Jurnal Riset Rumpun Matematika dan Ilmu Pengetahuan Alam (JURRIMIPA)
Vol.2, No.1 April 2023
e-ISSN: 2828-9390; p-ISSN: 2828-9382, Hal 61-75

Model Estimasi Produksi Padi Menggunakan Analisis Regresi Data Panel Di Provinsi Sumatera Utara Berdasarkan Data Tahun 2015-2019

Estimation Model Of Rice Production Using Panel Data Regression Analysis In North Sumatra Province Based On Data For 2015-2019

Fuji Hidayah
Universitas Sumatera Utara
Email : fujihidayah159@gmail.com

Mardiningsih
Dosen Universitas Sumatera Utara
Email : mardiningsih.math@gmail.com

Abstract. *Analysis of panel data regression is a regression method to determine the relationship between the independent variable and the dependent variable using combined data, namely between cross-section data and time-series data.Analysis of panel data regression can apply to the processing of rice production estimates for an area including the province of North Sumatera. North Sumatera is one of the provinces in Indonesia, at the very of the population focuses on the agricultural sector as the main livelihood in rice being the primary production. Based on data on the development of rice production at the Agriculture Office of North Sumatra Province, the level of rice production in 2019 for North Sumatra Province reached 4,693,563 Ton/Year.Throught Panel Data Regression Analysis with the three approach methods is Common Effect Model (CEM), Fixed Effect Model (FEM) and Random Effect Model (REM) , the best approach method is obtained, namely the Fixed Effect Model (FEM) with the best estimation model is* $\hat{Y} = 2{,}311909 + 0{,}950690X_{it} - 0{,}032532X_{it}$. *The coefficient of determination* (R^2) *of 0.950890, which means that the independent variable affects the dependent variable by 95.08% on rice production.Based on the estimation model, the development of rice production in the next 3 years, namely 2020, 2021 and 2022, has changes of around 39.88% (decrease), 10.68% (increase), 9.89% (increase).*

Keywords: *Rice Production, Analysis of Panel Data Regression, Estimation, Common Effect Model, Fixed Effect Model, Random Effect Model.*

Abstrak. Analisis regresi data panel merupakan metode regresi untuk mengetahui hubungan antara variabel independen terhadap variabel dependen dengan menggunakan data gabungan yaitu antara data cross section dan data time series. Analisis regresi data panel dapat diterapkan pada pengolahan estimasi produksi padi terhadap suatu wilayah termasuk provinsi Sumatera Utara. Sumatera Utara merupakan salah satu provinsi di

Received Febuari 01, 2023; Revised Maret 02, 2023; April 01, 2023
* Fuji Hidayah, fujihidayah159@gmail.com

Indonesia sebagian besar penduduknya berfokus pada sektor pertanian sebagai mata pencaharian utama dengan tanaman padi yang merupakan hasil produksi utama. Berdasarkan data perkembangan produksi padi pada Dinas Pertanian Provinsi Sumatera Utara, nilai rata-rata produksi padi pada tahun 2019 untuk Provinsi Sumatera Utara mencapai146.647 Ton. Melalui Analisis Regresi Data Panel dengan ketiga metode pendekatan yang meliputi Common Effect Model (CEM), Fixed Effect Model (FEM) dan Random Effect Model (REM) diperoleh pendekatan terbaik yaitu Fixed Effect Model (FEM) dengan model estimasi terbaiknya adalah $\hat{Y} = 2{,}311909 + 0{,}950690X_{it} - 0{,}032532X_{it}$. Dengan koefisien Determinasi (R^2) sebesar 0,950890 yang berarti variabel independen berpengaruh terhadap variabel dependen sebesar 95,08% pada produksi padi. Berdasarkan model estimasi diperoleh perkembangan produksi padi pada 3 tahun berikutnya yaitu tahun 2020, 2021 dan 2022 mengalami perubahan berturut-turut sekitar 39,88% (penurunan), 10,68% (peningkatan), 9,89% (peningkatan).

Kata kunci: Produksi Padi, Analisis Regresi Data Panel, Estimasi, Common Effect Model, Fixed Effect Model, Random Effect Model.

INTRODUCTION

Rice production is the main livelihood of the population in North Sumatra in the agricultural sector. However, the longer the condition of the harvested land area in the province of North Sumatra is decreasing with the increasing number of residents every year which causes the demand for residential land and infrastructure to increase. This also has an impact on increasing the amount of rice production, based on data from Agricultural Service of North Sumatra, the level of rice production in 2019 for North Sumatra Province reached 4,693,563 tons / year which made North Sumatra province third on the island of Sumatra after South Sumatra province. and Lampung. This is a consideration of how the government's efforts to increase rice production with adequate harvested land area and find out the optimal rice production estimation model for the next harvest season by using panel data regression analysis.

Panel data is a combination of cross section data and time series data (time series). Cross section data consists of several objects. For example, a company name with several characteristics, such as advertising costs, production, stock and others. Time series data is usually data collected at a certain time. For example, stock price and inflation data consisting of several periods, be it daily, weekly, monthly, quarterly and annually (Sofyan et al, 2004).

Panel data regression is a regression method to determine the relationship between the independent variable and the dependent variable by using a panel data structure that combines cross section data and time series data. Cross section data is data collected at one time from the sample, while time series data is a collection of observations within a certain time span where this data is collected in continuous time intervals. In other words, it can be used for forecasting in the future through data obtained in the past.

LITERATURE REVIEW

ANALYSIS OF PANEL DATA REGRESSION

(Refnaldo et al, 2018) panel data regression analysis is a regression method to determine the relationship between the independent variable and the dependent variable by using a panel data structure, which is a combination of cross section data and time series data.Cross section data is data that is collected at one time from the sample, in other words, data that has many objects in the same year (Supranto, 2008)

Time series data is a collection of observations within a certain time span where this data is collected in continuous time intervals. Time series data can be used for forecasting an object in the future through data taken in the past. Examples such as the annual total production of steel, the daily closing price of a stock in the capital market in a period of one month and so on. Mathematically, the time series is defined by the values $Y_1, Y_2, \ldots$ of a Y variable (air temperature, stock closing price and so on) for time points of a Y variable (air temperature, stock closing price and so on) for a point -point in time t_1, $t_2, \ldots$ thus Y is a function of t and is symbolized by $Y=F(t)$. (Spiegel & Stephens, 2004).

The approach in panel data regression analysis is as follows:

2.1. Common Effect Model (CEM)

CEM is the simplest approach technique for estimating panel data regression models. In this approach, all data are combined without regard to individuals and time. In the CEM method, the behavior of the data between individuals (cross section) is the same in various time series, using the Ordinary Least Square (OLS) approach to estimate the parameters.The equation of this method can be written as follows:

$$\sum_{i=1}^{n} Y_{i_t} = \alpha + \sum_{i=1}^{n} \beta_i X_{it} + \varepsilon_{it} \quad (1)$$

where :

Y_{it} : Dependent variable for individual i at time t

α : Intersep (constant)

i : Unit cross section as much as N

t : Unit time series as much as T, $(t = 1,2,3, \ldots, T)$

β_i : Parameters for the i-th variable

X_{it} : Independent variable for individual i at time t

ε_{it} : Error component for individual i at time t

2.2. Fixed Efect Model (FEM)

The FEM method estimates panel data using dummy variables to capture differences in intercepts, this model also assumes that the regression coefficients remain between individuals and over time. The parameter estimation method in the FEM approach is Least Square Dummy Variable (LSDV) where LSDV is a method used in estimating linear regression parameters using OLS on the dummy variable model for different intercepts for each individual and time.The equation of this method can be written as follows:

$$\sum_{i=1}^{n} Y_{i_t} = \alpha + \sum_{i=1}^{n} \beta_i X_{it} + \alpha_i D_{it} + \varepsilon_{it} \quad (2)$$

2.3. Random Effect Model (REM)

The REM method is a regression method that calculates the error value of the regression model using the Generalized Least Square (GLS) method. In contrast to the FEM method, where the effect specification of each individual is treated as part of the error component which is random and does not correlate with the observed independent variables.The equation of this method can be written as follows:

$$\sum_{i=1}^{n} Y_{i_t} = \alpha + \left[\sum_{i=1}^{n} \beta_i X_{it} + \varepsilon_{it}\right] \quad (3)$$

e-ISSN: 2828-9390; p-ISSN: 2828-9382, Hal 61-75

where, $\varepsilon_{it} = \mu_i + e_t + w_{it}$ (4)

with :

ε_{it} : The i-th individual error component at time t

μ_i : Error component cross section

e_t : Error component time series

w_{it} : Combined error component

2.4. Parameter Estimation

Parameters are measurement results that describe the characteristics of the population in other words are values that follow reference information or information that can explain certain boundaries or parts of a system of equations.While estimation or estimation is a process that uses a sample to estimate or estimate the relationship of unknown population parameters. Estimation is a statement about the known population parameters based on the population of the sample, in this case a random sample taken from the population in question. So with this estimation the condition of the population parameters can be known. In general, the parameter is given the symbol β and the estimator is given the symbol $\hat{\beta}$.

2.5. Ordinary Least Square (OLS)

Ordinary Least Square (OLS) is also called the method of ordinary least squares which is a very popular method that produces optimal properties and is easy to calculate. The estimator is found by minimizing the function (Drafer & Smith, 1972)**.** Based on the multiple linear regression equation model which if broken down into will be:

$$
\begin{aligned}
Y &= \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \\
Y_1 &= \alpha + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_n X_{1n} + \varepsilon_1 \\
Y_2 &= \alpha + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_n X_{2n} + \varepsilon_2 \\
&\vdots \quad \vdots \quad \vdots \quad \vdots \quad \cdots \quad \vdots \quad \vdots \\
Y_n &= \alpha + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_n X_{in} + \varepsilon_n
\end{aligned}
$$

By means of the matrix can be written as follows:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11}X_{21} & \cdots & X_{1n} \\ 1 & X_{12}X_{22} & \cdots & X_{2n} \\ \vdots & \vdots \quad \vdots & \cdots & \vdots \\ 1 & X_{n1}X_{n2} & \cdots & X_{in} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

For this purpose it is necessary to select a parameter so that the value of the function is as follows:

$$S = \varepsilon^T \varepsilon \quad ; \quad \varepsilon = Y - X\beta$$

$$S = [\varepsilon_1 \varepsilon_2 \varepsilon_3 \ \ldots \ \varepsilon_N] \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_N \end{bmatrix}$$

$$S = \varepsilon_1\varepsilon_1 + \varepsilon_2\varepsilon_2 + \cdots + \varepsilon_N\varepsilon_N$$

$$S = \sum_{i=1}^{N} \varepsilon_i^2 \qquad (5)$$

Substituting the value of $\varepsilon = Y - X\beta$ into the equation (5)

$$S = (Y - X\beta)^T (Y - X\beta)$$

$$S = (Y^T - X^T\beta^T) - (Y - X\beta)$$

$$S = Y^TY - Y^TX\beta - X^T\beta^TY + X^T\beta^TX\beta$$

$$S = Y^TY - (Y^TX\beta)^T - X^T\beta^TY + X^T\beta^TX\beta$$

$$S = Y^TY - X^T\beta^TY - X^T\beta^TY + X^T\beta^TX\beta$$

$$S = Y^TY - 2X^T\beta^TY + X^T\beta^TX\beta \qquad (6)$$

In order to minimize the function, we do the first differensial S terhadap β

$$\frac{\partial S}{\partial \beta} = \frac{\partial (Y^TY - 2X^T\beta^TY + X^T\beta^TX\beta)}{\partial \beta}$$

$$0 = -2X^TY + 2X^TX\hat{\beta}$$

$$2X^TX\hat{\beta} = 2X^TY$$

$$X^TX\hat{\beta} = X^TY$$

$$\hat{\beta} = (X^TX)^{-1}X^TY \qquad (7)$$

This method produces an unbiased estimate and the minimum variance among all unbiased linear estimates provided that the errors are independent, normal and identical. However, in the presence of multicollinearity, the estimator remains unbiased, but not anymore, because the variance is very large. The large value of variance caused by

e-ISSN: 2828-9390; p-ISSN: 2828-9382, Hal 61-75

multicollinearity will be very dangerous to use regression as a basis for hypothesis testing, estimation and forecasting (Sudiartanto et al, 2017)

2.6. Generalized Least Square (GLS)

Generalized Least Square (GLS) is also called the least common method. This method is used when the assumptions required by the OLS method (homocedasticity and non-autocorrelation) are not met, then GLS is one of the parameter estimation methods used to overcome the presence of autocorrelation if the value of the autocorrelation coefficient is known.

For the Random Effect Model (REM), the parameter estimation is carried out using GLS if the matrix is known, but if it is not known it is carried out with Feasible Generalized Least Square (FGLS) which is estimating the elements of the matrix. In REM incomplete information for each cross section unit is seen as an error so that μ_i is part of the disturbance element. Then the REM equation is:

$$\sum_{i=1}^{n} Y_{i_t} = \alpha + \left[\sum_{i=1}^{n} \beta_{x_{it}} + \varepsilon_{it}\right] \quad (8)$$

With assumption

$$\mu_i \sim N(0, \sigma_\mu^2)$$
$$e_t \sim N(0, \sigma_e^2)$$

For the i-th cross section data, the equation becomes:

$$\sum_{i=1}^{n} Y_{i_t} = \sum_{i=1}^{n} \beta_{x_{it}} + (\mu_i + e_i) \quad (9)$$

If all the residual assumptions required in the Ordinary Least Square (OLS) method are met, then the residual variance-covariance matrix is a diagonal matrix with elements on the main diagonal $\sigma_\mu^2 + \sigma_e^2$.

$$\boldsymbol{\Omega} = \begin{bmatrix} \sigma_\mu^2 + \sigma_e^2 & \sigma_\mu^2 & \cdots & \sigma_\mu^2 \\ \sigma_\mu^2 & \sigma_\mu^2 + \sigma_e^2 & \cdots & \sigma_\mu^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\mu^2 & \sigma_\mu^2 & \cdots & \sigma_\mu^2 + \sigma_e^2 \end{bmatrix}$$

The component variance $\boldsymbol{\Omega}$ is identical for each cross section unit so that the component variance for all observations can be written:

$$\mathbf{w} = \begin{bmatrix} \Omega & \Omega \cdots & \Omega \\ \Omega & \Omega \cdots & \Omega \\ \vdots & \ddots & \vdots \\ \Omega & \Omega \ldots & \Omega \end{bmatrix} \quad (10)$$

If the value of $\mathbf{\Omega}$ is known, the equation can be estimated using Generalized Leas Square (GLS) with:

$$\beta = (X^T W^{-1} X)^{-1} (Y W^{-1} X^T) \quad (11)$$

However, if the value of $\mathbf{\Omega}$ is known, it is necessary to estimate σ_μ^2 and σ_e^2. So the estimating equation becomes:

$$\beta = (X^T \hat{W}^{-1} X)^{-1} (Y \hat{W}^{-1} X^T)$$

Where the residue of the Least Square Dummy Variable (LSDV) is:

$$\hat{\sigma}_e^2 = \frac{\hat{e}^T \hat{e}}{NT - N - i}$$
$$\hat{\sigma}_1^2 = \frac{\hat{\sigma}^2}{T - N - i}$$
$$\hat{\sigma}_\mu^2 = \frac{\hat{\sigma}_1^2 - \hat{\sigma}_e^2}{T} \quad (12)$$

Fertilizer Use Recommendations

Based on the Regulation of the Minister of Agriculture Number40/Permentan/OT.40/4/2007. About Recommendations for Fertilizing N, P and K in site-specific lowland rice. Calculation of the need for N fertilizer in an area based on the level of productivity can be categorized into:

- Low productivity levels (<5 Ton/Ha) require 200 Kg/Ha urea.
- Medium productivity level (5-6 Ton/Ha) requires 250-300 Kg/Ha urea.
- High productivity levels (>6 Tons/Ha) require urea of 300-400 Kg/Ha.

Meanwhile, for the area of rice fields according to the nutrient status of class P and nutrient status of K with a map scale of 1:250,000 based on the Regulation of the Minister of Agriculture, it was obtained 526,719 Kg/Ha. So that the status of P and K nutrients in North Sumatra Province is included in the high class with a recommendation for the use of NPK fertilizer of 325 Kg/Ha.

RESEARCH METHODS

The data used in this study are secondary data obtained from the Department of Agriculture of North Sumatra Province. Agriculture Office of North Sumatra Province about the amount of rice production (appendix 2), land area(appendix 3) and the amount of fertilizer used (appendix 4) based on districts/cities of North Sumatra Province in 2015-2019 (Dinas Pertanian, 2015-2019).North Sumatra Province in 2015-2019 (Department of Agriculture, 2015-2019). The method used aims to obtain an estimate of the value of the estimate that will be used to forecast rice production in North Sumatra Province.The method used aims to obtain an estimate of the value of the estimate that will be used to forecast rice production in North Sumatra Province.used to forecast rice production in North Sumatra Province by selecting the best model approach to panel data regression analysis. The best model approach to panel data regression analysis. Testing will be assisted bySoftware Micrsoft Office Excel, SPSS and Eviews 9 so that data processing will be easier.

RESULT AND DISCUSSION

Based on data processing to do with Software Eviews 9, an estimation model is obtained for each approach model in analysis of panel data regression.

3.1. Commond Effect Model (CEM)

The results of the Common Effect Model (CEM) approach were obtained with Eviews 9 software.

Table 1. Approach result of Common Effect Model (CEM)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.368546	0.131761	10.38656	0.0000
X1	1.105740	0.091576	12.07455	0.0000
X2	-0.096117	0.090929	-1.057059	0.2921

Based on the table, the estimation equation is obtained:

$$\sum_{i=1}^{n} \hat{Y}_{it} = 1{,}368546 + 1{,}1057407X_{it} - 0{,}096117X_{it} + \varepsilon_{it}$$

3.2. Fixed Effect Model (FEM)

The results of the Fixed Effect Model (FEM) approach were obtained with Eviews9 software.

Table 2. Approach result of Fixed Effect Model (FEM)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.311909	0.258553	8.941738	0.0000
X1	0.950890	0.048675	19.53539	0.0000
X2	-0.032532	0.047494	-0.684979	0.4946

Based on the table, the estimation equation is obtained:

$$\sum_{i=1}^{n} \hat{Y}_{it} = 2{,}311909 + 0{,}950690X_{it} - 0{,}032532X_{it} + \varepsilon_{it}$$

3.3. Random Effect Model (REM)

The results of the Random Effect Model (REM) approach were obtained with Eviews9 software.

Table 3. Approach result of Random Effect Model (REM)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.683516	0.141680	11.88252	0.0000
X1	0.995649	0.046665	21.33597	0.0000
X2	-0.008709	0.046126	-0.188814	0.8505

Based on the table, the estimation equation is obtained:

$$\sum_{i=1}^{n} \hat{Y}_{it} = 1{,}683516 + 0{,}995649111X_{it} - 0{,}188814X_{it} + \varepsilon_{it}$$

3.4. Best Model Selection

There are three ways to choose the best approach model in panel data regression analysis, namely:

1. Chow Test

The Chow test was conducted to determine whether the CEM approach was better than the FEM approach.

Table 4. Result of Chow Test

Redundant Fixed Effects Tests
Equation: Untitled
Test cross-section fixed effects

Effects Test	Statistic	d.f.	Prob.
Cross-section F	18.050097	(31,126)	0.0000
Cross-section Chi-square	271.031024	31	0.0000

Based on the table, the probability value of Cross-section F in the Chow test is 0.0000 or $p - value < \alpha$. So in other words H_0 is rejected and H_1 is accepted, which means the Fixed Effect Model (FEM) approach is better than the Common Effect Model (CEM).

2. Hausman Test

Hausman test is used to choose between the FEM approach and the REM approach.

Table 5. Result of Hausman Test

Correlated Random Effects - Hausman Test
Equation: Untitled
Test cross-section random effects

Test Summary	Chi-Sq. Statistic	Chi-Sq. d.f.	Prob.
Cross-section random	11.424876	2	0.0033

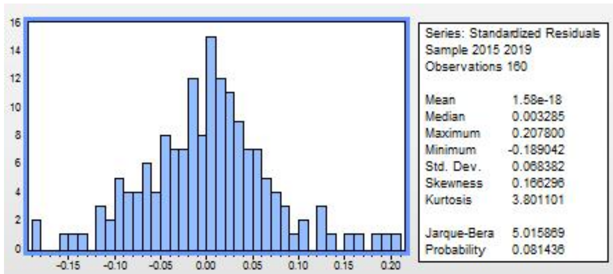
Based on the table, the probability value of Cross-section F in the Chow test is 0.0033 or $p - value < \alpha$. So in other words H_0 is rejected and H_1 is accepted, which means the Fixed Effect Model (FEM) approach is better than the Random Effect Model (REM).Based on the test selection, it was found that the Fixed Effect Model (FEM) approach was the best approach to estimate this research.

3.5. Classical Assumption Test

3.5.1 Normality Test

Normality test is a test used to see whether the data is normally distributed or not.

Figure 1. Result of Normality



The result of the output if the probability value in Jarque-Bera is 0.08146. This means that if the probability value is greater than 0.05 then the data is normally distributed.

3.5.2 Heteroscedasticity Test

Heteroscedasticity test is a test carried out to see whether there is an inequality of variance from the residuals for each observation in the regression model. The purpose of this test is to see any deviations that deviation in the regression model so that the heteroscedasticity test cannot be fulfilled.

Table 6. Result of Heteroscedasticity Test

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.374943	0.122995	3.048454	0.0028
X1	-0.003468	0.023155	-0.149756	0.8812
X2	-0.034454	0.022593	-1.524983	0.1298

Based on the table, the probability value indicator of X_1 and X_2 is greater than α (0.05) so that the data is data that does not meet heteroscedasticity or in other words, the data has similarities from the residuals (homoscedasticity).

3.5.3 Autocorrelation Test

Autocorrelation test is a test carried out to see the correlation between variables, both dependent and independent variables in the regression model with a predetermined time change.

Jurnal Riset Rumpun Matematika dan Ilmu Pengetahuan Alam (JURRIMIPA)
Vol.2, No.1 April 2023
e-ISSN: 2828-9390; p-ISSN: 2828-9382, Hal 61-75

Table 7. Result of Autocorrelation Test

Dependent Variable: Y
Method: Panel Least Squares
Periods included: 5
Cross-sections included: 32
Total panel (balanced)
observations: 160

R-squared	0.998095	Mean dependent var	11.12174
Adjusted R-squared	0.997596	S.D. dependent var	1.566646
S.E. of regression	0.076816	Akaike info criterion	-2.108694
Sum squared resid	0.743492	Schwarz criterion	-1.455219
Log likelihood	202.6955	Hannan-Quinn criter.	-1.843341
F-statistic	2000.282	Durbin-Watson stat	1.655473
Prob(F-statistic)	0.000000		

The Durbin Watson (DW) value obtained is 1.655473. Meanwhile, the total observations (n) are 160 observations and the independent variable (k) is 2. So that the Durbin Lower (DL) value is 1.7163 and the Durbin Upper (DU) is 1.7668 through the Durbin Watson (DW) table value with $\alpha = 5\%$.So based on the provisions of the value of Durbin Watson (DW) < Durbin Lower (DW) which means there is a positive autocorrelation to the research.

3.6. Significant Test

3.6.1 F Test

F test is a test conducted to see whether the independent variable affects the dependent variable.

Table 8. Result of F Test

R-squared	0.998095	Mean dependent var	11.12174
Adjusted R-squared	0.997596	S.D. dependent var	1.566646
S.E. of regression	0.076816	Akaike info criterion	-2.108694
Sum squared resid	0.743492	Schwarz criterion	-1.455219
Log likelihood	202.6955	Hannan-Quinn criter.	-1.843341
F-statistic	2000.282	Durbin-Watson stat	1.655473
Prob(F-statistic)	0.000000		

3.6.2 T Test

T test is a test that is conducted with the aim of knowing the independent variables partially (each one) affect the dependent variable.

Table 9. T Test Result

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.311909	0.258553	8.941738	0.0000
X1	0.950890	0.048675	19.53539	0.0000
X2	-0.032532	0.047494	-0.684979	0.4946

CONSLUSION

Based on the tests carried out in the study, the best approach model for panel data regression is the Fixed Effect Model (REM) approach with Eviews software, the regression estimation is as follows:

1. After fulfilling the diagnostic test and the assumption test of the regression model, the panel data regression model that is more suitable for Rice Production in North Sumatra Province in 2015-2019 is the Fixed Effect Model (FEM) approach with the following estimation equation mode.:

$$\sum_{i=1}^{n} \hat{Y}_{it} = 2{,}311909 + 0{,}950890X_{1it} - 0{,}032532X_{2it}$$

   This means that the variable area of harvested land has an effect on increasing the amount of rice production. While the variable amount of fertilizer use affects the decrease in the amount of rice production. This means that it can be assumed that the amount of fertilizer used is still insufficient to increase the amount of rice production in the harvested area.

2. Based on the Fixed Effect Model (FEM), the coefficient of determination (R^2) is 0,950890 which means that the independent variables X_1 and X_2, affect the dependent variable by 95,09% while the remaining 4,01% is influenced by other variables.Based on the estimated model obtained, it can be calculated the percentage of rice production development in the next 3 years, that is decreased in 2020 at 39,88%, increased in 2021 at 10,82% and increased 2022 at 9,89%.

REFFERENCES

Drafer, N.R. Smith, H. 1972. *Analisis Regresi Terapan Edisi Kedua*. Jakarta : PT Gramedia Pustaka Umum.

Gunawan, Imam. 2017. *Pengantar Statistika Inferensial*. Jakarta : Rajawali Pers.

Meutuah, S. M., Yasin, H., dan Maruddani, S.A.I. 2017. ”Pemodelan*Fixed Effect Geographically Weighted Panel Regression* untukIndeks Pembangunan Manusia di Jawa Tengah”.”*Jurnal Gaussian*”. Vol 6,no 2. Hal 241-250.

Refnaldo, Maiyastri, dan Asdi.2018. “Analisis Ketahanan Pangan Provinsi Sumatera Barat Dengan Metode Regresi Data Panel”. “*JurnalMatematika UNAND”*. vol 7, no. 2. Hal 39-49.

Sofyan, Y. Lien, A.R dan Kurniawan, H.2004.*Regresi dan Korelasi Dalam Genggaman Anda*. Jakarta : Salemba Empat.

Spiegel, M.R. Stephens, L.J. 2004.*Statistika Edisi Keiga*. Jakarta : Erlangga.

Sudiartanto. Suwarno, N. dan Taofik, A.2017. ”Ridge-MM Sebagai Salah Satu Metode Regresi Ridge Yang Robust Terhadap Data Pecilan”. ”*Jurnal FMIPA UNPAD*”. vol 10, no 1. Hal 37-51.

Suharyadi dan Purwanto, S.K.2003. *Statistika Untuk Ekonomi dan Keuangan Modern*. Jakarta : SalembaEmpat.

Supranto, J.2008. *Statistika Teori dan Aplikasi Edisi Ketujuh*. Jakarta : Erlangga.

Youssef, A.H., Abonazel, M.R., dan Ahmed, E.G.2020. “*Estimating The Number Of Patent in The World Using Count Panel Data Models*”.”*Asian Journal of Probability And Statistics”*. Vol 6 no 4. Hal 24-33.