Islamic International Conference on Education, Communication, and Economics
Mataram, 10-11 May 2025

Faculty of Islamic Studies
Universitas Muhammadiyah Mataram
Mataram City, Indonesia

Performance Evaluation of Quadratic Polynomial Regression Model for Population Growth Prediction

Nur Istiqamah[1], Syaharuddin[2], Vera Mandailina[3]
[1,2,3]Mathematic Education, Universitas Muhammadiyah Mataram, Indonesia
nuristiqamahhh.13@gmail.com

Abstract: This study aims to assess the effectiveness of the quadratic polynomial regression model in forecasting population growth in Dompu Regency. The approach used is quantitative-experimental with a focus on numerical analysis and application of mathematical models. Secondary data in the form of annual population numbers from 2016 to 2025 were obtained from the Central Statistics Agency (BPS) of Dompu Regency. Modeling was conducted using MATLAB software, while model accuracy was evaluated through two metrics: Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). The forecasting results showed irrational predicted values in 2026 and 2027, including negative population estimates. In addition, the MSE value of 3,456,657,665.76 and MAPE of 8281.27% indicate a very high level of prediction error. These findings indicate that the quadratic polynomial regression model is not suitable for describing the dynamics of population growth in the region. Therefore, it is necessary to explore alternative models, such as time series methods or machine learning-based approaches, in order to obtain more accurate prediction results and support population policy formulation more effectively.

Keywords: Quadratic Polynomial Regression, Demographic Prediction, Mse, Mape, Dompu District.

Article History:
Received: 30-04-2025
Online : 19-05-2025

A. INTRODUCTION

Estimating population size is one of the challenges in demographic analysis, where factors such as births, deaths, and migration affect population change (Setyonaturi & Aninditya, 2019). To obtain a more accurate prediction, quadratic polynomial regression is used, as this model can handle nonlinear relationships by including a power-of-two component in the independent variable (Sari et al., 2018). This approach is effective for describing curvature in the data, making it more suitable for forecasting population growth from 2016 to 2025 which is difficult to explain with a linear model.

In predicting population growth, various methods can be applied, such as linear regression, exponential models, and machine learning algorithms (Raharjo, Nugraha, et al., 2021). However, linear regression is often insufficient to capture nonlinear relationships in the data, while exponential models are more suitable for modeling consistent growth, although they are less effective in handling data fluctuations (Ni'matuzzahroh, 2018). Alternatively, quadratic polynomial regression is superior, as it is able to fit the data with a parabolic curve that more accurately describes nonlinear trends. The model is also evaluated using the coefficient of determination (R-squared) to measure how well the model matches the data at hand (Rohmah & Wilandari, 2023).

Previous research indicates that quadratic polynomial regression is effective in analyzing nonlinear trends in population data, as expressed by Akhmad et al., (2022) related to fluctuations in birth and migration rates. In a study of the Chinese population, polynomial regression successfully predicted demographic changes amidst declining fertility rates (Arumsari et al., 2021). Although polynomial regression gives good results, comparison with other models such as ARIMA shows that ARIMA gives more accurate results (Dinanto, et al., 2021). Therefore, further research is needed to address issues such as overfitting, by considering the application of hybrid models to improve prediction accuracy.

Previous research indicates that machine learning methods, such as Random Forest and Informer, are more effective than traditional models in forecasting population growth. According to Tantyoko et al., (2023) Random Forest is able to handle complex data and provide more precise predictions. Muflihah, (2017) also shows that this model is more accurate than logistic regression. The Informer model, which considers multiple variables, also showed better predictive results. Although logistic growth models and multiple linear regression are still used, they have limitations in handling more complex data (Balaka, 2022). While machine learning has shown significant progress, traditional models remain useful in situations with limited data (Cahyono et al., 2023).

While quadratic polynomial regression has been shown to be effective in the prediction of various phenomena, its application in predicting population growth is still under-researched, especially in terms of evaluation based on annual data. Most studies highlight the use of machine learning models or other statistical models, with quadratic regression often only being used as a comparison. This study aims to explore the performance of quadratic polynomial regression models in forecasting population growth between 2016 and 2025, and contribute to helping policy makers choose the right prediction model by considering the balance between complexity and accuracy.

B. METHOD

This study is a quantitative - experimental research that aims to analyze the population growth pattern in Dompu district through the application of quadratic polynomial regression model. The quantitative approach was chosen because this research focuses on numerical data processing and the use of mathematical models to forecast non-linear population growth patterns. This regression model allows the identification of trends that develop over a period of time and provides predictive results that are relevant for demographic planning. The general formula for quadratic polynomial regression is as follows:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon \quad (1)$$

The data used in this study is secondary data obtained from the Central Bureau of Statistics (BPS) of Dompu Regency, namely the annual population for the period 2016 to 2025. After the data was collected, the tabulation and preprocessing stages were carried out to ensure the completeness and consistency of the data, including the cleaning of invalid values. Then, the data was analyzed using MATLAB software to build a quadratic polynomial

regression model. This stage includes applying mathematical models, writing computational algorithms, calculating predictions, and creating graphical visualizations that compare actual data with predicted results.

To evaluate the accuracy of the model, two parameters are used, namely Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). MSE is used to measure the average squared difference between the actual value and the predicted value, while MAPE calculates the average percentage of prediction error against the actual data. The results of this study are expected to provide an overview of the model's accuracy in predicting population numbers, as well as provide insight into the growth patterns that occur. The findings are also expected to serve as a basis for more effective population policy and development planning in Dompu. The research flow describing the stages of the research process is presented in the following chart:

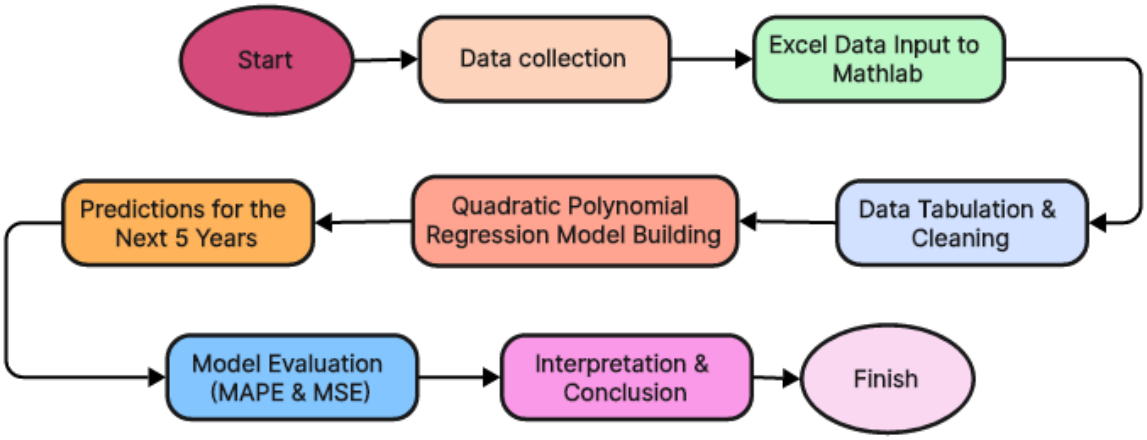

Figure 1. Computing System Design Procedure

C. RESULTS AND DISCUSSION

1. Data Description

Based on the results of the descriptive statistics, the data shows a variation in population throughout the period 2015 to 2024. The minimum value of 236.70 reflects the lowest population recorded, while the maximum value of 252,288.00 recorded the highest population during the same period. The average population over the ten years was 98,990.95, indicating that population figures tend to hover around this value. The standard deviation of 127,507.23 indicates a considerable spread between the years with the lowest and highest population numbers, indicating significant fluctuations. Overall, the data illustrates an unstable trend, with considerable fluctuations in population, although there is a tendency for changes to be influenced by certain factors each year.

Table 1. Descriptive Data

Data Descriptive Statistics	Value
N (Number of Years)	10
Minimum	236.70
Maximum	252288.00
Average (Mean)	98990.9500
Standard Deviation	127507.2343

2. Forecasting Results and Decision Making

Forecasting the population of Dompu district for the period 2025 to 2029, using a quadratic polynomial regression model, resulted in population predictions for each of these years. The results of this forecasting are expected to form the basis for formulating more effective policies for managing population growth in the future, which can be seen in the following table.

Table 2. Results of Population Forecasting in Dompu District for 2025-2029

Year	Population Forecasting (Thousand)
2026	-1.9381
2027	-1.2554
2028	0.1361
2029	2.2362
2030	5.0451

To clarify the predicted trends, a population prediction graph for the period 2025 to 2029 is presented. This graph provides a visual illustration of the projected growth or decline in population over the period, based on the results of the forecasting that has been done.

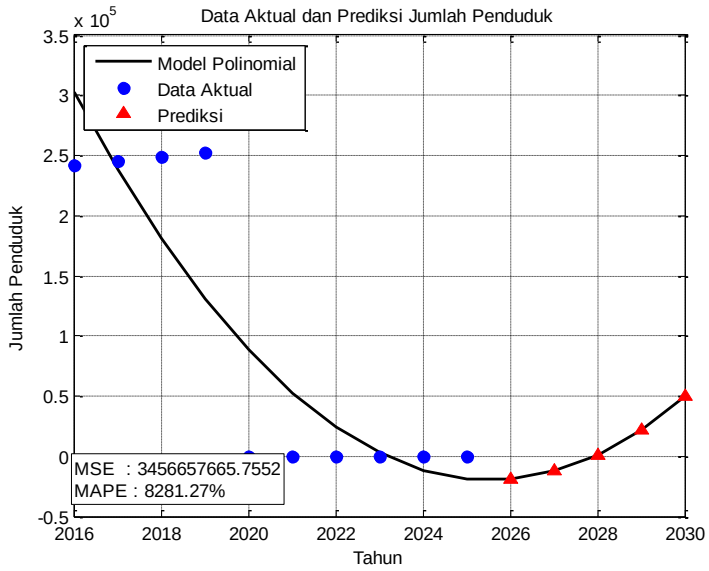

Figure 2. Actual and Predicted Data Approach

Based on the prediction results obtained through the application of the quadratic polynomial regression model, the estimated population in Dompu Regency for the period 2026 to 2030 can be obtained. The mathematical model used in this forecasting process is formulated as follows:

$$y = 69756 - 35697_x + 3543.6x^2 + \varepsilon \quad (2)$$

Using a quadratic polynomial regression model, the population of Dompu district was estimated for the years 2026 to 2030. However, the predictions in the first two years show negative values, namely -1.94 thousand people in 2026 and -1.26 thousand people in 2027, which is logically impossible. This shows that the model has not been able to accurately represent the real condition of the population at the beginning of the forecasting period. The very high Mean Squared Error (MSE) value of 3,456,657,665.76, and the Mean Absolute Percentage Error (MAPE) of 8281.27% also indicate a significant level of error, so the prediction results are not yet suitable as a basis for policy making. Therefore, it is necessary to evaluate or use alternative forecasting models that are more suitable so that the results can be used as a reference in regional development planning that is more targeted.

D. CONCLUSIONS AND SUGGESTIONS

The forecasting results using the quadratic polynomial regression model identified several significant problems. First, the population predictions for 2026 and 2027 show negative values, namely -1.94 thousand people in 2026 and -1.26 thousand people in 2027, which is logically impossible because the population cannot be negative. This indicates that the model has not been able to provide the correct estimation for that period. In addition, the model evaluation indicators, namely the Mean Squared Error (MSE) of 3,456,657,665.76 and the Mean Absolute Percentage Error (MAPE) of 8281.27%, reflect a very large error rate. Therefore, this model cannot be used as a reference in policy making. Thus, further analysis or the application of a more suitable alternative model is required to obtain more accurate predictions.

This study shows that the applied model cannot provide accurate predictions of population at the beginning of the forecasting period. This is due to the unsuitability of the model in describing the pattern of population development. Therefore, it is necessary to develop more appropriate forecasting models, such as time series models or machine learning approaches that can more effectively handle the fluctuations and complexity of demographic data patterns. Further research is essential to produce more accurate predictions and support the formulation of more appropriate policies in managing population growth in the future.

REFERENCES

Akhmad, D., Peirisal, T., & Komara, M. (2022). Faktor-faktor Determinan Kebijakan Pengendalian Laju Pertumbuhan Penduduk Kabupaten Subang 2017-2022. *JURNAL STUDI …*.

Arumsari, M., & Dani, A. (2021). Peramalan Data Runtun Waktu menggunakan Model Hybrid Time Series Regression – Autoregressive Integrated Moving Average. *Jurnal Siger Matematika*. https://doi.org/10.23960/jsm.v2i1.2736

Balaka, M. Y. (2022). Metode penelitian Kuantitatif. *Metodologi Penelitian Pendidikan Kualitatif*.

Muflihah, I. Z. (2017). Analisis Financial Distress Perusahaan Manufaktur Di Indonesia dengan Regresi Logistik. *Majalah Ekonomi*.

Najwa Fathiro Cahyono, Khurrotul 'Uyun, & Siti Mukaromah. (2023). Etika Penggunaan Kecerdasan Buatan Pada Teknologi Informasi. *Prosiding Seminar Nasional Teknologi Dan Sistem Informasi*. https://doi.org/10.33005/sitasi.v3i1.334

Ni'matuzzahroh, L. (2018). Pemodelan Dan Peramalan Jaringan Syaraf Tiruan Dengan Peubah Input Arima Dan Garch. In *Repository.Ub.Ac.Id*.

Raharjo, A. B., Dinanto, Z. Z., Sunaryono, D., & Purwitasari, D. (2021). Prediksi Akumulasi Kasus Terkonfirmasi Covid-19 Di Indonesia Menggunakan Support Vector Regression. *Techno.Com*. https://doi.org/10.33633/tc.v20i3.5062

Raharjo, A. B., Nugraha, E. A., Arunanto, F. X., Sunaryono, D., Baskoro, F., Purwitasari, D., & Irfan Subakti, M. M. (2021). Pendampingan Modul Pengumpulan dan Pelaporan Data pada Aplikasi Penelusuran COVID-19 untuk Dinas Kesehatan Jawa Timur. *SEWAGATI*. https://doi.org/10.12962/j26139960.v5i1.8110

Rohmah, D. D., & Wilandari, A. (2023). Pengaruh Proses Rekrutmen Dan Seleksi Terhadap Kinerja Karyawan Menggunakan Regresi Linier Berganda. *Jurnal Pariwisata Bisnis Digital Dan Manajemen*. https://doi.org/10.33480/jasdim.v2i2.4580

Sari, D., Asnah, A., & Siswadi, B. (2018). Introduction to the Implementation of Econometrics (Pengantar Penerapan Ekonometrika). *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3254684

Setyonaturi, D., & Aninditya, F. (2019). Kajian sektor kesehatan: Transisi Demografi dan Epidemiologi: Permintaan Pelayanan Kesehatan. *Direktorat Kesehat Dan Gizi …*.

Tantyoko, H., Sari, D. K., & Wijaya, A. R. (2023). Prediksi Potensial Gempa Bumi Indonesia Menggunakan Metode Random Forest Dan Feature Selection. *IDEALIS : InDonEsiA JournaL Information System*. https://doi.org/10.36080/idealis.v6i2.3036