FARABI
Jurnal Matematika dan Pendidikan Matematika
p-ISSN 2623-2332
e-ISSN 2798-5474

Comparison of Fuzzy Logic and Multiple Linear Regression in Forecasting Rice Production in Toba District

Agustaeys Pasaribu[1], M.R Syahputra[2]
[1]Student of Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Sumatera Utara Medan-Indonesia 20155
[3]Mathematics Undergraduate Study Program, Faculty of Mathematics and Natural Sciences, Universitas Sumatera Utara Medan-Indonesia 20155
Email: [1]agustaeysp1000@gmail.com

ABSTRAK

Tujuan dari penelitian ini adalah untuk meramalkan produksi padi di Kabupaten toba dimasa yang akan datang. Faktor yang mempengaruhi produksi padi yang diteliti adalah luas lahan, pupuk subsidi, populasi hama padi, rata-rata curah hujan dah hari hujan. Metode yang digunakan dalam penelitian ini adalah metode logika fuzzy dan regresi linier. Penelitian ini membandingkan hasil peramalan yang lebih akurat berdasarkan error dari kedua metode. Data yang digunakan merupakan data sekunder dari Dinas Pertanian Kabupaten Toba. Data yang diteliti berupa data produksi padi, luas lahan, pupuk subsidi, populasi hama, rata-rata curah hujan dan hari hujan tahun 2016-2021. Penelitian ini dilakukan dengan bantuan software matlab toolbox fuzzy dan SPSS. Dari hasil peramalan yang diperoleh, metode regresi linier berganda lebih mendekati hasil yang sebenarnya sebesar 878.428,07 Ton dibandingkan menggunakan logika fuzzy sebesar 1.032.300 Ton. Berdasarkan standar error dari kedua metode, dapat disimpulkan bahwa metode regresi linier berganda memiliki hasil yang lebih akurat dengan error 3.11% dibandingkan metode logika fuzzy dengan error sebesar 22,05%. Nilai koefisien determinasi menunjukkan pengaruh variabel luas lahan, pupuk subsidi, populasi hama, rata-rata curah hujan dan hari hujan sangat mempengaruhi produksi padi sebesar 99,9%.

Kata Kunci: Padi, Produksi, Logika Fuzzy, Regresi Linier, Standar Error

ABSTRACT

The purpose of this study is to forecast rice production in Toba Regency in the future. Factors affecting rice production studied are land area, subsidized fertilizers, rice pest populations, average rainfall on rainy days. The methods used in this study are fuzzy logic methods and linear regression. This study compared more accurate forecasting results based on errors from both methods. The data used is secondary data from the Toba Regency Agriculture Office. The data studied were in the form of data on rice production, land area, subsidized fertilizers, pest populations, average rainfall and rainy days in 2016-2021. This research was conducted with the help of matlab toolbox fuzzy and SPSS software. From the forecasting results obtained, the multiple linear regression method is closer to the actual result of 878,428.07 Tons compared to using fuzzy logic of 1,032,300 Tons. Based on the error standards of the two methods, it can be concluded that the multiple linear regression method has a more accurate result with an error of 3.11% compared to the fuzzy logic method with an error of 22.05%.

Keywords: Rice, Production, Fuzzy Logic, Linear Regression, Standard Error.

A. Pendahuluan

Forecasting is the activity of predicting the future value of the value obtained in the previous period. Forecasting is the ability to minimize errors or errors when predicting a problem optimally. The inability to accurately predict future events can be a big risk, for which strong forecasting is needed. One of them is the selection of the right forecasting method with high accuracy and minimal error (Arfina & Tomi, 2018). A commonly used method in forecasting is linear regression. The linear regression method is used to form the equations of several free variables that are considered to relate to bound variables. In addition to multiple linear regression the method used to foresee a future event is fuzzy logic. Fuzzy logic has the advantage of foreseeing future production with reasonings in the data of previous periods (Yulia R.S., dkk, 2021).

One of the rice producing areas is Toba Regency. Toba Regency is one of the regions in North Sumatra where the majority of people's income is from agricultural products, one of which is as a rice farmer. The Toba Regency area is suitable as a rice field because the area is in the mountains and there are many small rivers. In 2016, Toba Regency had a rice field area of 22,895 ha with a production of 146,701 Tons. Meanwhile, in 2021, Toba Regency is only able to produce rice production of 147,001.5 Tons with an area of 22,450.6 ha (tobakab.go.id, 2022).

The area of rice sowing land is one of the main factors of rice production, the wider the rice planting land, the greater the probability of the yield produced. In addition to land area, the triggering factor that causes rice production is the use of fertilizers. In this study, the fertilizer in question is the availability of government subsidized fertilizers. Many of the farmers do not get subsidized fertilizer from the government. Meanwhile, the price of non-subsidized fertilizers continues to increase in prices so that farmers are overwhelmed to fertilize rice crops. This makes the harvest obtained unsatisfactory (Yosefina, Made & Dewa, 2018). In addition to fertilizers, the disruptive factor of rice plants is plant pests. Pests of rice plants often become destroyers or make rice farmers fail to harvest, such as rats, sparrows, caterpillars, grasshoper and so on (Tri, Made & Lien, 2019). Another factor to be studied in this study is the average rainfall and the number of rainy days. Rainfall and rainy days that are too high or too low can cause less than perfect rice growth (Isnain, dkk 2021).

This study predicted rice production in Toba Regency using multiple linear regression methods and fuzzy logic. Factors affecting rice production studied were land area (X_1), the use of subsidized fertilizers (X_2), rice pests (X_3), average rainfall (X_4) and rainy days (X_5). The fuzzy logic method used is the Mamdani method, and the linear regression method used is the least squared The composition of the rules that is often used is the Max method. The solution of the Max method is the maximum value of the fuzzy inference of the membership function. Fuzzy inference is the merging of several rules from a fuzzy set. The Max method can be formulated in the following equation 2. The composition of the rules that is often used is the Max method. The solution of the Max method is the maximum value of the fuzzy inference of the membership function. Fuzzy inference is the merging of several rules from a fuzzy set. The Max method can be formulated in the following equation 2. multiple regression. From the results obtained, it can be known which method is better to predict the outcome.

B. Metode Penelitian

1. Fuzzy Mamdani

The fuzzy mamdani method is a forecasting method used to estimate the future based on previous data. The Mamdani method is often also referred to as the Max-Min method. In 1975 Ebrahim Mamdani began to introduce the mamdani method (Aulia Ishak, 2010). To get a good result or output, there are 4 basic stages that must be considered:

1. Formation of fuzzy sets
   The formation of a fuzzy set begins with determining the input and output variables of the set, determining the universe of speech and dividing the domain of each variable.
2. Application of the fuzzy implication function of mamdani
   An implication function that is often used in research is the Min function. The general form of the implication function is as follows:
   IF x is A THEN y is B

3. Create a combination or rule composition
   The composition of the rules that is often used is the Max method. The solution of the Max method is the maximum value of the fuzzy inference of the membership function. Fuzzy inference is the merging of several rules from a fuzzy set. The Max method can be formulated in the following equation 2.
   $\mu_{sf}(x_i)$ = (max $\mu_{sf}(x_i)$, $\mu_{kf}(x_i)$)
4. Affirmation function (defuzzyfication)
   The affirmation function is the output of an assertive number in the fuzzy set domain. The affirmation function is obtained from the regional solution of the composition of the rules. The affirmation function used is the centroid method. The general form of the centroid method is as follows:

$$y^* = \frac{\int_a^b y\,\mu(y)\,dy}{\int_a^b \mu(y)\,dy}$$

2. Regresi Linier Berganda

Multiple linear regression is a test method used to find out how the influence or relationship of free variables affects the production of bound variables. The results of forecasting using the method of multiple linear regression are obtained by forming an equation that connects free variables and bound variables (Yulia, dkk., 2019). The equation of multiple linear regression can be formulated in the following equation.

$$\hat{Y}_i = b_0 + b_1X_1 + b_2X_2 + \cdots + b_nX_n + \varepsilon_i$$

The stages to solve the multiple regression equation and the testing of the function of the multiple linear regression coefficient consisting of 5 independent variables and 1 bound variable include:

1. It starts by searching and determining the price of each variable to get the price of b_0, b_1, b_2, b_3, b_4 and b_5.
2. Create a multiple linear regression equation of 5 independent variables as in the general formula above:

$$\hat{Y}_i = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$$

3. Standard error

Standard error is a relative error in forecasting calculations. Here is a standard description of the error of each method:

1. Standar error fuzzy mamdani
   $$S_e = \sqrt{\frac{s^2}{n}}$$
2. Standar error regresi linier
   $$S_e = \sqrt{\frac{\sum(Y-\hat{Y}_i)^2}{n-k-1}}$$
3. Then the standard percentage of error can be calculated with the formula:
   Percentage = $\frac{S_e}{\bar{X}}$ x 100%

C. Hasil dan Pembahasan

The data used for this study were data on rice production, land area, availability of subsidized fertilizers, pest population, average rainfall and rainy days within 6 years (tobasamosirkab.go.id, 2022).

Table 1 Production Data, Land Area, Subsidized Fertilizers and Rice Pests 2016-2021.

Year	Rice Product (Tons)	Land Area (Ha)	Fertilizer Subsidy (Tons)	Pests (Ha)	Rain fall (mm)	Rainy Day
January-April 2016	14444	2808	3438	116.5	79	48
May-August 2016	123536	19448	2853.1	109.5	90,5	52
September-December2016	14572	2105	3854.9	102.5	138	61
January-April 2017	23354	3854	2920	76	172.75	73
May-August 2017	113548	17719	1697	73.5	99.5	52
September-December2017	24653	3780	3869	77	178.75	84
January-April 2018	14766.74	2687.70	3369	53	134.75	70
May-August 2018	105523.95	16574.70	1938	47	70.5	39
September-December2018	31292.84	4919.30	2884	46	218	75
January-April 2019	22722.95	3828.20	3759	60.5	140	52
May-August 2019	99449.11	15886.20	2128	60.5	195.75	62
September-December2019	21491.84	3725.10	2716	56.5	120.5	54
January-April 2020	19823.36	3400.60	2988	43.5	83.75	55
May-August 2020	97631.18	15559.10	2518.5	39	121.75	62
September-December2020	15973.77	2810.40	4138.5	38.5	171	71
January-April 2021	21016.56	3508.50	3437	107.5	119	42
May-Agustus 2021	92019.54	15043.10	2220	115.5	156.5	53
September-December2021	23732.95	3899.00	3389	110.5	218	68
Total	**879,552**	**141,555**	**54,117**	**1,333**	**2,508**	**1,073**

1. Fuzzy Set Formation

Based on Table 1, the sum of the data consists of 18 periods for which the universe value will be determined between the lowest value and the highest value. It can then be divided the domains of each fuzzy set.

Tabel 2 Fuzzy Set Formation

Function	Variabel	*Fuzzy set*	Universe of Talks	Domain	Information
Input	Land Area	Smallest Medium Most	[2,105-19,448]	[2,105-10,776.50] [6,440.75-15,112.25] [10,776.50-19,448]	Land Area (Ha)
	Land Area	Smallest Medium Most	[1,697-4,138.5]	[1,697-2,917.75] [2,307.375-3,528.125] [2,917.75-4,138.5]	Land Area (Tons)
	Pest Population	Lowest Medium Highest	[38.5-116.5]	[38.5-77.5] [58-97] [77.5-116.5]	Pest Population (Ha)
	Rainfall	Lowest Medium Highest	[70.5-218]	[70.5-144.25] [107.375-181.125] [144.25-218]	Rainfall (mm)
	Rainy Days	Smallest Medium Most	[39-84]	[39-61.5] [50.25-72.75] [61.5-84]	Rainy Days
Output	Rice Production	Decreased Constant Increased	[14,444-123,536]	[14,444-68,990] [41,717-96,263] [68,990-123,536]	Rice Production (Tons)

2. Function Implications

The implication function used is the min (minimum) function. In this study, 74 fuzzy rules were used based on the data in table 1. As an example, you can see the fuzzy rules in the data for the January-April 2016 period as follows:

[R1] IF Land Area IS SMALLES AND Fertilizer Subsidied IS MEDIUM AND Pest Population IS HIGHEST AND Rainfall IS LOWEST AND Rainy Days IS SMALLEST THEN Production IS DECREASED

[R2] IF Land Area IS SMALLEST AND Fertilizer Subsidied IS MOST AND Pest Population IS HIGHEST AND Rainfall IS LOWEST AND Rainy Days IS SMALLEST THEN Production IS DECREASED

In table 1, the land area for the January-April 2016 period is 2,808 Ha, subsidized fertilizers are 3,438 tons, pest populations are 116.5 Ha, average rainfall is 79 mm and rainy days are 48 days, then the membership function of each variable can be calculated, including:

1. For a land area of 2,808 Ha, it is calculated:

   μx_1Small(2,808) = $\frac{10,776.50-2,808}{8671.5}$ = 0.92

   μx_1Medium(2,808) = 0

   μx_1Most(2,808) = 0

2. For subsidized fertilizers as many as 3,438 then:

   μx_2Small(3,438) = 0

   μx_2Medium(3,438) = $\frac{3,528.125-x}{610.375}$

   = $\frac{3,528.125-3,438}{610.375}$

   = 0.15

   μx_2Most(3,438) = $\frac{(x-2,917.75)}{1,220.75}$

   = $\frac{(3,438-2,917.75)}{1,220.75}$ = 0.43

3. For a pest population of 116.5 Ha then:

   μx_3Low (116,5) = 0

   μx_3Medium(116,5) = 0

   μx_3High(116,5) = 1

4. For an average rainfall of 79 mm then:

   μx_4Low(79) = $\frac{144.25-79}{73.75}$

   = $\frac{65.25}{73.75}$ = 0.88

   μx_4Medium(79) = 0

   μx_4High(79) = 0

5. For a 48-day rainy day then:

   μx_5Small(48) = $\frac{61.5-48}{22.5}$ = $\frac{13.5}{22.5}$ = 0.6

   μx_5Medium(48) = 0

   μx_5Most(48) = 0

α-predikat1 = μx_1Small ∩ μx_2Medium ∩ μx_3High ∩ μx_4Low ∩ μx_5Small

= *min* (μx_1Small(2,808), μx_2Most(3,438), μx_3Hihg(116.5), μx_4Low(79), μx_5Small(48))

= *min* (0.92;0.15;1;0.88;0.6)

= 0.15

α-predikat2 = μx_1Small ∩ μx_2Most ∩ μx_3High ∩ μx_4Low ∩ μx_5Small

= *min* (μx_1Small (2,808), μx_2Most (3,438), μx_3High(116.5), μx_4Low(79), μx_5Small(48))

= *min* (0.92;0.43;1;0.88;0.6)

= 0.43

3. Rule Composition

The composition of the rules used is a non-zero min function. From the α-predicate values obtained in the implication function above, α-predicate2 is greater than α-predicate1, then the composition of the rules can be calculated as follows:

α-predikat 2 = *min* (0.92;0.43;1;0.88;0.6) = 0.43

μyReduced = $\frac{68,990-y}{54,546}$ = 0.43

y = 68,990- 54,546 x 0.43 = 45,535.22

So that the fuzzy area solutions for the period January-April 2016 include:

µyReduced =

$$\begin{cases} 1 & ; \quad y \le 14,444 \\ \frac{(68,990-y)}{54,546} & ; 14,444 < y < 68,990 \\ 0 & ; \quad y \ge 68,990 \end{cases}$$

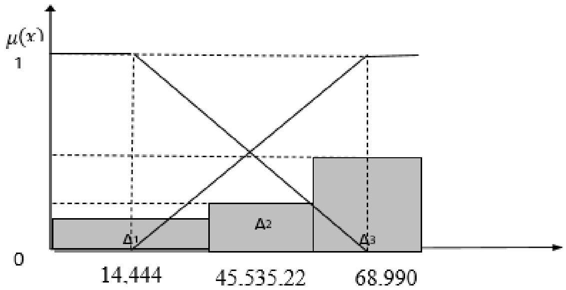

Figure 2 Fuzzy area solution for the period January-April 2016

The affirmation function used in fuzzy mamdani is the centroid method. To calculate the affirmation function in the period January-April 2016, it is calculated the moment (M) and area of the area (A) to calculate the value of the center point (y*) from figure 2. The first step is to calculate the M_1 moment and M_2:

1. The first inference is a linear function, then it can be calculated:

$$M_1 = \int_{14,444}^{45,535.22} 0.43\ y\ dy$$

$$M_1 = \frac{0.43\ y^2}{2} \Big|_{14,444}^{45,535.22}$$

$$M_1 = 0.215(45,535.22)^2 - 0.215(14,444)^2$$

$$M_1 = 400,937,832$$

2. The second inference is the descending function, then it can be calculated:

$$M_2 = \int_{45,535.22}^{68,990} \frac{68,990-y}{54,546}\ y\ dy$$

$$M_2 = \frac{68,990\ y^2}{2 \text{ X } 54,546} - \frac{y^3}{3\text{X}54,546} \Big|_{45,535.22}^{68,990}$$

$$M_2 = \left(\frac{68,890\ (68,990)^2}{109,092} - \frac{(68,990)^3}{163,638}\right) - \left(\frac{68,890\ (45,535.22)^2}{109,092} - \frac{(45,535.22)^3}{163,638}\right)$$

M_2 = (3,009,993.315 – 2,006,662,210) – (1,311,257,905 – 576,976,539.6)

M_2 = 269,049,739

Next will be calculated the area of the fuzzy solution area (A_1 and A_2) from figure 2:

A_1 = (0.43) x (45,535.22 – 14,444)

= 13,369.22

$A_2 = (0.43) \text{ x } \left(\frac{68,990-45,535.22}{2}\right)$

= 11,727.39

Furthermore, the central point is calculated by summing the M_1+M_2 divided by A_1+A_2, then obtained:

$$y^* = \frac{400,937,832 + 269,049,739}{13,369.22 + 11,727.39}$$

= 36,388.63172

The value of the central point (y*) obtained was 36,388.63172, meaning that the optimum rice production in the January-April 2016 period was 36,388.63172 Tons. Figure 3 is the result of testing rice production in the period January-April 2016 using the matlab toolbox (Sri Kusumadewi, 2002).

Figure 3 Production output for the period January-April 2016 by centroid method

4. Multiple Linear Regression Equation

The normal equation of 5 free variables is $\hat{Y}_i = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$. With the help of the SPSS program, the value of the regression coefficient can be calculated from the table data in table 3, including:

Table 3 Regression coefficient of 5 independent variables

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta
1 (Constant)	-4800.976	3711.185	
Land Area	6.449	.097	1.008
Subsidy Fertilizer	.372	.940	.006
Rice Pests	-1.045	14.465	-.001

Rainfall	-4.830	12.718	-.005
Rainy Day	43.243	53.792	.012

Furthermore, the value of the coefficient obtained is substitution into equation 11 so as to form a linear regression equation, including:

$$\hat{Y}_i = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$$

$$\hat{Y}_i = -4{,}800.976 + 6.449X_1 + 0.372X_2 - 1.045X_3 - 4.83X_4 + 42.243X_5$$

Than can be calculated multiple regression forecasting values $\hat{Y}_i$ by substituting the values of variables in table 4.

Table 4 Production results of fuzzy mamdani and regression

Y	$\hat{Y}_i$*fuzzy*	$\hat{Y}_i$*regresi*
14,444	36,100	16,111.10
123,536	102,000	123,325.62
14,572	36,300	12,011.36
23,354	36,600	23,309.65
113,548	97,900	111,739.38
24,653	35,700	23,620.10
14,766.74	36,800	16,036.05
105,523.95	103,000	104,068.05
31,292.84	35,700	30,063.65
22,722.95	36,200	22,742,65
99,449.11	102,000	100,051.11
21,491.84	37,300	21,872.61
19,823.36	37,700	20,114.42
97,631.18	100,000	98,466.80
15,973.77	36,900	16,995.91
21,016.56	37,100	20,191.00
92,019.54	87,300	94,400.10
23,732.95	37,700	23,308.49
879,552	**1,032,300**	**878,428.07**

5. Standar Error

From the results of the production obtained, the error standards of the two methods can be calculated by the following equation:

Table 5 Standard Error

Standard Error Fuzzy Mamdani	Standard Error Linear Regresion
$S_e = \sqrt{\frac{s^2}{n}}$	$S_e = \sqrt{\frac{\sum(Y-\hat{Y}_i)^2}{n-k-1}}$
$= \sqrt{\frac{(41{,}729.89)^2}{18}}$	$S_e = \sqrt{\frac{27{,}826{,}328.97}{18-5-1}}$
= 10,774.6112	= 1,522.78
Percentage = $\frac{S_e}{\bar{X}}$	Percentage = $\frac{S_e}{\bar{X}}$
Percentage = $\frac{10{,}774.6112}{48{,}863.99}$ x 100%	Percentage = $\frac{1{,}522.78}{48{,}863.99}$ x 100%
Percentage = 22.05%	Percentage = 3.11 %

D. Kesimpulan dan Saran

1. Kesimpulan:

Using 74 propositions of rules, the result of forecasting using fuzzy mamdani logic reached 1,032,300 Tons. While the forecasting results using linear regression with the equation $\hat{Y}_i$ = -4,800.976 + 6.449X_1 + 0.372X_2 + 1.045X_3 - 0.83X_4 + 42.243X_5 is 878,428.07 Tons. Based on the standard error average, multiple linear regression forecasting has a smaller error of 1,522.78 with a percentage of 3.11%. Meanwhile, forecasting using the fuzzy logic method has an error of 10,774.61 with a percentage of 22.05%. It can be concluded that multiple linear regression forecasting has more accurate results than mamdani's fuzzy logic.

2. Saran

For readers of this study, it is hoped that it is better to use data that tends to be more constant so that the forecasting results obtained are more accurate. If the data used is not constant, then the standard deviation or standard error value of the study is greater, resulting in the forecasting result far from the actual result.

For researchers who want to conduct further research, it is expected to use data on time intervals longer or a period of more than 6 years, so that the forecast obtained is more accurate.

E. Daftar Pustaka

Arfina, S., & Tommi, P. (2018). Pemilihan Metode Peramalan dalam Manajemen Persediaan Produk Pertanian. *Agroteknologi dan Agribisnis*, *2*(1).

Biyanto, D. W., & Siti, R. (2019). Pengaruh Employee Terhadap Kinierja Karyawan Direktorat Sistem Informasi dan Trnasformasi Digital Institut Pertanian Bogor. *Jurnal Manajeman dan Organisasi*, *10*(2), 133-146.

Ishak, A. (2010). *Manajemen Operasi*. Yogyakarta: Graha Ilmu.

Isnain, G., Ainie, K., Muhammad, F. M., Ari, B. P., & Jatmiko, E. S. (2021). Penentuan Curah Hujan Berdasarkan Input Cuaca Menggunakan Metode Logika Fuzzy Mamdani. *Jurnal Pengabdian Vokasi, 02*(01).

Jay, H., & Barry, R. (2009). *Manajemen Operasi Buku 1 Edisi 9.* Jakarta: Salemba .

Sri, K. (2002). *Analisis Desain Sistem Fuzzy Menggunakan Tol Box Matlab.* Yogyakarta: Graha Ilmu.

tobakab.go.id. (2022). Retrieved Februari 17, 2022, from https://tobakab.go.id/

tobasamosirkab.go.id. (2022). Retrieved April 10, 2022, from https://tobasamosirkab.bps.go.id/

Tri, A., Made, A., & Lien, D. (2019). ANALISIS FAKTOR-FAKTOR YANG MEMEGARUHI PRODUKSI PADI SAWAH DI DESA UETOLI KEACAMATAN AMPANA TETE KABUPATEN TOJO UNA-UNA. *Jurnal Agroland, 2*(26), 111-122.

Yosefina, M. F., I Made, N. T., & I Dewa, G. A. (2018). Faktor Faktor yang Mempengaruhi Produksi Padi Sawah di Kecamatan Biboki Moenleu Kabupaten Timor Tengah Utara Provinsi Nuasa Tenggara Timur. *Jurnal Agribisnis Lahan Kering, 34*, 64-66.

Yulia, R. S., Lidya, R. A., & Maha, R. (2021). Perbadningan Analisis Logika Fuzzy dan Regresi Linier Berganda dalam Menentukan Produksi Beras Nasional. *Komtekinfo, 4*, 239-248.