Computational Modeling of ARIMA-based G-MFS Methods: Long-term Forecasting of Increasing Population

Habib Ratu Perwira Negara[1], Maximus Tamur[2], Syaharuddin[3], Tri Herdiawan Apandi[4], Jaka Wijaya Kusuma[5], Hamidah[6]
[1]Universitas Bumigora Mataram, Indonesia, habib.ratu27@gmail.com
[2]Universitas Katolik Indonesia Santu Paulus Ruteng, Indonesia, maximustamur@unikastpaulus.ac.id
[3]Universitas Muhammadiyah Mataram, Indonesia, syaharuddin.ntb@gmail.com
[4]Politeknik Negeri Subang, Indonesia, h.apandi@gmail.com
[5,6]Universitas Bina Bangsa, Indonesia, jakawijayak@binabangsa.ac.id , hamidah@binabangsa.ac.id

ABSTRACT

Every year the population in each region has increased, no exception in the province of NTB. Therefore, to take the policies of upgrading the quality of life of residents in NTB province it is necessary to predict the number of people in the future as a benchmark decision making by the Government. The method used is ARIMA. This research aims to construct the G-MFS-based mathematical model on the ARIMA method which is then continued to determine the number of population predictions in NTB province for the next few years. The data used is the population data in 10 districts in the province of NTB for the last 11 years. Based on the results of the simulation found that of the four mathematical models that became the output of the G-MFS turned out to be the 4th model gives the highest level of accuracy on each data simulated with an average MSE of 0207, while the increase in population in the year 2020 average of 1.15% and for the next 20 years the average increase of 0.62%. The forecasting of the population is expected to provide an important contribution to the Government as a weighting material in the planning, implementation, repair, and formulation of subsequent policies, both in the field of economy, education, health, and reducing the level of poverty in the NTB province in the future.

Key words : ARIMA method; GUI Multimodel Forecasting System; Population forecasting

1. INTRODUCTION

The residents are the ones who are in a region bound by the rules that apply and interact with each other continuously or continually [1], [2]. In sociology, a population is a group of human beings occupying a particular geographical and spatial area. The occurrence of population increase affects the process of development and activity development of a region as well as increasing the need for space and land area [3]. With the increasing population, it also demands the provision of life needs of both physical needs such as housing, facilities and infrastructure, and non-physical such as education, economics, and recreation [4]. Population density often raises problems in the arrangement of the population due to the amount of pressure on land area. In areas of dense population and uneven spread will face problems such as housing problems, job problems, educational problems, food problems, security issues and can impact environmental problems, as well as other social problems.

Therefore, the Government should strive to determine the most appropriate policy to overcome the problems that occur. One government policy is to know the forecasting of the population in the coming year. The forecasting of the population is expected to provide important input for the government as an ingredient in the implementation, repair, and formulation of subsequent policies.

Forecasting is defined as an activity to predict future events by using and considering previous data [5], [6]. Forecasting is a tool in effective and efficient planning, so it can determine when an event occurs and the right decision can be made [7]. Therefore, forecasting is very important by government and private agencies to see the development of institutions that have implications for the direction of policy to be taken in the future [8]. In addition, forecasting is often combined with linear programming to produce maximum gains. Generally, forecasting is done on data in the form of time series that is managed by the Centre of Statistics (called BPS) both the district, the province and the nation because it affects the policy determination by the Government.

In conducting forecasting the required model, method, or approach should be proven, if the accuracy of a model has been tested, more and more people are using it. Nowadays, the development of various methods is presented with algorithms that facilitate users to operate it [9]. However, not all methods can be used in all situations, especially with regard to the data time series. The most frequently used method of data forecasting process in the form of time series includes Autoregressive Integrated Moving Average (ARIMA) [10],

Commented [U1]: Please briefly describe the ARIMA research steps.

[11]. The ARIMA method is a very precise method to overcome the complexity of time series and other forecasting situations [12], [13]. The method of ARIMA can be used to predict historical data with conditions that are difficult to understand the impact on data technically and very accurately for both short-term and long-term forecasting through a given mathematical model. XXX

The research team has developed a computer program is called GUI Multimodel Forecasting System (G-MFS) used to facilitate forecasting processes with representative output and facilitate the research team to read the predicted results. Therefore, the objective to achieve from this research is to find a mathematical model of the population, describe the output of G-MFS, perform long-term data simulations, and analyze policies that are able to be recommended in policy determination by Governments.

2. METHODS

This is a quantitative study by simulating population data to find future predictions using the G-MFS-based ARIMA method. The total population data used is data from 10 districts in the province of NTB for the last 11 years taken from the official link of the NTB Statistical Center (https://ntb.bps.go.id). The data patterns can be seen in Figure 1 below.

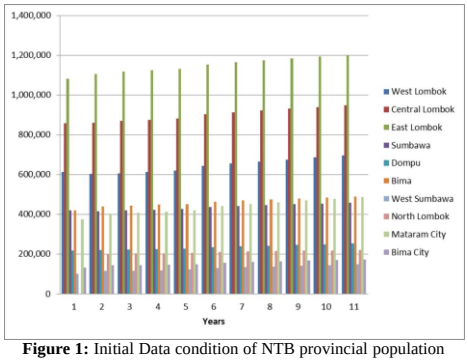


Figure 1: Initial Data condition of NTB provincial population

The forecasting measures are as follows:

2.1 Data collection

At this stage, the research team collects population data in NTB Province by district/city from the official website of the provincial statistics agency of NTB and each district/city BPS from 2009-2019, as well as collecting references on methods of prediction using ARIMA method.

2.2 Data Simulation and Prediction

At this stage, researchers perform data simulations using the G-MFS computer program for the ARIMA method. As for the steps in the simulation phase of data can be seen in Figure 2 below.

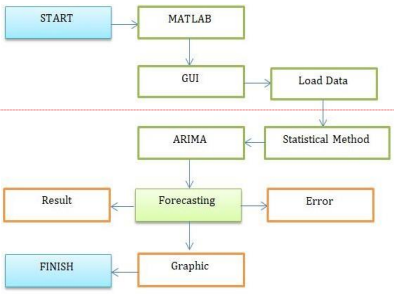
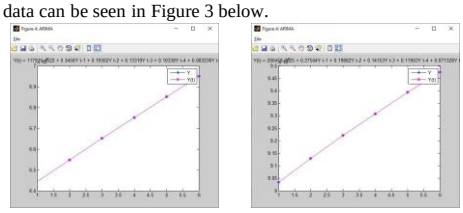


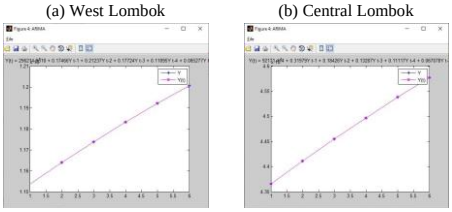
Figure 2: Phase Data Simulation Using ARIMA method

3. RESULT AND DISCUSSIONS

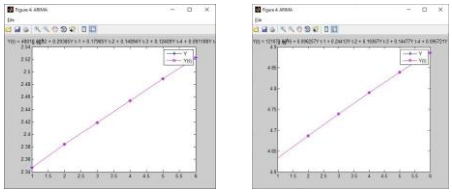
After the researchers conducted a simulation with the total population data of 10 districts/cities obtained 4 models or mathematical equations and the graphical output of the actual data approach and predictive data. Of the four models acquired, the predicted result of the 4th equation is known to have the highest level of accuracy among the 1st, 2nd, and 3rd models, as it has the smallest MSE value. The output of the graph and the mathematical model of the total population data can be seen in Figure 3 below.



(a) West Lombok (b) Central Lombok



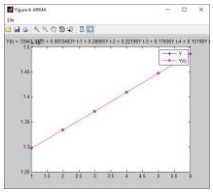
(c) East Lombok (d) Sumbawa



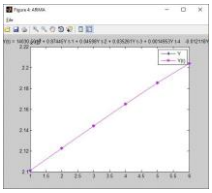
(e) Dompu (f) Bima

Commented [U2]: Your research problem has not been well defined. The novelty of your research has also not been explained..

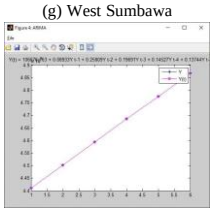
Commented [U3]: Why did you choose this province?



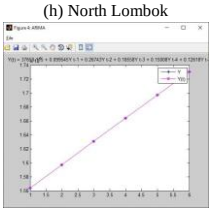
(g) West Sumbawa



(h) North Lombok



(i) Mataram City



(j) Bima City

Figure 3: Graphics Output and Mathematical Models the Population of Each District/City.

Based on Figure 3 above, it is known mathematical model of the population data of each district/city:

1. West Lombok District
   $Y(t_{4\text{-WL}}) = 117915.4522 + 0.3456\ Y_{t-1} + 0.19302\ Y_{t-2} + 0.13319\ Y_{t-3} + 0.10338\ Y_{t-4} + 0.083238\ Y_{t-5}$
2. Central Lombok District
   $Y(t_{4\text{-CL}}) = 200458.3025 + 0.27504\ Y_{t-1} + 0.19882\ Y_{t-2} + 0.14153\ Y_{t-3} + 0.11903\ Y_{t-4} + 0.071328\ Y_{t-5}$
3. East Lombok District
   $Y(t_{4\text{-EL}}) = 296214.9518 + 0.17466\ Y_{t-1} + 0.21237\ Y_{t-2} + 0.17724\ Y_{t-3} + 0.11895\ Y_{t-4} + 0.085277\ Y_{t-5}$
4. Sumbawa District
   $Y(t_{4\text{-S}}) = 92171.414 + 0.31979\ Y_{t-1} + 0.18426\ Y_{t-2} + 0.13287\ Y_{t-3} + 0.11117\ Y_{t-4} + 0.067078\ Y_{t-5}$
5. Dompu District
   $Y(t_{4\text{-D}}) = 48016.0892 + 0.29385\ Y_{t-1} + 0.17985\ Y_{t-2} + 0.14894\ Y_{t-3} + 0.12409\ Y_{t-4} + 0.091188\ Y_{t-5}$
6. Bima District
   $Y(t_{4\text{-B}}) = 121079.0016 + 0.096257\ Y_{t-1} + 0.24413\ Y_{t-2} + 0.19367\ Y_{t-3} + 0.14477\ Y_{t-4} + 0.095721\ Y_{t-5}$
7. West Sumbawa District
   $Y(t_{4\text{-WS}}) = 33043.3323 + 0.0072483\ Y_{t-1} + 0.28805\ Y_{t-2} + 0.22196\ Y_{t-3} + 0.17836\ Y_{t-4} + 0.15198\ Y_{t-5}$
8. North Lombok District
   $Y(t_{4\text{-NL}}) = 14030.5018 + 0.87445\ Y_{t-1} + 0.04598\ Y_{t-2} + 0.035261\ Y_{t-3} + 0.0014953\ Y_{t-4} - 0.012118\ Y_{t-5}$
9. Mataram City
   $Y(t_{4\text{-MC}}) = 106676.763 + 0.08933\ Y_{t-1} + 0.25809\ Y_{t-2} + 0.19691\ Y_{t-3} + 0.14527\ Y_{t-4} + 0.13744\ Y_{t-5}$
10. Bima City
    $Y(t_{4}\text{.BC}) = 37658.115 + 0.099545\ Y_{t-1} + 0.26743\ Y_{t-2} + 0.18558\ Y_{t-3} + 0.15008\ Y_{t-4} + 0.12618\ Y_{t-5}$

From this mathematical model is then obtained prediction results in each district/city in NTB year 2020 according to Table 1 follows.

Table 1: Result Prediction and MSE Value

District/City	Models	Population	MSE
West Lombok	$Y(t_{4\text{-WL}})$	703,558	0.7524
Central Lombok	$Y(t_{4\text{-CL}})$	954,683	0.0863
East Lombok	$Y(t_{4\text{-EL}})$	1,208,734	0.1844
Sumbawa	$Y(t_{4\text{-S}})$	461,010	0.0080
Dompu	$Y(t_{4\text{-D}})$	255,379	0.0001
Bima	$Y(t_{4\text{-B}})$	492,472	0.0302
West Sumbawa	$Y(t_{4\text{-WS}})$	151,796	0.5193
North Lombok	$Y(t_{4\text{-NL}})$	222,201	0.1447
Mataram City	$Y(t_{4\text{-MC}})$	494,245	0.0001
Bima City	$Y(t_{4\text{-BC}})$	175,785	0.3460

If referring from the initial data in Figure 1 and the predicted result in Table 1 then it can be known to increase the population of West Lombok District amounted to 1.23%; Central Lombok District amounted to 0.76%; East Lombok District amounted to 0.68%; Sumbawa district amounted to 0.73%; Dompu District amounted to 1.23%; Bima District amounted to 0.80%; West Sumbawa District amounted to 2.15; North Lombok Regency amounted to 0.81%; Mataram City amounted to 1.55%; and the city of Bima amounted to 1.59%. This results in the fact that West Sumbawa district has an increase in population in 2020, while the lowest in Sumbawa district has been low.

From the simulations and predictions that have been done, researchers highly recommend using ARIMA's method of forecasting in the future using the 4th mathematical model in each G-MFS simulation, as it has the smallest MSE from other equations. Furthermore, using this 4th model can be done predictions for the next few years according to Table 2 follows.

Table 2: Estimated 20 years in the Future of NTB Population

Years	WL	CL	EL	S	D	B	WS	NL	MC	BC
2020	703,558	954,434	1,207,679	461,010	255,214	492,472	151,796	222,201	494,245	175,785
2021	711,636	960,864	1,214,250	464,099	257,960	496,178	155,030	223,900	501,652	178,480
2022	719,259	966,803	1,220,269	466,954	260,553	499,591	158,099	225,511	508,643	181,014
2023	26,437	972,273	1,225,739	469,587	262,986	502,728	161,036	227,040	515,236	183,413
2024	733,169	977,272	1,230,703	471,997	265,251	505,592	163,821	228,490	521,452	185,662
2025	739,446	981,836	1,235,182	474,199	267,348	508,204	166,430	229,867	527,196	187,749
2026	45,331	986,024	1,239,265	476,222	269,306	510,611	168,942	231,173	532,645	189,724
2027	750,849	989,866	1,242,973	478,079	271,134	512,816	171,320	232,412	537,757	191,576
2028	756,022	993,387	1,246,338	479,783	272,838	514,837	173,583	233,588	542,561	193,316
2029	760,867	996,611	1,249,393	481,345	274,426	516,688	175,731	234,704	547,071	194,948
2030	765,406	999,566	1,252,167	482,778	275,906	518,385	177,770	235,763	551,301	196,480
2031	769,660	1,002,274	1,254,685	484,092	277,286	519,941	179,712	236,768	555,280	197,918
2032	773,645	1,004,756	1,256,972	485,298	278,572	521,367	181,555	237,721	559,016	199,269
2033	777,380	1,007,030	1,259,049	486,404	279,771	522,675	183,308	238,625	562,525	200,537
2034	780,879	1,009,114	1,260,934	487,419	280,888	523,873	184,973	239,484	565,821	201,727
2035	784,157	1,011,023	1,262,645	488,350	281,930	524,971	186,555	240,298	568,916	202,844
2036	787,229	1,012,773	1,264,200	489,204	282,901	525,978	188,059	241,071	571,824	203,893
2037	790,107	1,014,377	1,265,611	489,987	283,806	526,901	189,488	241,804	574,554	204,878
2038	792,804	1,015,847	1,266,892	490,706	284,650	527,747	190,846	242,500	577,119	205,803
2039	795,331	1,017,193	1,268,055	491,365	285,437	528,523	192,136	243,160	579,528	206,671

Table 2 above shows that the increasing number of people for the top 20 years in West Sumbawa District is the lowest value in West Lombok with an average total increase of 0.62%. Complete results can be seen according to table 3 below.

Table 3: Increased population rate

District	Average (%)
West Lombok	0.68
Central Lombok	0.36
East Lombok	0.27
Sumbawa	0.36
Dompu	0.62
Bima	0.39
West Sumbawa	1.29
North Lombok	0.49
Mataram City	0.88
Bima City	0.89

Because of this, the government must try to determine the most appropriate policy to tackle the problems that occurred. One government policy is to know the forecasting of the population in the coming year. The forecasting of the population is expected to provide important input for the government as an ingredient in the implementation, repair, and formulation of subsequent policies [4]. XXX

4. CONCLUSION

Forecasting is very important by government and private institutions because forecasting can provide important input and role for the government as a material in the planning, implementation, improvement, and formulation of future policies. From the results of the simulation of population data in the province of NTB using the G-MFS-based ARIMA method is obtained that the equation or the 4th mathematical model provides the highest level of accuracy with an average MSE of 0.207. Furthermore, the predicted result showed that in 2020 the highest population increase occurred in the West Sumbawa district, while the lowest occurred in the Sumbawa district. Then by using a mathematical model found done predictions for the next 20 years, it was obtained the highest increase occurred in West Sumbawa district while the lowest increase occurred in West Lombok district.

From the results of predictions that have been obtained, it is hoped that the Government can make the right decision to do proper development in various areas such as economy, education, health, poverty reduction and many more developments that need to be done according to the number of people in the future.

The advice that needs to be considered for further research is expected to study the sources and references related to forecasting and methods used in this research in order to get even better results.

Commented [U4]: This study has not reviewed the data properly. There is still minimal theory and expert opinion involved in the discussion of this data.

REFERENCES

[1] J. Dai and S. Chen, "The application of ARIMA model in forecasting population data," in *Journal of Physics: Conference Series*, 2019, vol. 1324, no. 1. https://doi.org/10.1088/1742-6596/1324/1/012100

[2] A. R. Alias, N. Y. Zainun, and I. Abdul Rahman, "Comparison between ARIMA and DES methods of forecasting population for housing demand in Johor," in *MATEC Web of Conferences*, 2016, vol. 81. https://doi.org/10.1051/matecconf/20168107002

[3] W. C. Sanderson, S. Scherbov, W. Lutz, and B. C. O'Neill, "Applications of probabilistic population forecasting," in *The End of World Population Growth in the 21st Century: New Challenges for Human Capital Formation and Sustainable Development*, 2013, pp. 85–120.

[4] H. R. P. Negara, Syaharuddin, K. Kiki Riska Ayu, and R. P. N. Habibi, "Analysis of nonlinear models for the acceleration of increasing HDI in Asia," *Int. J. Sci. Technol. Res.*, vol. 8, no. 1, pp. 60–62, 2019.

[5] Syaharuddin, D. Pramita, T. Nusantara, and Subanji, "Computational of Distribution of Wind Speed as Preliminary Information for Fishers: Case Study in Lombok Sea," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 3, pp. 339–342, 2020.

[6] M. Tawarish and K. Satyanarayana, "A review on pricing prediction on stock market by different techniques in the field of data mining and genetic algorithm," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 1. pp. 23–26, 2019. https://doi.org/10.30534/ijatcse/2019/05812019

[7] S. B. Pooja and R. V. Siva Balan, "Linear program boosting classification with remote sensed big data for weather forecasting," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 4, pp. 1405–1415, 2019. https://doi.org/10.30534/ijatcse/2019/58842019

[8] Syaharuddin, D. Pramita, T. Nusantara, and Subanji, "Testing alpha-beta parameters of holt method for time series forecasting," *Int. J. Sci. Technol. Res.*, vol. 8, no. 12, pp. 3653–3656, 2019.

[9] T. Nyoni, "Forecasting total population in Yemen using Box-Jenkins ARIMA models," *Munich Pers. RePEc Arch.*, no. 92433, 2019.

[10] E. P. Patulin, "Scholarship grants prediction using Autoregressive Integrated Moving Average (ARIMA) algorithm," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 3, pp. 551–555, 2019. https://doi.org/10.30534/ijatcse/2019/33832019

[11] D. K. Silalahi, "Forecasting of Poverty Data Using Seasonal ARIMA Modeling in West Java Province," *JTAM | J. Teor. dan Apl. Mat.*, vol. 4, no. 1, p. 76, 2020. https://doi.org/10.31764/jtam.v4i1.1888

[12] T. Alam, "Crude oil export: A comparative analysis using artificial neural network and ARIMA," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 5, pp. 2546–2550, 2019. https://doi.org/10.30534/ijatcse/2019/102852019

[13] E. P. Patulin and R. E. Talingting, "Crime prediction using autoregressive integrated moving average (ARIMA) algorithm," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 3, pp. 720–724, 2019. https://doi.org/10.30534/ijatcse/2019/59832019

Commented [U5]: Please cite several articles from the journal you are referring to.