FORECASTING USING FUZZY LOGIC RELATIONSHIP GROUPS METHOD ON IMMIGRATION PUBLIC SERVICE'S PASSPORT PRODUCTION DATA

Aidil Akbar[1], Andi Farmadi[2], Muliadi[3], Dwi kartini[4], M. Itqan Mazdadi[5]
[1,2,3,4,5] Computer Science FMIPA ULM
A. Yani St. KM 36 Banjarbaru, South Kalimantan
e-mail: j1f115021@mhs.ulm.ac.id

Abstract

Stationarity is a term used to describe the pattern of trend in time series data. In time series data, this term known as stationary and non-stationary. Non-stationary data is a data that has an unstable pattern of increase and decrease. This condition makes forecasting more difficult. Fuzzy Time Series is one of many forecasting methods that can be used. In this algorithm, adding order is an option that can be used to increase the accuracy of the method. Application up to order three are carried out to determine the effect of addition order to the resulted accuracy value. Experiment is done by applying the used method to the data which is divided into several amounts of data. From the experiment, the average accuracy value of the three Order of Fuzzy Logic Relationship Groups (FLRG) Order-1, Order-2, and Order-3 are 84.06719%, 85.77546%, 92.01034%. FLRG Order-3 has the largest accuracy value while the smallest accuracy value is owned by FLRG Order-1. From this, it is proven that the addition of order able to reduce the error in accuracy value while forecasting using non-stationary data but the accuracy produced by different amounts of data are erratically increasing and decreasing. the experiment concluded that the order, the amount of data, and the data pattern are factors that affect the accuracy result.

Keywords: Passport Data, Non-Stationary Data, Prediction, *Fuzzy*, *Fuzzy Logic Relationship Groups*.

1. INTRODUCTION

Stasionerity is a term to describe the pattern of increases and decreases in data. In data *time series*, the terms stationary data and non-stationary data are known. Nachrowi & Usman [7] revealed that a data set is said to be stationary if the mean and variance of the data *time series* do not change systematically over time, or some experts state that the mean and variance are constant. For forecasting methods, especially *Fuzzy Time Series*, will produce better accuracy when applied to stationary data compared to using non-stationary data. This is because the mean and variance of the data are not constant so that the regression for these data usually leads to ambiguous regression.

The research that tested themethod *Fuzzy Time Series* on non-stationary data has been carried out by several researchers. Desmonda et al. [3] in his research using rainfall data in the city of Pontianak which is known to experience climate change due to the effects of globalization. The data is then divided into 15 time periods, namely from 1 year, 2 years, 3 years, to 15 years. The results of the study showed that the accuracy obtained was high, but the accuracy continued to decrease with the increasing number of time periods.

In a study conducted by Boaisha & Amaitik [2] they useddata *Arabian Oil Gulf* taken every year from 1977 to 1999. The research aimed to test the method *Fuzzy Time Series* to forecast data problems based on the average the average interval of the data. The results of this study indicate that the method tested has a low accuracy value.

From the two studies above, it can be concluded that non-stationary data will produce an indefinite accuracy value. This is in accordance with what was expressed by Efendi et al. [4] that forecasting non-stationary data is more difficult than using stationary data. From the statement put forward, it implies that forecasting should only use stationary data to make it easier. However, in reality a lot of data found in the field is non-stationary data such as transaction data, market financial data, rainfall data, and others. It is estimated that the difficulty experienced is that data that is not stagnant will have less accurate accuracy, so itis necessary to apply techniques to strengthen the performance of forecasting methods in order to produce more accurate accuracy. One technique that can be used is the addition of order.

For themethod *Fuzzy Time Series*, the addition of an order can help improve the accuracy of the forecast. In the research of Kartini et al. [5] they use five orders, namely order-1 to order-5. The results show that, orders 4 and 5 have the highest accuracy while order-1 has the lowest accuracy. However, the research conducted by Singh [8] showed that the accuracy obtained from the 4th order was smaller than the 3rd order.

From these two studies, it can be assumed that the addition of order is able to help themethod *Fuzzy Time Series* to produce better accuracy but the resulting accuracy value still depends on the structure.

Based on the description above, regarding non-stationary data and the addition of orders, it is necessary to conduct research to analyze the effect ofadding orders to themethod *Fuzzy Time Series* to determine the value of forecasting accuracy when tested using non-stationary data.

2. RESEARCH METHODOLOGY

The research methodology carried out in this study is as follows:

a. The sample used is the passport data of the Class I TPI Banjarmasin Immigration office which is known to be non-stationary from thestationary test using the Dickey-Fuller method. The sample obtained is then divided into several amounts according to research needs.

b. Data *Processing*, in this step a class and interval search is carried out to obtain the number of classes, intervals, middle values, and class sets for the fuzzyfication process.

c. Fuzzyfication, this process converts data in the form of numbers into linguistic variables (fuzzy variables).

d. *Fuzzy Logic Relationship,* perform fuzzy relationships by looking at the time sequence of fuzzy values. The first sequence is called the *Current State* and the next sequence is called the *next State*. In the next line the *next state* becomes the *current state* and the sequence after it becomes the *next state*. The process is carried out until all data has a relation.

e. *Fuzzy Logic Relationship Groups*, the relationships that have been obtained are grouped. The group name is taken from the *current state* and the contents of the group are thevalues *next state* that have *current state* the same.

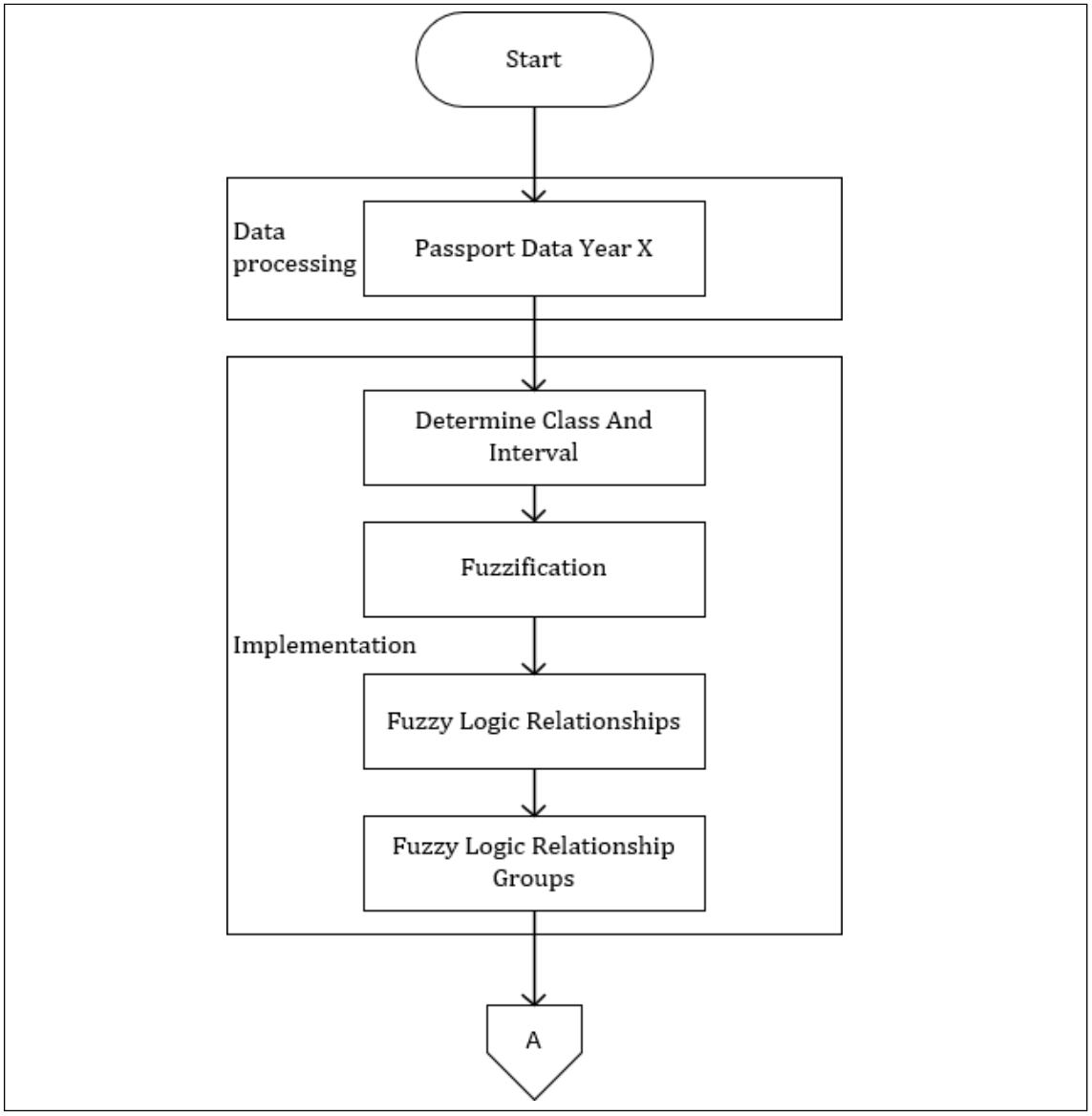
f. Weighting is done by counting the number of occurrences of the *next state* in the *current state*.

g. Forecasting value is obtained by calculating the weight value with the middle value of the class.

h. Deffuzification, The deffuzification value is obtained by changing the fuzzy value into a forecasting value according to the fuzzy value of thefuzzyfication process

i. Accuracy Search, the forecasting value of the deffuzification is compared with the original data for accuracy by using theformula *Mean Absolute Percentage Error*.

The scheme and modeling of the research flow are presented in the flow below.



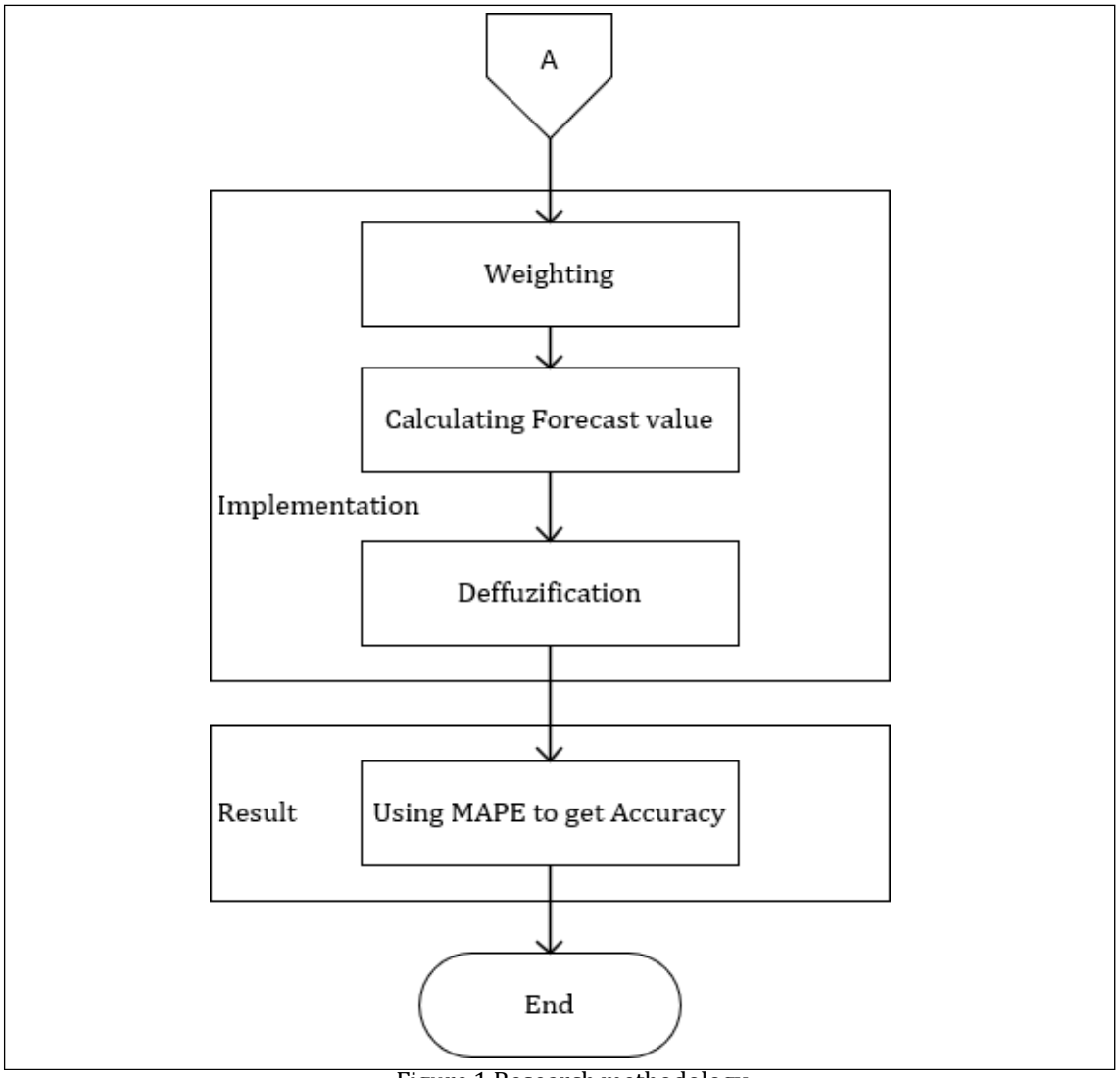


Figure 1 Research methodology

3. RESULTS AND DISCUSSION

3.1. Data collection

Research data taken from the immigration office is passport data per month from 2010 to 2019 with a total of 120 data. This data is then divided into several amounts as shown in the table below.

Table 1 data distribution table

No	Year Data	Range Year	Amount of Data
1	Passport Data for 2019	1 Year	12
2	Passport Data for 2018-2019	2 Years	24
3	Passport Data for 2017-2019	3 Years	36
4	Passport Data for 2016-2019	4 Years	48
4	Passport Data for 2015-2019	5 Years	60
6	Passport Data for 2014-2019	6 Years	72
7	Passport Data for 2013-2019	7 Years	84
8	Passport Data for 2012-2019	8 Years	96
9	Passport Data for 2011-2019	9 Years	108
10	Passport Data for 2010-2019	10 Years	120

3.2. Data *Processing*

In data processing, class and interval searches are carried out to search for class sets. Class search is done using the following formula:

$$k = 1 + 3.3 \log (n) \quad ...(1)$$

then search for the values of N1 and N2 by rounding the maximum and minimum values. The formula is:

$$N1 = Min(x) - a$$
$$N2 = Max(x) + b \quad ...(2)$$

The formulation of the values of a and b for rounding follows the rounding rules in the following table:

Table 2 Rounding Base

Range	Basis
0.1 – 1.0	0,1
1.1 – 10	1
11 – 100	10
101 - 1000	100

After obtaining the number of classes and the values of N1 and N2, then the interval value can be searched using the formula:

$$c = (N2-N1) / k \quad ...(3)$$

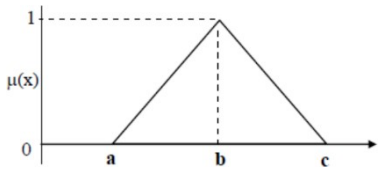
All The above process is applied to the data that has been divided in table 1 and the result is the class and interval values in the following table below:

Table 3 Class and Interval

No.	Year Data	Amount of Data	N1	N2	Class	Interval
1	2019	12	1500	3800	5	460
2	2018-2019	24	800	4400	5	720
3	2017-2019	36	800	4400	6	600
4	2016-2019	48	800	4400	6	600
4	2015-2019	60	800	4400	7	514.2
6	2014-2019	72	800	4400	7	514.2
7	2013-2019	84	800	4400	7	514.2
8	2012-2019	96	800	5500	7	671.4
9	2011-2019	108	800	5500	8	587.5
10	2010-2019	120	800	5500	8	587.5

3.3. Fuzzyfication

Fuzzyfication is done by looking at the location of the numbers in the class set. This is done by using the fuzzy membership degree equation [1].



$$\mu(x) = \begin{cases} 0, & x \le a \; atau \; x \ge c \\ \frac{x-a}{b-a}, & a \le x \le b \\ \frac{c-x}{c-b}, & b \le x \le c \end{cases} \quad ...(4)$$

Following are the results of fuzzyfication for 2019 data.

Table 4 Fuzzyfication of 2019 Data

No.	Data	Class	Fuzzy Value
1	3663	5	A5
2	3187	4	A4
3	3695	5	A5
4	1881	1	A1
5	1774	1	A1
6	1547	1	A1
7	3091	4	A4
8	3132	4	A4
9	2836	3	A3
10	3159	4	A4
11	2912	4	A4
12	2691	3	A3

3.4. Fuzzy Logic Relationship

Fuzzy Variables are then related to the next data (*Next State)*. Below is an example of the relationship for fuzzy Fuzzy Order-1 method.

Table 5 Fuzzy Order-1 Relationships

No	Data	Fuzzy Value	Current State	→	Next State
1	3663	A5			
2	3187	A4	A5	→	A4
3	3695	A5	A4	→	A5
4	1881	A1	A5	→	A1
5	1774	A1	A1	→	A1
6	1547	A1	A1	→	A1
7	3091	A4	A1	→	A4
8	3132	A4	A4	→	A4
9	2836	A3	A4	→	A3
10	3159	A4	A3	→	A4
11	2912	A4	A4	→	A4
12	2691	A3	A4	→	A3

Table 6 Fuzzy Order-2 Relationships

No	Data	Fuzzy Value	Current State			Next State
1	3663	A5				
2	3187	A4				
3	3695	A5	A5	A4	→	A5
4	1881	A1	A4	A5	→	A1
5	1774	A1	A5	A1	→	A1
6	1547	A1	A1	A1	→	A1
7	3091	A4	A1	A1	→	A4
8	3132	A4	A1	A4	→	A4
9	2836	A3	A4	A4	→	A3
10	3159	A4	A4	A3	→	A4
11	2912	A4	A3	A4	→	A4
12	2691	A3	A4	A4	→	A3

Table 7 Fuzzy Order-3 Relationships

No	Data	Fuzzy Value	Current State				Next State
1	3663	A5					
2	3187	A4					
3	3695	A5					
4	1881	A1	A5	A4	A5	→	A1
5	1774	A1	A4	A5	A1	→	A1
6	1547	A1	A5	A1	A1	→	A1
7	3091	A4	A1	A1	A1	→	A4
8	3132	A4	A1	A1	A4	→	A4
9	2836	A3	A1	A4	A4	→	A3
10	3159	A4	A4	A4	A3	→	A4
11	2912	A4	A4	A3	A4	→	A4
12	2691	A3	A3	A4	A4	→	A3

3.5. Fuzzy Logic Relatonship Groups The

relations that have been obtained are then grouped based on the current state and next state. Below is a table for grouping Orders 1 to 3.

Table 8 FLRG Order-1

Current State	Next State
A1	A1 A1 A4
A2	-
A3	A4
A4	A3 A3 A4 A4 A5
A5	A1 A4

Table 9 FLRG Order-2

Current State	Next State
A1 A1	A1 A4
A1 A4	A4
A3 A4	A4
A4 A3	A4
A4 A4	A3 A3
A4 A5	A1
A5 A1	A1
A5 A4	A5

Table 10 FLRG Order-3

Current State	Next State
A1 A1 A1	A4
A1 A1 A4	A4
A1 A4 A4	A3
A3 A4 A4	A3
A4 A3 A4	A4
A4 A4 A3	A4
A4 A5 A1	A1
A5 A1 A1	A1
A5 A4 A5	A1

3.6. Weighting

Weighting is done by counting the number of occurrences of the *next state* in a group. The following is an example of the weighting of Order-1.

Table 11 Weighting Order-1

Current State	Weighting *Next State*	Total Weight
A1	A1(2) A4	3
A2	-	0
A3	A4	1
A4	A3(2) A4(2) A5	5
A5	A1 A4	2

Table 12 Weighting Order-2

Current State	Weighting *Next State*	Total Weight
A1 A1	A1 A4	2
A1 A4	A4	1
A3 A4	A4	1
A4 A3	A4	1
A4 A4	A3 A3	2
A4 A5	A1	1
A5 A1	A1	1
A5 A4	A5	1

Table 13 Weighting Order-3

Current State	Weighting *Next State*	Total Weight
A1 A1 A1	A4	1
A1 A1 A4	A4	1
A1 A4 A4	A3	1
A3 A4 A4	A3	1
A4 A3 A4	A4	1
A4 A4 A3	A4	1
A4 A5 A1	A1	1
A5 A1 A1	A1	1
A5 A4 A5	A1	1

3.7. Forecasting Value

Obtained by calculating the class value with the weight value that has been obtained in the previous stage. Forecasting results can be seen in the following table.

Table 14 Forecasting Value

Current State	Next State	Weighting Total	Forecasting Value
A1	A1(2) A4	3	2190
A2	-	-	-
A3	A4	1	3110
A4	A3(2) A4(2) A5	5	3018
A5	A1 A4	2	2420

3.8. Deffuzification

The last process in implementing the fuzzy time series method is deffuzification, which is to change the fuzzy value back into a number form based on the forecast value obtained.

Table 15 Deffuzification

No	Data	Value Fuzzi	Deffuzification
1	3663	A5	
2	3187	A4	3018
3	3695	A5	2420
4	1881	A1	2190
5	1774	A1	2190
6	1547	A1	1880
7	3091	A4	3018
8	3132	A4	3018
9	2836	A3	3110
10	3159	A4	3018
11	2912	A4	3018
12	2691	A3	3110

3.9. Accuracy Search Accuracy

search was carried out by comparing the value of the deffuzification result with the original data. Accuracy is obtained from the difference in MAPE values with a value of 100% and then compared to each method in order to obtain the best accuracy value [6]. Calculations are carried out using the formula below.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{At-Ft}{At}\right| \quad ...(5)$$

$$Akurasi = 100\% - (MAPE)\% \quad ...(6)$$

3.10. Implementation Results

By implementing the 1st, 2nd, and 3rd order FLRG methods to the data that has been divided in table 1, the accuracy results are obtained:

Table 16 FLRG Implementation Results Order-1, Order-2 and Order-3

Accuracy Comparison			
Year Data	***FLRG* Order 1**	***FLRG* Order 2**	***FLRG* Order 3**
Passport Data 2019	85.40144%	89.0834%	95.54588%
Passport Data 2018-2019	77 0.05709%	89.9901%	92.39824%
Passport Data 2017-2019	78.01308%	82.71049%	90.7984%
Passport Data 2016-2019	84.16109%	87.7225%	94.34523%
Passport Data 2015-2019	84.15696%	86.43642%	94.18172%
Passport Data 2014-2019	86.80585%	86.83319%	93.02901%
Passport Data 2013-2019	87.0417%	85.70953%	92.81054%
Passport Data 2012-2019	86.77132%	84.91964%	89.53733%
Passport Data 2011-2019	86.77279%	83.34624%	88.97495%
Passport Data 2010-2019	84.49061%	81.00304%	88.48208%
Average accuracy			
Passport Data accuracy	84.06719%	85.77546%	92.01034%

3.11. Discussion

Testing to find the accuracy of the method is carried out by implementing the *Fuzzy Logic Relationship Groups* Order-1, 2, and 3 method to the data that has been divided into several amounts of data. All accuracy results obtained from each amount of data in one order are then searched for the average accuracy. Theresults and the process of finding the average value of this accuracy will be observed to obtain conclusions.

The test is carried out on several amounts of data to then find the average accuracy so that the accuracy results obtained are more reliable. This is because when data with an erratic trend is tested on themethod *Fuzzy Logic Relationship Groups* , the resulting accuracy value does not have a fixed pattern of increase or decrease, but tends to rise and fall along with different amounts of data. So if the search for accuracy values is only done with one sample of data, the resulting accuracy value will be less reliable for the case of non-stationary data.

From the test, the average accuracy for themethod is *Fuzzy Logic Relationship Groups* Order-1, 2, and 3 . respectively 84.06719%, 85.77546%, 92.01034%. The highest average accuracy is owned by FLRG Order-3 and the smallest average accuracy value is owned by *FLRG* Order-1. It can be concluded that the addition of the order can increase the average accuracy of the resulting forecast value. However, if the accuracy value is observed for every change in the amount of data, then in some cases it can be found that the accuracy values follow each other. In the study conducted, the case was found on the error value of *FLRG* Order 1 and 2. For example, in a data 2010-2019, error value *FLRG* Order-1 obtained by 84.49061% while the value of error *FLRG* 2nd Order of 81.00304%. This shows that increasing the Order will not always increase the resulting accuracy.

Increasing the amount of data does not help in increasing the accuracy of the results. This can be seen from the resulting accuracy value does not have a fixed pattern of increase or decrease with the increasing amount of data. In some cases, increasing the amount of data actually increases the value of the resulting error.

4. Conclusion

From the tests that have been carried out, it can be concluded that the addition of order can increase the accuracy of themethod *Fuzzy Logic Relationships Group* when applied to non-stationary data. The average accuracy results by the method of *Fuzzy Logic Relationships Group* Order-1, 2, and 3, respectively, are 84.06719%, 85.77546%, and 92.01034%.

REFERENCES

[1] Bahroini, A., Farmadi, A. and Nugroho, RA, 2016. **Prediction of demand for instant noodle products using the fuzzy Takagi-Sugeno method**. Click-Computer Science Journal Collection, 3(2), pp. 220-230. [1]

[2] Boaisha, SM & Amaitik, SM, 2010. **Forecasting model based on fuzzy time series approach**. In Proceedings of the 10th International Arab Conference on Information Technology-ACIT (pp. 14-16).

[3] Desmonda, D., Tursina, T. & Irwansyah, MA, 2018. **Prediction of Rainfall Amount Using the Fuzzy Time Series Method**. JUSTIN (Journal of Information Systems and Technology), 6(4), pp.145-149.

[4] Efendi, R., Deris, MM & Ismail, Z., 2016. **Implementation of fuzzy time series in forecasting of the non-stationary data**. International Journal of Computational Intelligence and Applications, 15(02), p.1650009.

[5] Kartini, D., Rusdiani, H. & Farmadi, A., 2019. **Analysis of the Effect of Multiple Orders on the Multivariate High-Order Fuzzy Time Series Method for Predicting Reservoir Water Level Predictions**. Journal of Informatics Education and Research, 5(1), pp. 9-17.

[6] Mada, M., Farmadi, A., Budiman, I., Faisal, MR and Mazdadi, MI, 2021. **GRU, AdaGrad, RMSprop, Adam Implementation of GRU and Adam Optimization Method For Stock Price Prediction**. Journal of Data Science and Software Engineering, 2(01), pp.36-45.

[7] Nachrowi, D. & Usman, H., 2006. **Econometrics: Popular and Practical Approaches To Economic and Financial Analysis**. Jakarta: Publishing institution of the Faculty of Economics, University of Indonesia.

[8] Singh, SR, 2007. **A simple time variant method for fuzzy time series forecasting**. Cybernetics and Systems: An International Journal, 38(3), pp. 305-321.