Analysis of factors affecting strawberry farming (case: Dolat Rayat District, Karo Regency)

N F Pinem*, A R Pohan and R Gunawan
Agribusiness Department, Faculty of Agriculture, Universitas Sumatera Utara, Medan, Indonesia.

E-mail: *nurulfajriahpinem@usu.ac.id

Abstract. This study aims to analyse the factors affecting strawberry farming in Dolat Rayat District. The analytical method used in finishing the study is the multiple linear regression model through the Cobb Douglas production function. The variables used in this study were land area, use of seedlings, use of NPK fertilizer, total labours and use of Joker pesticide. The results obtained indicate that land area, use of seedlings, use of NPK fertilizer, total labours and use of Joker pesticide together have a significant effect on strawberry production with a coefficient of determination of 0.699. Partially, land area, labours and the use of Joker pesticide did significantly influence the strawberry production. Meanwhile, the use of seedlings and the use of NPK fertilizer did not significantly affect the strawberry production.

1. Introduction
Strawberries are one of the most important fruit commodities in the world, especially for countries with subtropical climates. Consumer demand for strawberries tends to increase from year to year. The higher market absorption rate reflects that the strawberry agribusiness has bright prospects in the future. Based on Agricultural Statistics in 2017, strawberry production in Indonesia has decreased every year. Strawberry production reached 169,796 tons in 2012, 90,352 tons in 2013, 58,882 tons in 2014, 31,798 tons in 2015 and 12,091 tons in 2016 [1].

The high demand for strawberries is inversely proportional to the lower production of strawberries, thereby increasing the import of strawberries to Indonesia. To prevent the increasing in strawberry imports, some efforts should be made to increase the production of strawberries in Indonesia. The authors will analyse factors affecting the production of strawberry farming.

2. Research methods
To analyse the effect of several factors on the farmers’ strawberry farm production in the study area, multiple linear regression models were used, where previously classical assumptions were tested in order to meet the requirements of the econometric criteria such as normality, multicollinearity and heteroscedasticity tests.

2.1. Multiple linear regression model
To determine the factors that affecting the production of strawberries, the model used is a multiple linear model through the Cobb-Douglas production function. The formula of the production function in the linear model mathematically is as follows:
\[ Y = f(x_1, x_2, x_3, x_4, x_5) \]  

(1)

In order to estimate the parameters used, which were initially in a multiple linear model, the formula must be transformed first into a natural double logarithm (ln) by using the ordinary least square method as formulated as follows:

\[ \ln Y = \ln \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \beta_3 \ln x_3 + \beta_4 \ln x_4 + \beta_5 \ln x_5 + \epsilon \]  

(2)

Information:
\[ \begin{align*}
Y & = \text{Strawberry Production (kg)} \\
X_1 & = \text{Land Area (m}^2) \\
X_2 & = \text{Seedlings (trunk)} \\
X_3 & = \text{NPK Fertilizer (kg)} \\
X_4 & = \text{Labours (People Working Day)} \\
X_5 & = \text{Joker Pesticide (litre)} \\
\ln & = \text{Logarithm} \\
e & = \text{Error Term} \\
0 & = \text{Constant} \\
\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_n & = \text{Independent Variable Coefficient}
\end{align*} \]

2.2. Ordinary least square (OLS) assumption test

OLS linear regression is a model of linear regression with the method of calculating the smallest square. In this regression model, there are several requirements that must be met in order for the forecasting model made to be valid as a forecasting tool. When these conditions are met, then the linear regression model is said to be BLUE. BLUE stands for Best Linear Unbiased Estimation.

2.2.1. Multicollinearity test. The multicollinearity test is intended to avoid linear relationships between independent variables. Multicollinearity can be detected by several methods, among them are by looking at:
1. If the value of the determination coefficient (R²) is high; in the simultaneous test (F-test), exogenous variables simultaneously have a significant effect on endogenous variables; but in the partial test (t-test), many exogenous variables have no real effect on endogenous variables, so this indicates the occurrence of multicollinearity.
2. View the standard error values. Large standard error values indicate the occurrence of multicollinearity.
3. If the VIF (Variance Inflation Factor) or the value of Tolerance is less than 0.1 or the value of VIF exceeds 10 indicates the occurrence of multicollinearity.
4. There is a simple correlation coefficient that reaches or exceeds 0.8 if the F-count value exceeds the F table of regression between independent variables.

2.2.2. Heteroscedasticity test. The way to detect the occurrence of heteroscedasticity in the regression model with the SPSS Program is by using the graph analysis. The way to do the graph analysis is by
2.2.3. Normality test. The use of the normality test is to determine whether the data distribution approaches or follows the normal distribution. In principle, normality can be discovered by looking at the data distributions or the points on the diagonal axis of the graph or by looking at the histogram of the residual. If the data is spreading around the diagonal line or histogram graph, it means the regression model has met the assumption of normality [2].

2.3. Model compatibility test

2.3.1. Determinant test (R²). The determination coefficient (R²) is a statistical value calculated from the sample data. This coefficient indicates the percentage variation of all dependent variables that can be identified by looking at the changes in independent variables (explanatory variables). The coefficient is a measure of the extent to how far an independent variable can change a dependent variable in a relationship [3].

2.3.2. Simultaneous influence of variable test (F test). The F test is a test that is used to show the simultaneous influence of the related variables in accordance with all independent variables used in the model. With the following criteria:

a. If calculated F < F statistics or if the significance of F > α: H₀ is accepted.

b. If calculated F > F statistics or if the significance of F < α: H₁ is accepted

Information:
H₀ = variables of land area, labours, seedlings, NPK fertilizer and Joker pesticide simultaneously have no real effect on strawberry production
H₁ = variables of land area, labours, seedlings, NPK fertilizer and Joker pesticide simultaneously have a real effect on strawberry production

2.3.3. Partial variable influence test (t test). The T test is a test used to show how far the influence of an independent variable is partially in explaining the variation of dependent variables. With the following test criteria:

a. If calculated t < t statistics or if the significance of t > α: H₀ is accepted.

b. If calculated t > t statistics or if the significance of t < α: H₁ is accepted

Information:
H₀ = variables of land area, labours, seedlings, NPK fertilizer and Joker pesticide simultaneously have no real effect on strawberry production
H₁ = variables of land area, labours, seedlings, NPK fertilizer and Joker pesticide simultaneously have a real effect on strawberry production

3. Results and discussion

The number of samples used in this study is 30 samples of strawberry farmers. This is directly proportional with Agung Statement said that the central limit theorem has been applied to a sample size of at least 30. Even stated for a sample size greater than 20, the normal distribution has been used to approach the binomial distribution [4].

The production function used to look at the factors that influence strawberry farming is the Cobb-Douglas function which consists of dependent variables and independent variables. The dependent variables are strawberry production (Y) and its independent variables are land area (X₁), seedlings
(X_2), NPK fertilizer (X_3), labours (X_4) and Joker pesticide (X_5). Based on the analysis that has been done, that the result of the Kolmogorov-Smirnov test is obtained the significance value of 0.200 > 0.05 (α). It can be deduced that the data used are distributed normally. This is also in accordance with the approximation graph, the display of P-P normal plot of regression standardized residual of data is said to be normally distributed plots depicted data spread or docked to the diagonal line.

![Normal P-P plot of regression standardized residual](image1)

Figure 1. Normal P-P plot of regression standardized residual

Multicollinearity test is used to see whether there is a great linear correlation between some or all of the variables that included in the regression model. The existence of multicollinearity can be identified from the VIF (Variance Inflection Factor) value or the tolerance value. Multicollinearity does not occur when the VIF value is smaller than 10 and also the tolerance value is larger than 0.1. The VIF values obtained are 2.311, 4.003, 1.865, 1.662 and 2.695. While the tolerance values obtained were 0.433, 0.250, 0.536, 0.602, and 0.371 respectively. This indicates that the data used has met the requirements of the multicollinearity test. This clearly proves that there is no multicollinearity in the regression model. The heteroscedasticity test can also be done by looking at the distribution pattern of the dots on the diagram scatterplot. Based on Figure 2, it can be seen the dots are spreading out without forming any certain clear pattern. Therefore, it can be concluded that there are no symptoms of heteroscedasticity.

![Scatterplot](image2)

Figure 2. Scatterplot

The R^2 value is 0.699 based on the analysis of the coefficient of determination (R^2) has been done. It means variations in bean production can be explained by variations in production factors of 69.9%. This means that the ability of the independent variables to be included in the production function model, namely land area, seedlings, NPK fertilizer, labours and Joker pesticide has a considerable
effect on increasing strawberry production. The remaining percentage (30.1%) is influenced by other factors such as natural conditions, pests or diseases.

Based on the results of the F test analysis, the calculated F (11.141) > F table (3.75) and a significant value of 0.000, at the 95% confidence level and error rate α = 5% so that the use of production factors of land area, seedlings, NPK fertilizer, labours and Joker pesticide together have a significant effect on strawberry production.

Table 1. Significance test results of strawberry farming production factors coefficient

| Variable            | Regression Coefficient | T-Test | Significance |
|---------------------|------------------------|--------|--------------|
| Constant            | -334.303               | -0.854 | 0.401        |
| Land Area (X1)      | 0.316                  | 4.121  | 0.000        |
| Seedlings (X2)      | -0.037                 | -1.779 | 0.088        |
| NPK Fertilizer (X3) | 0.083                  | 0.102  | 0.920        |
| Labours (X4)        | 8.095                  | 2.362  | 0.027        |
| Joker Pesticide (X5)| 55.714                 | 3.373  | 0.003        |

Source: Primary Data Processing, 2020

Based on Table 1, the results of the analysis that influence the amount of use of chemical Joker pesticide are included in the equation for the use of Cobb Douglas production inputs as follows:

\[
\ln Y = 334.303 + 0.316 \ln x_1 - 0.037 \ln x_2 + 0.083 \ln x_3 + 8.095 \ln x_4 + 55.714 \ln x_5 + \epsilon \tag{3}
\]

Information:

- **Y** = Strawberry Production (kg)
- **X1** = Land Area (m²)
- **X2** = Seedlings (trunk)
- **X3** = NPK Fertilizer (kg)
- **X4** = Labours (People Working Day)
- **X5** = Joker Pesticide (litre)

3.1. Land area (X1)

The results of the partial test analysis that have been carried out show that the coefficient value of land area is 0.316. The coefficient of power X1 (0.316) shows the magnitude of the effect of land area on strawberry production. If X1 (land area) increases by 1%, then Y (strawberry production) will increase by 0.316%. Conversely, if X1 (land area) decreases by 1%, then Y (strawberry production) will decrease to 0.316%. The significance value of t is X1 (0.000) < α (0.05). This shows that H1 is accepted and H0 is rejected. This means that the independent variable of land area has a significant effect on changes in strawberry production.

3.2. Seedlings (X2)

The results of the partial test analysis that have been carried out show that the coefficient of seedlings is -0.037. The coefficient of power X2 (-0.037) shows the magnitude of the effect of the seedlings on strawberry production. If the use of X2 (seedlings) increases by 1%, then Y (strawberry production) will decrease by 0.037%. Conversely, if the use of X2 (seedlings) decreases by 1%, then Y (strawberry production) will increase to 0.037%. The significance value of t is X2 (0.088) > α (0.05). This shows that H0 is accepted and H1 is rejected. This means that the independent variable use of seedlings has no significant effect on changes in strawberry production.
3.3. NPK fertilizer ($X_3$)
The results of the partial test analysis that have been carried out show that the coefficient of NPK fertilizer is 0.083. The coefficient of power $X_3$ (0.083) shows the magnitude of the effect of NPK fertilizer on strawberry production. If the use of $X_3$ (NPK fertilizer) increases by 1%, then $Y$ (strawberry production) will increase by 0.083%. Conversely, if the use of $X_3$ (NPK fertilizer) decreases by 1%, then $Y$ (strawberry production) will decrease to 0.083%. The significance value of $t$ is $X_3$ (0.920) > $\alpha$ (0.05). This shows that $H_0$ is accepted and $H_1$ is rejected. This means that the independent variable use of NPK fertilizer has no significant effect on changes in strawberry production.

3.4. Labours ($X_4$)
The results of the partial test analysis that have been carried out show that the labours coefficient value is 8.095. The coefficient of power $X_4$ (8.095) shows the magnitude of the influence of labours on strawberry production. If the use of $X_4$ (labours) increases by 1%, then $Y$ (strawberry production) will increase by 8.095%. Conversely, if the use of $X_4$ (labour) decreases by 1%, then $Y$ (strawberry production) will decrease to 8.095%. The significance value of $t$ is $X_4$ (0.027) < $\alpha$ (0.05). This shows that $H_1$ is accepted and $H_0$ is rejected. This means that the independent variable of use of labours has a significant effect on changes in strawberry production.

3.5. Joker pesticide ($X_5$)
The results of the partial test analysis that have been carried out show that Joker pesticide coefficient value is 55.714. The coefficient of power $X_5$ (55.714) shows the magnitude of the influence of Joker pesticide use on strawberry production. If the use of $X_5$ (Joker pesticide) increases by 1%, then $Y$ (strawberry production) will increase by 55.714%. Conversely, if the use of $X_5$ (Joker pesticide) decreases by 1%, then $Y$ (strawberry production) will decrease to 55.714%. The significance value of $t$ is $X_5$ (0.003) < $\alpha$ (0.05). This shows that $H_1$ is accepted and $H_0$ is rejected. This means that the independent variable of use of Joker pesticide has a significant effect on changes in strawberry production.

4. Conclusions
The variables used in this study were land area, use of seeds, use of NPK fertilizer, total labours and use of Joker pesticide. The land area, use of seedlings, use of NPK fertilizer, total labours and use of Joker pesticide together have a significant effect on strawberry production with a coefficient of determination of 0.699. Partially, land area, labours and use of Joker pesticide have a significant effect on strawberry production. Meanwhile, the use of seedlings and the use of NPK fertilizer did not significantly affect the strawberry production. This explains that help from government in the form of capital and provision of production facilities to increase production in strawberry farming is still very much needed.

References
[1] [BPS] Badan Pusat Statistik [BPS-Indonesian Statistics] 2017 Produksi Buah-buahan di Indonesia (2012-2016) [Fruit Production in Indonesia (2012-2016)] (Jakarta, Indonesia: Agricultural Statistics)
[2] Nachrowi and Hardius U 2002 Penggunaan Teknik Ekonometrika [Using of Econometric Techniques] (Jakarta, Indonesia: Rajawali Press)
[3] Syafirizal H S 2010 Data Penelitian Menggunakan Program SPSS [Research Data Using SPSS Program] (Medan, Indonesia: USU Press)
[4] Supriana T 2016 Metode Penelitian Sosial Ekonomi [Socio-Economic Research Method] (Medan, Indonesia: USU Press)