MIMS Majalah Ilmiah Matematika dan Statistika 23 (2), 2023, 172 – 187
eISSN 2722-9886 pISSN 1411-6669
https://jurnal.unej.ac.id/index.php/MIMS/index

Mathematical modeling of farmer household income in Singosari Village, Tanggamus District

Komang Ariyanto

Department of Mathematics, Indonesia Open University
South Tangerang, Banten 15437, Indonesia

corresponding: komangariyanto731@gmail.com

Received: dd-mm-2022, accepted: dd-mm-2022

Abstract

Farmers engage in various sectors (agriculture and non-agriculture) to earn income in order to meet their diverse household needs. Based on various pieces of literature, the agricultural sector serves as a gauge for the income level of rural farming households and provides a greater contribution than the non-agricultural sector. The purpose of this research is to analyze the mathematical modeling of household income among farmer households and to assess the contribution of the agricultural sector to households in Singosari Village, Tanggamus Regency. This study employs a quantitative method with a case study approach. The population in this study consisted of 464 heads of households, from which a sample of 30 households was selected. The respondents in this study were farmer households, and data were collected through surveys using questionnaires, documentation, and literature studies related to the mathematical modeling of income and the contribution of the agricultural sector to people's livelihoods. Data processing was conducted using SPSS software version 26, assisted by Microsoft Excel, and data analysis was performed using a simple linear regression model. The findings of this study indicate that, in Singosari Village, the agricultural income variable significantly influences the household income variable, as determined from the significance value of the simple linear regression results.

Keywords: Agricultural income, household income, mathematical modeling, simple linear regression, singosari farmers

MSC2020: 62P25

1. Introduction

The purpose of this research is to analyze the mathematical modeling of farmer household income. The relationship between the discipline of mathematics and other fields, such as social sciences, economics, and agriculture, is dynamic and complex. However, mathematical modeling, as a mathematical approach, can be applied to construct a model related to farmers' issues. Based on [1], it is said that the methodological aspects in mathematics can be applied to phenomena and social processes in the conditions of contemporary society. Concepts and approaches from mathematical modeling can contribute to the social sciences while taking into account methodological validity and normative adequacy [2]. Thus, it is important to carry out mathematical modeling of the dynamics of farmers' lives (in this case, it is the farmer household income) in Indonesia today.

Mathematical modeling is defined as the process of deriving a mathematical model to simplify a problem using mathematical concepts and symbols. Studies in mathematics, including applied mathematics, mathematical theory, operational research, and mathematical modeling, serve as tools to predict crop productivity and the area of land planned for cultivation [3]. Similarly, as indicated by the findings [4], making land use decisions based on mathematical models holds the potential to significantly enhance sustainable agricultural performance and resource use efficiency in developing countries. Therefore, mathematical models can enhance farmers' efficiency. Consequently, this study aims to analyze the feasibility of modeling real-life agricultural income problems through a set of mathematical equations.

Singosari Village in Tanggamus Regency has significant potential in the agricultural sector. According to Singosari Village Regulation No. 6 of 2021, the village is primarily dryland and known for its productive cultivation of rice, salak, cocoa, coffee, bananas, corn, and pepper [5]. With various potentials in this region, the agricultural sector is the main livelihood of the community in supporting household income.

Several studies examining mathematical modeling have been conducted. A study that tested mathematical modeling related to Local Revenue (PAD) [6] and General Allocation Fund (DAU) [7], poverty analysis using variables such as GRDP, HDI, and Per Capita Expenditure [8], Gross Regional Domestic Product (GRDP) [9], [10], [11]. Furthermore, studies on mathematical modeling have also examined various parametric models, such as nonparametric spline regression [6], nonparametric regression model [10], nonparametric regression model based on a local estimator with a linear kernel [12], genetic algorithm [13], panel regression model [8], multiple linear regression [7], [9], [11], [14], [15], [16], [17], [18], [19], and simple linear regression [11]. Meanwhile, in this study, the focus will be on simple linear regression. Simple linear regression analysis is used because it is simple yet provides a good interpretation in understanding the single influence of one independent variable on the dependent variable, which in this study is the agricultural income variable and the household variable. For farmers, land use decision-making is based on mathematical models in rural areas [4], and even urban land use can be mathematically modeled using hybrid interval-probabilistic programming methods [20].

From several previous studies, research on modeling still focuses on methodological aspects, issues of poverty, and income in a broader spectrum (GRDP and GDP). Mathematical modeling also needs to be carried out in the conditions of rural community livelihoods dominated by the agricultural sector. Therefore, researchers in this study highlight mathematical modeling by considering household income. As in several relevant previous studies conducted on income and expenses in the fields of economics and business [15], [21], [22], [23] and income of transportation service drivers [14]. In the agricultural sector, mathematical modeling is used to measure agricultural household expenditure [24], cattle rancher income and its contribution to overall household income [16]. Studies have also been conducted on the structure and factors influencing the income and expenditure of farming households in rain-fed areas [17], the food expenditure of households utilizing yard land with horticultural cultivation [18], and net income of coffee farmers [19]. However, the difference in this study is that it will focus on mathematical modeling of income in a farmer's household unit, with characteristics of dryland farming in the highlands of Singosari Village, Tanggamus.

Furthermore, based on the description above, we can formulate the main problems that will be studied, which will then become the primary objectives of this research. The main objective of this research is to determine the mathematical model that represents the household agricultural income of farmers in Singosari Village, Tanggamus Regency. In order to answer this research question, researchers will use simple linear regression analysis to model its mathematical form and examine the connections and contributions of agricultural sector income to household income in meeting their daily needs.

Theoretically, the results and discussions in this study can contribute to findings in the field of applied mathematics, especially those related to mathematical modeling on a broader scale. Additionally, it can support the existing literature regarding the contribution of the agricultural sector to the household income of rural farmers. From a practical perspective, it is expected that the results of this study can be used as initial reference material for describing the welfare conditions based on household income indicators for relevant stakeholders.

2. Methods

In this study, researchers attempted to use a quantitative method with a case study approach to explore and provide detailed insights into the development of the situation and context [25], [26]. Quantitative research is a technique or approach to measure a variable using specific instruments and analyze it through statistical procedures [27] given by the respondents and the results of the interpretation of the mathematical model are interpreted. In this study, data were obtained from primary and secondary sources. Data collection from primary sources was carried out using a socio-economic survey of farmer household income and documentation. Meanwhile, secondary data were obtained from a literature study by reviewing previous literature relevant to the research topic. The researcher used a socio-economic survey to collect data on farm household income from agricultural sources. Then, secondary data were obtained using a literature study. The literature study is used to elaborate on previous research and field findings in the form of the interpretation of the output or data that has been processed. The procedure in this literature study was carried out using the keywords: 'mathematical modeling,' 'mathematical model,' 'agricultural income,' and 'household income [28].

This research was conducted in Singosari village with a population in this survey totaling 464 heads of households (based on the Singosari Village profile 2022), and the sample was taken using the formula from Gay and Diehl [29] a total of 30 household heads were selected as samples, with a minimum requirement of 30 respondents in the questionnaire to approximate the distribution closer to the normal curve [29], [30], [31]. The error rate in determining the sample is 17,6%. Typically, research is conducted with an error rate of 5% to 10%, which requires a larger sample size. However, due to the limited number of researchers (only one researcher) and the wide scope of the survey across six hamlets, a higher error rate was chosen to reduce the sample size. Nevertheless, this sample remains quite representative because the researchers conducted a survey by selecting samples that represented the dominant types of farmers in Singosari, distributed across every hamlet.

Respondents were selected through purposive sampling based on specific characteristics. The survey targeted farmer households, referencing household data. The primary criterion for inclusion was that the households had income from the agricultural sector. Therefore, the study's respondents were members of farmer households. A survey questionnaire was used, including questions about agricultural income. The following table presents data on agricultural sector income and total household income from 30 respondents, as shown in Table 1.

To process survey data, Microsoft Excel was used to create tables and graphical data representations. Subsequently, the data from Microsoft Excel was processed using SPSS software version 26 to obtain a mathematical model. Therefore, this study utilizes simple linear regression analysis, which is an analytical method used to identify a linear relationship between the independent variable and the dependent variable, also known as the predictor variable. Based on [32], [33], in this study, one predictor variable is employed. Simple linear regression analysis, a statistical technique used to model the relationship between variables, is used. In this context, X represents agricultural income, Y represents household income, and the equation is as follows:

$$Y = a + bX.$$

Before we proceed with mathematical modeling using simple linear regression analysis, we will conduct a series of classic assumption tests, including the normality test, linearity test, and heteroscedasticity test. The normality test is used to determine whether the data is normally distributed or not. In this study, the normality test was performed using the Kolmogorov-Smirnov test [34]. The Kolmogorov-Smirnov test is one of the statistical tests used to test the normality of data distribution. This test attempts to determine whether a sample of data comes from a normally distributed population. The Kolmogorov-

Smirnov test is conducted using the SPSS statistical software. SPSS will generate various outputs, including the Kolmogorov-Smirnov test statistics table. You need to pay attention to the p-value in this output. If the p-value is greater than the predetermined significance level (α), such as 0,05, then you do not have enough evidence to reject the null hypothesis (H_0), which states that the data is normally distributed. If the p-value is less than α, then you can reject H_0, indicating that the data is not normally distributed. On the contrary, if the p-value is greater than α, then you can reject the alternative hypothesis (H_1), indicating that the data is normally distributed.

Table 1. The mapping of respondents based on their income from the agricultural sector and total household income

No.	Agricultural Sector Income	Total Household Income
1	875.000	2.941.666
2	3.343.333	3.493.333
3	733.333	3.591.666
4	821.250	1.321.250
5	2.464.166	2.664.166
6	664.166	664.166
7	5.750.000	9.150.000
8	2.514.583	2.514.583
9	1.075.000	1.975.000
10	175.000	1.675.000
11	2.960.000	3.660.000
12	3.288.333	6.288.333
13	946.527	2.746.527
14	3.284.583	3.284.583
15	2.068.333	2.568.333
16	2.050.000	2.550.000
17	4.662.500	5.162.500
18	3.041.666	3.041.666
19	2.372.916	3.572.916
20	2.002.500	2.702.500
21	2.279.166	2.279.166
22	875.000	2.600.000
23	7.314.583	13.314.583
24	600.000	6.100.000
25	5.281.666	7.281.666
26	155.833	1.455.833
27	1.937.500	1.937.500
28	108.333	708.333
29	33.333	1.933.333
30	3.286.666	4.644.999

In order to find out whether the data we are going to analyze is linear or non-linear, we use the linearity test [35], [36], linearity test using ANOVA [37]. The linearity test is conducted using the SPSS statistical software. Interpret the results of the ANOVA test generated. Pay attention to the F-value and p-value. A p-value smaller than the predetermined significance level (typically 0,05) indicates a significant difference between the tested groups. The F-value is a statistical measure used in analysis of variance (ANOVA) and regression analysis. It is used to test the overall significance of a group of independent variables (also known as predictors or factors) in explaining the variation in a dependent variable (the outcome or response variable). In the context of regression analysis, the F-value is typically used to assess whether the overall regression model, which includes multiple independent variables, is statistically significant. It does not directly indicate the effect of individual independent variables on the dependent variable, but rather it evaluates whether the combination of all independent variables as a group has a significant impact on the dependent variable. The last test, namely the heteroscedasticity test. Thus, the test at the last stage is fulfilled and we can apply the simple linear regression analysis method to the data from the field survey.

The R-squared value, also known as the coefficient of determination, is a measure used in regression analysis to assess the extent to which the variability in the dependent variable data can be explained by the variability in the independent variables used in the regression model. The R-squared value ranges from 0 to 1. R-squared is used to evaluate how well a linear regression model fits the observed data. The higher the R-squared value, the better the model is at explaining the variation in the data. It provides an indication of how much of the variation in the dependent variable can be accounted for by the model, thus serving as a measure of the quality and fit of the regression model [36].

The T-value is a measure used to test the statistical significance of a coefficient in a regression model. It measures how far the estimated coefficient differs from zero in standard deviation units. The T-value is typically used in hypothesis testing to determine whether an independent variable has a significant effect on the dependent variable in a regression model [38]. The T-value is used in conjunction with degrees of freedom to calculate the p-value. The p-value is a measure of the statistical significance of the coefficient. If the p-value is small (usually less than 0,05), then we can conclude that the independent variable has a significant effect on the dependent variable. Conversely, if the p-value is large, then we do not have enough evidence to conclude that the effect of the independent variable is significant.

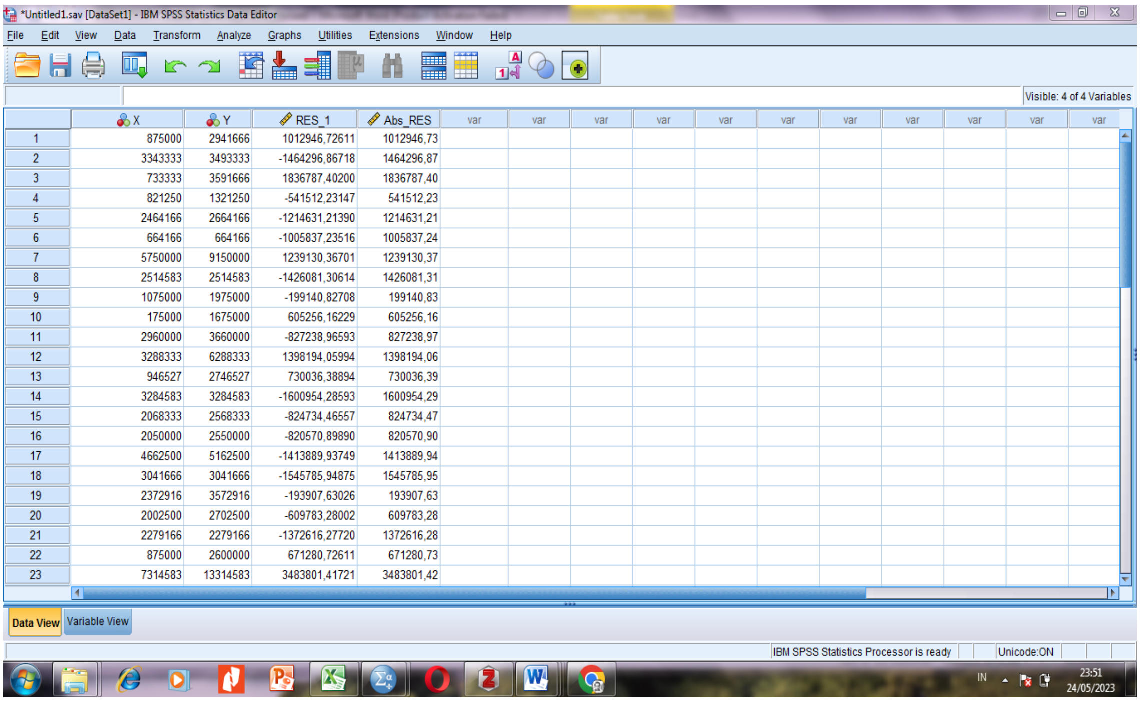
Heteroskedasticity test is a statistical test used to assess whether the variance of errors (residuals) in a regression model significantly changes or not across the range of values of the independent variable. Heteroskedasticity occurs when the residual variance is not constant, indicating a certain pattern in error variability along the regression line. To determine whether the data passes the heteroskedasticity test, we look at the following indicator: if the significance value (sig value) > 0,05, it indicates the presence of heteroskedasticity symptoms; conversely, if sig value < 0,05, it indicates the absence of heteroskedasticity symptoms. The heteroskedasticity test is typically conducted using various statistical methods. One of the common tests used is the White, Breusch-Pagan, or Goldfeld-Quandt test, depending on the data characteristics and assumptions used. In some cases, graphical tests such as a scatterplot of residuals versus predicted values can also be used to detect heteroskedasticity. If the test indicates the presence of heteroskedasticity, the regression model may need to be modified to better fit the observed data [35].

3. Results and Discussions

This section will present survey data and explain the data processing steps, followed by the interpretation of the output, as well as a discussion of the results of other studies that have been conducted. Results and discussion will be presented in the form of tables, figures, and narratives to facilitate data processing and output interpretation. In accordance with the research method, the following aspects will be described: 1) classic assumption testing in simple linear regression; 2) interpretation of the simple linear regression model; and 3) decision-making in a simple linear regression test.

3.1 Three Tests for Testing Classical Assumptions in Simple Linear Regression

Below, data on agricultural sector income and total household income from 30 respondents are presented in Table 2, and the data processing results through SPSS software version 26 are displayed in Figure 1.



	X	Y	RES_1	Abs_RES
1	875000	2941666	1012946,72611	1012946,73
2	3343333	3493333	-1464296,86718	1464296,87
3	733333	3591666	1836787,40200	1836787,40
4	821250	1321250	-541512,23147	541512,23
5	2464166	2664166	-1214631,21390	1214631,21
6	664166	664166	-1005837,23516	1005837,24
7	5750000	9150000	1239130,36701	1239130,37
8	2514583	2514583	-1426081,30614	1426081,31
9	1075000	1975000	-199140,82708	199140,83
10	175000	1675000	605256,16229	605256,16
11	2960000	3660000	-827238,96593	827238,97
12	3288333	6288333	1398194,05994	1398194,06
13	946527	2746527	730036,38894	730036,39
14	3284583	3284583	-1600954,28593	1600954,29
15	2068333	2568333	-824734,46557	824734,47
16	2050000	2550000	-820570,89890	820570,90
17	4662500	5162500	-1413889,93749	1413889,94
18	3041666	3041666	-1545785,94875	1545785,95
19	2372916	3572916	-193907,63026	193907,63
20	2002500	2702500	-609783,28002	609783,28
21	2279166	2279166	-1372616,27720	1372616,28
22	875000	2600000	671280,72611	671280,73
23	7314583	13314583	3483801,41721	3483801,42

Figure 1. The display of data processed through SPSS software version 26

Table 2. Farm income and household income of respondents

No.	Agricultural Sector Income	Total Household Income
1	875.000	2.941.666
2	3.343.333	3.493.333
3	733.333	3.591.666
4	821.250	1.321.250
5	2.464.166	2.664.166
6	664.166	664.166
7	5.750.000	9.150.000
8	2.514.583	2.514.583
9	1.075.000	1.975.000
10	175.000	1.675.000
11	2.960.000	3.660.000
12	3.288.333	6.288.333
13	946.527	2.746.527
14	3.284.583	3.284.583
15	2.068.333	2.568.333
16	2.050.000	2.550.000
17	4.662.500	5.162.500
18	3.041.666	3.041.666
19	2.372.916	3.572.916
20	2.002.500	2.702.500
21	2.279.166	2.279.166
22	875.000	2.600.000
23	7.314.583	13.314.583
24	600.000	6.100.000
25	5.281.666	7.281.666
26	155.833	1.455.833
27	1.937.500	1.937.500
28	108.333	708.333
29	33.333	1.933.333
30	3.286.666	4.644.999

The importance of the classical assumption test in simple linear regression analysis is to determine whether there are any issues and to ensure the validity of the obtained linear regression model [38]. In other words, checking whether the regression equation obtained is accurate in estimation, unbiased, consistent, and valid as an estimator. In the classic assumption test in simple linear regression, there are three types, namely: 1) normality test, 2) linearity test, and 3) heteroscedasticity test.

Table 3. Output of the normality test

One-Sample Kolmogorov-Smirnov Test		
		Unstandardized Residual
N		30
Normal Parameters[a,b]	Mean	0
	Std. Deviation	1.474.470,482
Most Extreme Differences	Absolute	0,152
	Positive	0,152
	Negative	-0,139
Test Statistic		0,152
Asymp. Sig. (2-tailed)		0,074[c]
a. Test distribution is Normal.		
b. Calculated from data.		
c. Lilliefors Significance Correction.		

To find out whether the data is (in) normally distributed or not, we use the normality test. In this study, the normality test was carried out using the One-sample Kolmogorov-Smirnov test [34]. To determine this, we can use significance indicators, namely: data can be considered normally distributed if the significance value is > 0,05, and conversely, if it's < 0,05, then the data is not normally distributed. Based on Table 3, the significance value is 0,074. Since the value is more than 0,05, then we can say that the data is normally distributed. Therefore, the normality test has been satisfied, and we will proceed to the linearity test.

Table 4. Output of the linearity test

ANOVA Table

			Sum of Squares	Df	Mean Square	F	Sig.
Household Income * Farm Income	Between Groups	(Combined)	203.700.080.218.161,000	28	7.275.002.864.934,321	124,641	0,071
		Linearity	140.710.615.182.700,950	1	140.710.615.182.700,950	2410,756	0,013
		Deviation from Linearity	62.989.465.035.460,050	27	2.332.943.149.461,483	39,970	0,125
	Within Groups		58.367.827.778,000	1	58.367.827.778,000		
	Total		203.758.448.045.939,000	29			

In order to find out whether the data we are going to analyze is linear or non-linear, we use the linearity test [35], [36]. Linearity test using ANOVA [37]. To assess linearity, we examine the "sig value of the deviation from linearity". If the "sig value of the deviation from linearity" is > 0,05, it indicates a linear relationship between the response variable (X) and the predictor variable (Y). Conversely, if the "sig value of the deviation from linearity" is < 0,05, it suggests a non-linear relationship between the response variable (X) and the predictor variable (Y). In Table 4, the "sig value of the deviation from linearity" is stating the p-value, whereas the value is 0,125. Since the value is less than 0,005, then we can say that a linear relationship between the response variable (X) and the predictor variable (Y). Since the tests for normality and linearity have been satisfied, we also need to conduct the heteroscedasticity test.

Table 5. Output of the heteroscedasticity test

Coefficients[a]

	Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	T	Sig.
1	(Constant)	850.419,859	268.889,952		3,163	0,004
	Farm Income	0,125	0,094	0,242	1,320	0,197

a. Dependent Variable: Abs_RES

The heteroscedasticity test is essential to assess whether there is variation in residual variance among all observations within the linear regression model. If the heteroscedasticity assumption test cannot be met, it renders the regression model invalid as a forecasting tool. To determine if the data satisfies the heteroscedasticity test, we look at the following indicators: if the significance value (sig value) > 0,05, it indicates the presence of heteroscedasticity symptoms; conversely, if the sig value < 0,05, it suggests the absence of heteroscedasticity symptoms. Table 5 shows that the sig value is 0,197. Since the value is more than 0,05, then we can say that the absence of heteroscedasticity symptoms. Therefore, the final test stage is met, and we can proceed with a simple linear regression analysis of the survey data.

3.2 Mathematical Model Interpretation of Simple Linear Regression

In modeling the mathematical form, simple linear regression analysis tests the relationship between the response variable (X) and the predictor variable (Y). Here, X represents agricultural income, and Y represents household income. To determine this relationship, we use the rule: if the significance value is < 0,05, then the X variable affects the Y variable. Conversely, if the significance value is > 0,05, then the X variable has no effect on the Y variable.

Table 6. The value of the correlation/relationship ®

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0,831[a]	0,691	0,680	1.500.569,331	1,513

a. Predictors: (Constant), Farm Income

b. Dependent Variable: Household Income

Table 6 displays a correlation value (R) of 0,831. Since the value is more than 0,67, then we can say that the indicating a strong relationship. From the SPSS output, the R-square coefficient is 0,691. This demonstrates that the agricultural income variable contributes to explaining 69,1 percent percent of the variance in household income, while the remaining 30,9 percent is influenced by other variables.

Table 7. F value

ANOVA[a]						
Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	140.710.615.182.700,840	1	140.710.615.182.700,840	62,491	0[b]
	Residual	63.047.832.863.238,160	28	2.251.708.316.544,220		
	Total	203.758.448.045.939,000	29			

a. Dependent Variable: Household Income

b. Predictors: (Constant), Farm Income

From the output results in Table 7, the calculated F-value is 62,491 with a significance value of 0, which is less than 0,05. Therefore, based on the F-test, this linear regression model can be applied to predict the agricultural income variable (X) based on the household income variable (Y).

Table 8. Regression coefficient constant value

Coefficients[a]						
Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	T	Sig.
1	(Constant)	854.999,979	441.723,273		1,936	0,063
	Farm Income	1,227	0,155	0,831	7,905	0

a. Dependent Variable: Household Income

From Table 8, we observe that the constant (a) is 854.999,979 and the agricultural income value (variable X) is 1,227. This allows us to formulate a simple linear regression equation:

$$Y = 854.999{,}979 + 1{,}227X.$$

From this equation, we can interpret that the constant value of 854.999,979 represents the baseline for agricultural income (X), meaning that when farm income is 0, the household income (Y) is 854.999,979 Indonesian Rupiah. The regression coefficient of 1,227 indicates that for every 1 Indonesian Rupiah increase in farm income, household income increases by 1,227 Indonesian Rupiah. The positive sign of the regression coefficient (+) signifies a positive relationship between agricultural income (X) and household income(Y).

3.3 Decision Making in Simple Linear Regression Test

We will make decisions in this regression test based on the significance value and the T value. From Table 6, we observe the following: 1) the significance value is 0, which is less than 0,05. This indicates that the agricultural income variable (X) significantly affects the household variable (Y). 2) the calculated T value is 7,905, which is greater than the critical T value of 2,042. This suggests that the agricultural income variable (X) indeed influences the household income variable (Y). This finding aligns with [39], which suggests that the agricultural sector significantly contributes to household income. Therefore, it's crucial to focus on agricultural sector development to mitigate rural population migration. However, [40] argues that while large-scale rural labor migration has short-term benefits, it poses long-term challenges for village development. On the other hand, [41] reports that expanding off-farm income opportunities can help lift smallholder households out of poverty. Regarding the comparison between the agricultural and non-agricultural sectors, [42] finds that families engaged in non-agricultural activities tend to have higher average incomes than farming families. Nevertheless, it's worth noting that income from the non-agricultural sector can also positively and significantly influence agricultural productivity or contribute to agricultural sector income, as suggested by [43], [44], [45], [46].

4. Conclusion

From the presentation in section 3 (results and discussion), several key findings emerge: 1) the mathematical model in simple linear regression analysis is represented as $Y = 854.999{,}979 + 1{,}227X$; 2) he agricultural income variable (X) significantly impacts the household income variable (Y), as indicated by the significance value in the simple linear regression results (0 < 0,05). This implies that agricultural income can influence household income in Singosari Village; and 3) based on the T value, we can conclude that the calculated T value (7,905) exceeds the critical T value (2,042). Consequently, it can be inferred that the agricultural income variable (X) has a significant effect on household variables, thereby affecting household income in Singosari Village.

This research offers significant contributions to the field of applied mathematics, particularly in the realm of mathematical modeling on a broader scale. Furthermore, it lends support to existing literature regarding the impact of agricultural sector income on the overall household income of rural farmers and in broader contexts. It can also serve as a valuable reference or introduction, providing insights into welfare conditions based on household income indicators for relevant stakeholders.

The implications of this research underscore the value of constructing a mathematical model through simple linear regression analysis. This model enhances our understanding of the determinants of rural household income, making it a valuable resource for designing targeted program interventions aimed at poverty reduction and improving overall welfare. This is particularly pertinent given the well-documented challenges of high poverty rates in many developing rural areas.

However, it's essential to acknowledge the limitations of this research, primarily its focus on mathematical modeling of agricultural sector income and farmer household income. The sample size was constrained by limited survey staff. Therefore, future research should consider expanding its scope to include non-agricultural income and expenditure patterns for a more comprehensive analysis. Methodologically, efforts to reduce error percentages in surveys and the inclusion of additional researchers in future surveys would contribute to a more robust study with a larger sample size.

References

[1] A.Y. Bulatetskaya, "Sociological and mathematical models as tools of social processes applied sociological research," *International Journal of Recent Technology and Engineering (IJRTE)*, pp. 5462-5473, 2019. [CrossRef]

[2] A. Saltelli and A. Puy, "What can mathematical modelling contribute to a sociology of quantification?," *Humanit. Soc. Sci. Commun.*, vol. 10, no. 1, Art. no. 1, 2023. [CrossRef]

[3] M.A. Vodounon, "Mathematicals models can make farmers more efficient." Diakses: 26 September 2023. [Online]

[4] M.T. Mellaku and A.S. Sebsibe, "Potential of mathematical model-based decision making to promote sustainable performance of agriculture in developing countries: a review article," *Heliyon*, vol. 8, no. 2, p. e08968, 2022. [CrossRef]

[5] K. Ariyanto, "Participatory socio-economic mapping of singosari village, Talang Padang District, Tanggamus Regency," *Int. J. Soc. Sci. Econ. Art*, vol. 13, no. 1, pp. 34-49, 2023. [CrossRef]

[6] V.E.D. Pribadi, "Pemodelan pendapatan asli daerah Kabupaten Banyuwangi tahun 2014 menggunakan regresi nonparametrik spline," Skripsi, Institut Teknologi Sepuluh Nopember Surabaya, 2016.

[7] A. Asraf, M.S. Lubis, Z.S. Perapatih, and W. Sari, "The effect of regional original income (Pad) and general allocation funds (DAU) on capital goods expenditure in West Pasaman," *E-J. Apresiasi Ekon.*, vol. 7, no. 1, pp. 30-37, 2019. [CrossRef]

[8] R. Aprilianti, G.C. Messakh, S.N. Asiah, and D.A. Nohe, "Analisis regresi data panel pada kasus persentase kemiskinan di Kalimantan Timur," *Pros. Semin. Nas. Mat. Dan Stat.*, vol. 2, 2022. [CrossRef].

[9] R.J. Tamba, “Menentukan model matematika pendapatan domestik regional bruto (PDRB) di Kabupaten Deli Serdang,” Skripsi, UNIMED, 2013.

[10] A. Mansyur and E. Simamora, “Mathematical modeling of gross regional domestic product growth rate of North Sumatra Province by business field using local polynomial regression,” *AIP Conf. Proc.*, vol. 2659, no. 1, p. 110020, 2022. [CrossRef]

[11] B.T. Bulayo, “Mathematical Modelling of the Gross Domestic Product of the Philippines,” *Int. J. Soc. Sci. Educ. Res. Stud.*, vol. 03, no. 05, pp. 750-754, 2023. [GreenVersion]

[12] M.N. Tilova, “Pemodelan pendapatan asli daerah menggunakan regresi nonparametrik lokal linier kernel,” Skripsi, Muhammadiyah University, Semarang, 2019.

[13] M.S. Sinaga, Y.M. Rangkuti, and S. Manullang, “Using genetic algorithms to optimize regional original income in the tourism sector,” *AIP Conf. Proc.*, vol. 2659, no. 1, p. 110012, 2022. [CrossRef]

[14] A.D. Rahayu, “Model matematika pendapatan pengemudi angkutan kota (angkot) jalur ADL dan AL di Malang,” Skripsi, Universitas Brawijaya, 2008.

[15] D. Djeni, T. Zamania, and I.S.N. Hasanah, “Pemodelan matematika terhadap keuntungan harian pada penjualan jajanan dikafe nihayah Pondok Putri Pesantren Zainul Hasan,” *BAHTSUNA J. Penelit. Pendidik. Islam*, vol. 3, no. 1, pp. 58-70, 2021. [CrossRef]

[16] E. Prasetyo, T. Ekowati, and S. Gayatri, “An income analysis of beef cattle fattening system and its contribution to the total household income in Central Java Province,” *J. Indones. Trop. Anim. Agric.*, vol. 45, no. 4, pp. 365-372, 2020. [CrossRef]

[17] B. Hartoyo, Komalawati, and D. Sahara, “Analysis of income and expenditure of farmers’ household in the rain-fed area of Boyolali district,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 653, no. 1, p. 012007, 2021. [CrossRef]

[18] M. Martina, R. Praza, and Z. Zuriani, “Factors affecting household food expenditure through the utilization of yard land in Dewantara Subdistrict, Aceh Utara,” *Int. J. Econ. Bus. Account. Agric. Manag. Sharia Adm. IJEBAS*, vol. 1, no. 1, pp. 62-70, 2021. [CrossRef]

[19] I. Irmeilyana, N. Ngudiantoro, and S.I. Maiyanti, “Regression model on pagaralam coffee farmers’ income with the influence of the use of herbicide reductant variable,” *BAREKENG J. Ilmu Mat. Dan Terap.*, vol. 16, no. 2, pp. 409-420, 2022. [CrossRef]

[20] S. Tan *dkk.*, “A hybrid mathematical model for urban land-use planning in association with environmental–ecological consideration under uncertainty,” *Environ. Plan. B Urban Anal. City Sci.*, vol. 44, no. 1, pp. 54-79, 2017. [CrossRef]

[21] H. Chintia, "Pemodelan matematika terhadap pendapatan usaha di koperasi mahasiswa UIN RIL," Skripsi, UIN Raden Intan Lampung, 2021.

[22] M.D. br Barus, M.N. Sari, and F.S. Thahirah, "Pemodelan matematika terhadap pendapatan usaha di koperasi unit desa (KUD) Tomuan Holbung Kabupaten Asahan," *EKSAKTA J. Penelit. Dan Pembelajaran MIPA*, vol. 7, no. 1, pp. 162-168, 2022. [CrossRef]

[23] N. Fadilah, "Pemodelan matematika terhadap keuntungan harian penjualan produk di toko Kholidi," *JUMANT*, vol. 12, no. 2, pp. 45-59, 2020. [CrossRef]

[24] M. Fajar, "Model kurva Lorenz pada pengeluaran rumah tangga pertanian di Provinsi Papua," *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 1, no. 3, pp. 153-158, 2020. [CrossRef]

[25] N.K. Denzin (ed) and Y.S. Lincoln, *The SAGE Handbook of Qualitative Research*. United State of America: Sage Publications, 2005. [Online]

[26] S. Suryabrata, *Metodologi penelitian*. Jakarta: Rajawali Pers, 2016. [Online]

[27] J.W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. United State of America: Sage Publications, 2014.

[28] K. Ariyanto, "Literature Review: Urban Poverty in a Sociological Perspective," *Antroposen J. Soc. Stud. Hum.*, vol. 2, no. 1, pp. 24-32, 2023. [CrossRef]

[29] L.R. Gay and P.L. Diehl, *Research Methods for Business and Management*. New York: Macmillan Publishing Company, 1992.

[30] M. Singarimbun and S. Effendi, *Metode Penelitian survei*, Rev.ed. Jakarta: PT. Pustaka LP3ES, 1989. [Online]

[31] L. Cohen, L. Manion, and K. Morrison, *Research Methods in Education*, 8 ed. London: Routledge, 2017. [CrossRef]

[32] D. Maulud and A.M. Abdulazeez, "A review on linear regression comprehensive in machine learning," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 4, pp. 140-147, 2020. [CrossRef]

[33] S. Etemadi and M. Khashei, "Etemadi multiple linear regression," *Measurement*, vol. 186, p. 110080, 2021. [CrossRef]

[34] X.J. Bilon, "Normality and significance testing in simple linear regression model for large sample sizes: a simulation study," *Commun. Stat. - Simul. Comput.*, vol. 52, no. 6, pp. 2781-2797, 2023. [CrossRef]

[35] U. Khasanah, *Analisis Regresi*. Yogyakarta: UAD Press, 2021.

[36] Sudjana, *Teknik Analisis regresi dan Korelasi*. Bandung: Tarsito, 2003. [Online]

[37] J. (JD) Long and P. Teetor, *15 Simple Programming | R Cookbook, 2nd Edition*. California, USA: O'Reilly Media, 2019. Diakses: 26 September 2023. [Online]

[38] T. Dahiru, "P–value, a true test of statistical significance? A cautionary note," *Ann. Ib. Postgrad. Med.*, vol. 6, no. 1, pp. 21-26, 2008.

[39] R. Mukaila, A. Falola, and L. Egwue, "The Determinants of Rural Households' Income in Nigeria," *Fiscaoeconomia*, vol. 5, no. 3, pp. 978-989, 2021. [CrossRef]

[40] Y. Li, P. Fan, and Y. Liu, "What makes better village development in traditional agricultural areas of China? Evidence from long-term observation of typical villages," *Habitat Int.*, vol. 83, pp. 111-124, 2019. [CrossRef]

[41] S.-A. Emran, T.J. Krupnik, S. Aravindakshan, V. Kumar, and C.M. Pittelkow, "Factors contributing to farm-level productivity and household income generation in coastal Bangladesh's rice-based farming systems," *PLOS ONE*, vol. 16, no. 9, p. e0256694, 2021. [CrossRef]

[42] D.F.S. Mesquita, C.A. do Nascimento, and P.V.P.S. Lima, "Non-agricultural income, infrastructure and access to consumer goods in rural households in the northeast and south," *Rev. Econ. E Sociol. Rural*, vol. 61, p. e263246, 2023. [CrossRef]

[43] Md. S. Rashidin, S. Javed, B. Liu, and W. Jian, "Ramifications of households' nonfarm income on agricultural productivity: evidence from a rural area of Pakistan," *SAGE Open*, vol. 10, no. 1, p. 2158244020902091, 2020. [CrossRef]

[44] A. Tandjigora, "Impact of non-agricultural activities on farmers' income: evidence from the senegalese groundnut area," dalam *The Palgrave Handbook of Agricultural and Rural Development in Africa*, E. S. Osabuohien, Ed., Cham: Springer International Publishing, 2020, pp. 395-423. [CrossRef]

[45] A. Rahman and S. Mishra, "Does non-farm income affect food security? Evidence from India," *J. Dev. Stud.*, vol. 56, no. 6, pp. 1190-1209, 2020. [CrossRef]

[46] Z. Chen, A. Sarkar, M.S. Hossain, X. Li, and X. Xia, "Household labour migration and farmers' access to productive agricultural services: a case study from Chinese Provinces," *Agriculture*, vol. 11, no. 10, p. 976, 2021. [CrossRef]