UNIVERSITI TEKNOLOGI MARA

MATHEMATICS AND STATISTICS

UNDERGRADUATE RESEARCH PROCEEDINGS 2025

UiTM CAWANGAN NEGERI SEMBILAN

Forecasting Maternal Mortality Rate in Malaysia

Nurul Asyiqin Sahrulnizam[1], Nor Akmal Md Noh [1,*]
[1]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Negeri Sembilan Seremban Campus, 70300 Negeri Sembilan*
[*]norak029@uitm.edu.my

Abstract

Maternal mortality remained a significant issue in Malaysia despite the advancements in healthcare. A key factor affecting maternal mortality rates is the decline in the number of pregnancies and live births. As Malaysia faces economic challenges such as rising living costs, stagnant wages, and financial instability, more women are choosing to delay or avoid pregnancy altogether. The primary objective of this study is to examine the pattern of maternal mortality in Malaysia from the year 1946 to 2023 and to estimate the maternal mortality rates for the next three years in Malaysia using Box-Jenkins methodology. The findings of this study showed that the pattern of maternal mortality rates in Malaysia from 1946 to 2023 was seen to be exponentially decaying. The data was divided into training and test sets for ARIMA analysis to assess the accuracy of the model. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) were used to evaluate the data's stationarity to identify the model. Several ARIMA models were proposed, including ARIMA (1,1,1), ARIMA (1,1,2), ARIMA (2,1,1) and ARIMA (2,1,2). Based on its capacity to produce white noise errors, its effectiveness in fitting the data, and its precision in producing a precise forecast, the ARIMA (1,1,1) model was determined to be the best model. ARIMA (1,1,1) was determined to be the best model for forecasting maternal death rates in Malaysia based on the forecast accuracy and how well the model's prediction matched the actual data. Maternal mortality rates were estimated to be 20.2, 24.6 and 21.1 in 2024, 2025 and 2026 respectively. Reducing maternal mortality requires better healthcare access, training, and public awareness. Future research should explore advanced models and higher-frequency data for improved accuracy.

Keywords*:* Maternal Mortality, Forecasting, Box-Jenkins Methodology, ARIMA model

Introduction

Maternal mortality remains a significant health issue globally, especially in developing countries like Malaysia. Maternal mortality, defined as the death of a woman during pregnancy or within 42 days after childbirth, regardless of the stage or location of the pregnancy, resulted from a cause linked to or worsened by the pregnancy or its care, rather than from an accidental or unrelated cause. In Malaysia, maternal mortality rates are measured by the number of maternal deaths per 100,000 live births [1]. Despite medical advancements, socioeconomic factors, healthcare accessibility, and quality of maternal services, it still influences maternal mortality rates [2].

The United Nations' Sustainable Development Goals (SDGs) aims to the worldwide maternal mortality ratio until it reaches less than 70 per 100,000 live births by 2030, with no country's rate surpassing double the global average [3]. As a member state, Malaysia had committed to meeting this goal. Nonetheless, achieving this goal requires identifying high-risk groups, implementing certain policies to address issues and consistent monitoring of patterns in maternal mortality. While Malaysia has shown progress in reducing these rates, challenges persist due to economic pressures, limited rural healthcare access, and disparities in maternal care [4]. Declining fertility rates, influenced by economic uncertainty and improved access to family planning, have contributed to fewer pregnancies and reduced maternal deaths [5].

The decrease in pregnancy and birth rates affected maternal mortality statistics, but it also brought broader concerns about reproductive health and demographic patterns in the country. Although fewer pregnancies meant a reduction in pregnancy-related health complications, there was growing concern that a sustained decline in birth rates could have significant social and economic impacts over time [6]. However, this demographic shift raises concerns about an aging population and economic impacts, including labor shortages and increased pressure on public resources [7]. Addressing maternal health care gaps and improving support for working mothers are critical steps towards encouraging family planning [8]. The main objective of the study is to forecast the maternal mortality rates in Malaysia for the years 2024 until 2026.

Methodology

Data Set

The data used in this study are the yearly maternal mortality rates in Malaysia recorded by the Department of Statistics Malaysia (DOSM) from the year 1946 to 2023. The analysis includes the year and the annual maternal mortality rates. The data on maternal death cases in Malaysia from 1946 to 2023 were split into two parts in this research. In the estimation phase, the model was trained to identify the trends and patterns using 75% of the data, which covered 1946 to 2003. The evaluation phase took place over the remaining 25%, which ran from 2004 to 2023. By contrasting the model’s predictions with the actual values from this frame, this phase evaluated the model’s performance.

Autoregressive Integrated Moving Average

In the Box-Jenkins methodology, stationarity is one of the key assumptions, as it ensures that the statistical properties of a time series remain constant over time. Graphical tools such as the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) are useful for identifying stationarity. The ACF shows the variation of the autocorrelation with different lags, where consistently low or near-zero values across all lags indicate stationarity. The PACF highlights the correlation between two points after accounting for the effects of intermediate points. If the first PACF value is close to one and subsequent values are near zero, the series is likely non-stationary. In addition, the Augmented Dickey-Fuller (ADF) test is a statistical method used to formally test for stationarity by checking for the presence of a unit root. The null hypothesis states that the data is non-stationary. If the p-value obtained from the test is less than 0.5, the null hypothesis is rejected, indicating that the data is stationary.

Model Identification

The Box-Jenkins methodology was used to identify the best-fitting model for the data. During this process, the ACF and PACF were the key tools. If the data were not stationary, differencing was applied to make it stationary. After achieving stationarity, the ACF and PACF diagrams were analyzed for patterns. If the ACF decayed exponentially and the PACF had significant spikes, the data was suited for an Autoregressive (AR) model, denoted as AR(p), where p is the number of spikes, a Moving Average (MA) model, denoted as MA(q), was used, where q represents the number do significant spikes in the ACF. The ARIMA model is expressed in equation 1 [9].

$$Y_t = Y_{t-1} + \Phi_1 Y_{t-1} + \ldots + \Phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{1}$$

where Y_t is the maternal mortality rate from 1946 to 2023, $Y_{t-1},\ldots,Y_{t-p}$ is the lagged values of the time series, $\varepsilon_{t-1},\ldots,\varepsilon_{t-q}$ is the lagged values of the error term, while, $\Phi_1,\ldots,\Phi_p$ and $\theta_1,\ldots,\theta_q$ represent the autoregressive parameters and the moving average parameters, respectively.

Model Estimation and Validations

Box-Jenkins models were estimated using sample statistics, which needed validation to ensure accuracy in the representation of the population parameters. Common statistical measures used to select and validate the best ARIMA model included Akaikes's Information Criteria (AIC), Bayesian Information Criterion (BIC), Ljung-Box Test, Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). In general, the AIC value is used to measure the fitness of the Model. While the BIC value was designed to select a model that provided the most accurate forecasts.

$$AIC = e^{\frac{2k}{T}} \left(\frac{\sum_{t=1}^{T} e_i^2}{T} \right) \tag{2}$$

The AIC equation in equation 2, $k = p + q$ gives the number of parameters estimated in the model, where p and q refer to the number of parameters for AR and MA components respectively. While T is the number of series in the dataset.

$$BIC = T^{\frac{k}{T}} \left(\frac{\sum_{i=1}^{T} e_i^2}{T} \right) \tag{3}$$

The Bayesian Information Criteria (BIC) in equation 3, k is the total number of parameters in the model, including the constant term while T is the number of observations in the series. The modified Q statistics, developed by Box and Pierce (1970) and refined by Ljung and Box (1978), tests model adequacy. Box-Jenkins assumes residuals are uncorrelated, random, and normally distributed. If not, the model may be mis-specified, with missing or irrelevant parameters. One example of such a test was the Ljung-Box Test [9]. The Null Hypothesis is residuals were white noise.

$$Q = T(T+2) \cdot \sum_{k=1}^{m} \frac{r_k^2}{T-k} \tag{4}$$

Equation (4) shows the Ljung-Box test where r is the estimated total number of parameters, T is the number of residuals and k is the time lag. In addition, m is the number of time lags to be tested while r_k is the residual autocorrelation at lag k.

Model Comparison

The Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), which are provided in equations (5) and (6), were used to evaluate the accuracy of the model.

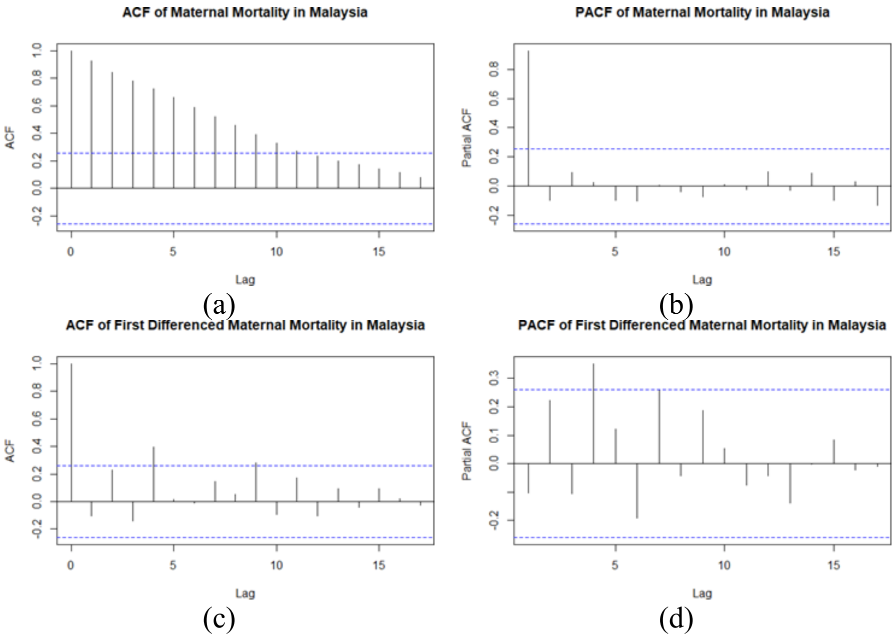
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} e_i^2} \tag{5}$$

$$MAPE = \frac{\sum_{i=1}^{n} \left| e_i / y_i \right|}{n} \cdot 100 \tag{6}$$

In equations (5) and (6) the e_i represents the difference between the actual and fitted value at time period i, while is the length of the evaluation part. The error values on both equations are calculated on the fitted values for a particular forecasting method.

Results and Discussion

The first step in applying the ARIMA model is to check for the stationarity of the data. Figure 2 (a) and (b) reveal that the series is not stationary, where there are decaying patterns on the ACF diagram and the ADF test gives a p-value of 0.8653. Therefore, to proceed with Box-Jenkins modeling, it is necessary to transform the series into a stationary form. After performing first-order differencing, Figure 2 (c) and (d) illustrate that the series becomes stationary, where the decaying pattern disappears and around 1 or 2 autocorrelation values exceed the significance limits. Also, the ADF test gives p-value 0.01. As a result, the series is integrated of order one, which indicates that the data becomes stationary after one differencing. The model obtained represents the general form of ARIMA (p , d , q) with d = 1, where d represents the degree of differencing required to achieve stationarity.



ADF p-value 0.8653

ADF p-value 0.01

Figure 1: ACF and PACF Plots for Maternal Mortality Rates in Malaysia.

Table 1 illustrates the p-value result of the Ljung-Box Test. Based on the results, ARIMA (1,1,1), ARIMA (1,1,2), ARIMA (2,1,1), and ARIMA (2,1,2) yield p-values exceeding the threshold (0.05 value), confirming that their residuals exhibit white noise properties and validating their suitability for forecasting purpose.

Table 1: Summary of Ljung-Box Test.

Model	p-value
ARIMA (0,1,1)	0.0144
ARIMA (1,1,0)	0.0002
ARIMA (1,1,1)	0.0691
ARIMA (1,1,2)	0.0780
ARIMA (2,1,1)	0.0508
ARIMA (2,1,2)	0.1771

The best ARIMA model was assessed with the use of Akaike's Information Criteria (AIC) and Bayesian Information Criterion (BIC) to be considered for suitability in the estimation of the final parameters. Evaluating also included RMSE and MAPE values wherein preference for that model, the lowest in those values will determine the selection for the best model. As shown in Table 2 among the 6 models proposed, ARIMA (1,1,1) is identified as the most suitable model, as it exhibits the lowest BIC, Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), with values of 539.8530, 28.8517, and 0.9883, respectively. Although the AIC from ARIMA (1,1,1) is not the smallest (533.7770), the marginal difference does not outweigh the model's superior forecasting performance, making it the optimal choice.

Table 2: Summary of The Estimated ARIMA Model.

Model	AIC	BIC	RMSE	MAPE
ARIMA (1,1,1)	533.7770	539.8530	28.8517	0.9883
ARIMA (1,1,2)	533.2548	541.3562	29.2452	1.0043
ARIMA (2,1,1)	535.2686	543.3700	29.1788	0.9992
ARIMA (2,1,2)	535.9113	546.0381	29.1788	1.0009

The best model to represent the yearly maternal mortality rate in Malaysia from year 1946 to 2023 is expressed as equation (7). The same equation is used to generate the rate of maternal mortality in 2024, 2025 and 2026 as shown in Table 2.

$$Y_t = Y_{t-1} + \Phi_1 Y_{t-1} - \Phi_1 Y_{t-2} - \theta_1 \varepsilon_{t-1} + \varepsilon_t \qquad (7)$$

Table 3: Summary of Forecast Value of Maternal Mortality Rate in Malaysia.

Year	2024	2025	2026
Forecast	20.2	24.6	21.1

Table 3 indicates a significant increase and decrease in the future maternal mortality rate in Malaysia per 100,000 live births. For the year 2024, there will be a significant decrease in the maternal death rate with only 20.2 mortality rate compared to 25.7 maternal mortality rate cases reported in the previous year, which is the year 2023. And in 2025, however, the trend shifted when the forecasted rate increased to 24.6 compared to the previous year. The forecast value for the maternal mortality rate in Malaysia decreases from 24.6 to 21.1 for the year 2026.

Conclusion

The maternal mortality rates in Malaysia pattern shown a distinct trend, revealing consistent changes influenced by healthcare improvements, policy implementation and socio-economic factors. This analysis highlighted areas where targeted interventions are needed to reduce maternal deaths. Additionally, the study aimed to forecast future maternal mortality rates using the Box-Jenkins Methodology, with ARIMA (1,1,1) identified as the best model based on AIC, BIC, RMSE, and MAPE values. Using this model, maternal mortality rates were forecasted for 2024 to 2026, showing a consistent range of 20 to 24 cases per 100,000 live births. To address the persistent issue of maternal mortality, the study emphasizes the need for educational campaigns to raise awareness, significant investments in healthcare infrastructure, equitable access to quality maternal care, and partnerships with international organizations to adopt comprehensive strategies. These efforts are essential for reducing maternal mortality and improving public health outcomes in Malaysia.

Acknowledgments

I would like to express my deepest gratitude to my supervisor for her guidance and support throughout this study. I am also thankful to my family and friends for their encouragement and understanding during this journey. Lastly, I extend my appreciation to everyone who contributed directly or indirectly to the completion of this project. Thank you all for your invaluable support.

References

[1] Department of Statistics Malaysia. (2024, October). Annual maternal deaths — opendosm. https://open.dosm.gov.my/data-catalogue/deaths maternal.

[2] Alvarez, J. L., Gil, R., Hern´andez, V., & Gil, A. (2009). Factors associated with maternal mortality in sub-Saharan Africa: An ecological study. BMC public health, 9, 1–8.

[3] UNICEF. (2024, May). UNICEF annual report 2023 — UNICEF. https : / / www.unicef.org / reports/unicef-annual-report-2023.

[4] Stanton, M. E., Higgs, E. S., & Koblinsky, M. (2013). Investigating financial incentives for maternal health: An introduction. Journal of Health, Population, and Nutrition, 31(4 Suppl 2), S1.

[5] Bernstein, A. (2021, May). Maternity, maternal health, and the economy during the pandemic. https : / / tcf . org / content / commentary / maternity - maternal – health - economy-pandemic/

[6] Wu, Y., Su, B., & Li, J. (2023). The impact of low fertility rates on labor demand and socioeconomic development in China. China CDC Weekly, 5(27), 599.

[7] Fang, H. (2023). The impact of population aging on financial services and economic development. Available at SSRN 4574199.

[8] Karim, Z. A., Nuruddin, N. A. M., Karim, B. A., Mohamad, M., & Ishak, I. (2023). The impact of population aging and fertility rate on economic growth in Malaysia. Economic Journal of Emerging Markets, 199–211.

[9] Lazim, M. A. (2011). Introductory business forecasting : A practical approach (3rd ed).UiTM Press.