and citric acid concentration on orange syrup’s pH.
- $F$ sugar and storage $(0.191) < F$ table $(0.668)$, which means there is no significant effect given by sugar concentration and storage time on orange syrup’s pH.
- $F$ citric acid and storage $(0.191) < F$ table $(0.668)$, which means there is no significant effect given by citric acid concentration and storage time on orange syrup’s pH.
- $F$ sugar, citric acid and storage $(0.532) > F$ table $(0.476)$, which means there is significant effect given by sugar concentration, citric acid concentration and storage time on orange syrup’s pH.

From the results above we can conclude that only sugar concentration, citric acid concentration and storage time can give significant effect on orange syrup’s pH.

The pH results obtained will be input to get the calculation of the regression equation by using SPSS. The result can be seen at Table 3.

Table 3. Regression equation with SPSS

| Model     | Unstandardized Coefficients | Standardized Coefficients | t    | Sig. |
|-----------|-----------------------------|---------------------------|------|------|
|           | B       | Std. Error | Beta |       |       |
| (Constant)| 3.317   | 0.187      | -    | 17.755 | 0.000 |
| Sugar     | 0.002   | 0.003      | 0.095 | 0.789 | 0.440 |
| CitricAcid| -0.717  | 0.106      | -0.813 | -60.782 | 0.000 |
| Storage   | -0.046  | 0.026      | -0.208 | -10.735 | 0.098 |

The conclusions from Table 3 are constants, the coefficients X1, X2, and X3 in the regression equation are $3.317$, $0.0021$, $-0.7167$, and $-0.0458$. Thus, the regression equation is as follows:

$$Y = 3.317 + 0.002 X1 + (-0.717) X2 + (-0.046) X3$$

(2)
Regression linearity test is conducted to determine whether the regression line obtained is linear or not linear. Hypothesis that will be tested in this experiments are:

- \( F \) result < \( F \) table, \( H_0 \) is accepted, which means regression line is not linear.
- \( F \) result > \( F \) table, \( H_0 \) is rejected, which means regression line is linear.

A simple linear regression model is used to model the linear relation between the independent variable \( X \) and the dependent variable \( Y \).

Calculation results of regression linearity test by using SPSS can be seen in the Table 4 below.

| Model   | Sum of Squares | df | Mean Square | \( F \) | Sig. |
|---------|----------------|----|-------------|-------|-----|
| Regression | 0.831          | 3  | 0.277       | 16.542 | 3.0280 |
| Residual | 0.335          | 20 | 0.017       |        |      |
| Total   | 1.166          | 23 |             |        |      |

From the Table 4 above, it obtained that the value of \( F \) result (16.542) > \( F \) table (3.0280), \( H_0 \) is rejected which means regression line is linear. line Input the data obtained to calculate correlation coefficient. The higher correlation coefficient between two variables (the closer to 1), the higher level of closeness of the relationship between the two variables. Otherwise, the lower correlation coefficient between the two types of variables (the closer to 0), the lower level of closeness of the relationship between the two variables. The level of correlation value categories can be seen at Table 5.

| No. | Correlation Value | Category     |
|-----|-------------------|--------------|
| 1.  | 0                  | No Correlation |
| 2.  | >0 – 0.25          | Very Low     |
| 3.  | >0.25 – 0.5        | Moderate     |
| 4.  | >0.5 – 0.75        | Strong       |
| 5.  | >0.75 – 0.99       | Very Strong  |
| 6.  | 1                  | Perfect Correlation |

The result of single correlation coefficient can be seen at Table 6.

|          | Sugar | CitricAcid | Storage | pH    |
|----------|-------|------------|---------|-------|
| Sugar    | Pearson Correlation 1 | 0.000 | 0.000  | 0.095 |
|          | Sig. (2-tailed)       | 1.000 | 1.000  | 0.660 |
|          | N                  | 24    | 24     | 24    |
| CitricAcid | Pearson Correlation | 0.000 | 1      | -0.813** |
|          | Sig. (2-tailed)       | 1.000 | 1.000  | 0.000 |
|          | N                  | 24    | 24     | 24    |
| Storage  | Pearson Correlation  | 0.000 | 0.000  | 1     |
|          | Sig. (2-tailed)       | 1.000 | 1.000  | 0.330 |
|          | N                  | 24    | 24     | 24    |
| pH       | Pearson Correlation  | 0.095 | -0.813** | -0.208 | 1 |
|          | Sig. (2-tailed)       | 0.660 | 0.000  | 0.330 |
|          | N                  | 24    | 24     | 24    |
The results obtained from the Table 6 are:

- On sugar concentration coefficient correlation, the r value = 0.095, which means it given very low effect on the amount of orange syrup’s pH.
- On citric acid concentration coefficient correlation, the r value = -0.813, which means it given very strong opposite effect on the amount of orange syrup’s pH.
- On storage time coefficient correlation, the r value = -0.208, which means it given very low opposite effect on the amount of orange syrup’s pH.

Multiple correlation is the correlation of several independent variables (X) simultaneously with the dependent variable (Y). The result of multiple correlation can be seen at Table 7.

| Model | R    | R Square | Adjusted R Square | Std. Error of the Estimate | R Square Change | Change Statistics | F Change | df1 | df2 | Sig. F Change |
|-------|------|----------|-------------------|-----------------------------|-----------------|-------------------|----------|-----|-----|--------------|
| 1     | 0.844a | 0.713     | 0.670             | 0.12942                     | 0.713           | 16.542            | 3        | 20  |     | 0.000        |

From the Table 7 above, it can concluded that the r value = 0.844, which means sugar concentration, citric acid and storage time was given very strong effect on the amount of orange syrup’s pH.

4. Conclusion
The quality of orange syrup can determine the success rate of a company. Related to this, the company needs to find factors that can give significant effect to the orange syrup’s pH. The data obtained is normally distributed after conducted normality test by using Anderson Darling test with p-value = 0.108. Based on the ANOVA table, there is significant effect given by sugar concentration, citric acid concentration and storage time factor. There are no significant effect by given one or two factor at the same time.

The regression equation used is multiple linear regression, the regression equation obtained is $Y = 3.317 + 0.002 \times X1 + (-0.717) \times X2 + (-0.046) \times X3$. From the calculation of the regression equation obtained for relation among sugar concentration (X1), citric acid concentration (X2), and storage time (X3) on orange syrup’s pH. The regression line is linear.

Correlation coefficient calculations obtained from the three factors are correlation coefficient calculation between sugar concentration (X1) and pH (Y) obtained the value of 0.095, which means it given very low effect on the amount of orange syrup’s pH, correlation coefficient calculation between citric acid concentration (X2) and pH (Y) obtained the value of -0.813, which means it given very strong opposite effect on the amount of orange syrup’s pH and correlation coefficient calculation between storage time (X3) and pH (Y) obtained the value of -0.208, which means it given very low opposite effect on the amount of orange syrup’s pH.

Acknowledgment
The author would like to thank the University of Sumatera Utara who has provided financial and site support to publish this paper. The author appreciates all the support system for the effort and contribution for this experiment according to this study.

References
[1] Ansel H et al 2005 Ansel Farmacutical Dosage Form and Drug Delivery System Eight Edition (Philadelphia: Lippincott Williams and Wilkins a Watters Kluver Company)
[2] Mun’im et al 2012 Fitoterapi Dasar (Jakarta: Dian Rakyat)
[3] Kohlmann, F J 2003 What is pH and How is it Measured USA: Hach Company
[4] Sudjana 1994 Desain dan Analisis Eksperimen (Surabaya: ITS) pp 1-2
[5] Hanrahan G and Lu K 2006 Application of Factorial and Response Surface Methodology in Modern Experimental Design and Optimization *Critical Reviews in Analytical Chemistry* **36** (3-4) pp 141–151

[6] Yuwana A R and Hadi Y 2019 Experimental Design To Analyze The Effect Of Writing Format On Student Reading Comprehension On Paper Media Using Completely Randomized Factorial Design *Journal of Engineering and Management in Industrial System* **6** (2) pp 95-104

[7] Sugandi E and Sugiaro 1994 *Rancangan Percobaan* (Yogyakarta: Andi Offset) pp 21-22

[8] Fallo J *et al* 2013 Uji Normalitas berdasarkan Metode Anderson Darling Cramer-Vonmises dan Lilliefors Menggunakan Metode Bootstrap pp 152-153

[9] Allus M A *et al* 1989 The Use of Experimental Design, Multilinear Regression, ANOVA, Confidence Bands and Leverage in A Study of The Influence of Metals on the Growth of Barley Seedlings *Chemometrics and Intelligent Laboratory Systems* **6**(1) pp 65–80

[10] Montgomery D C 1997 *Design and Analysis of Exsperiments* (New York: John Wiley & Sons) pp 81-82

[11] Nuryadi *et al* 2017 *Dasar Dasar Statistik Penelitian* (Yogyakarta: Sibuku Media) pp 89-90

[12] Priyatno D 2009 *5 Jam Belajar Olah Data dengan SPSS* (Yogyakarta: Elex Media Komputindo)

[13] Walpole R E 2005 *Pengantar Statistika* (Jakarta: PT Gramedia Pustaka Utama) p 369

[14] Indrayanto G 2018 Recent development of quality control methods for herbal derived drug preparations *Natural Product Communications*, **13** (12)