Evaluating Selected Univariate Forecasting Models for Accurate Prediction of Fertility Rates

Olugbenga G. Ojo & Stephen O. Are
Department of Mathematics & Statistics, Federal Polytechnic, Ilaro, Nigeria
E-mail: gabriel.ojo@federalpolyilaro.edu.ng

Abstract

Accurate forecasting of live birth trends is essential for effective planning and policy-making in the health sector. In Nigeria, understanding these trends can support resource allocation and strategic interventions. This study analyses live birth trends from 2005 to 2023 using Holt’s linear trend, Damped Holt’s method, and ARIMA models to identify the most suitable forecasting approach. The results reveal an overall upward trend in live births, with a dip around 2015 followed by a rise after 2020. Model evaluations indicate that ARIMA provides the most reliable forecasts, characterized by unbiased predictions, lower AIC/BIC values, and reduced error variability, making it ideal for long-term projections. Although Holt’s and Damped Holt’s methods capture trends effectively, they exhibit higher forecast error variability and slight biases, limiting their reliability. Comparative forecasts highlight ARIMA's superior performance with lower forecast errors (RMSE: 30.77, MAPE: 8.38%, AIC: 100.91 & BIC:101.51) and stable projections. The findings recommend using ARIMA for future live-birth forecasting, especially for mid- to long-term planning.

Keywords: live-births, model evaluation, Holt’s linear trend, Damped Holt’s method, ARIMA

Introduction

Fertility rate remains a vital demographic indicator, with significant implications for population growth, economic development, and social planning (Bermudez et al., 2022). Human fertility, however, is a complex and dynamic process, shaped by a range of biological, socio-economic, and cultural factors such as education, urbanization, and changes in social norms around family size and gender roles (Olatoregun et al., 2014; Maulida et al., 2023). Forecasting fertility is a challenge due to this complexity. It requires models that can account for shifts in behavior, economic conditions, and policy interventions. In many low- and middle-income countries, high fertility persists due to preferences for large families and limited access to reproductive health services (Olatoregun et al., 2014). Fertility rates often follow a natural pattern beginning in early adulthood, peaking in the twenties, and gradually declining in the forties (Maulida et al., 2023). Fertility is one of the three primary drivers of population dynamics, alongside mortality and migration. Accurate fertility forecasts are crucial for planning in health, education, and labor markets. Yet, the task is difficult due to the variability and unpredictability of reproductive behavior. This has led to the adoption of univariate forecasting methods that rely solely on historical fertility data to project future trends. The central problem is identifying which univariate time series model offers the most reliable and accurate forecasts of fertility rates. Numerous methods exist each with specific strengths and limitations depending on the characteristics of the data, such as trend, seasonality, and randomness (Hyndman & Athanasopoulos, 2018). Lam and Wang (2021) explored demographic forecasting using Gaussian Process Regression. Their model used a spline-based mean function and spectral mixture kernel to track age-specific rates. Results showed improved forecast accuracy for mortality and fertility in developed nations. The approach outperformed traditional time-based demographic models. Tzitiridou-Chatzopoulou et al. (2024) investigated the use of adaptive machine learning to forecast monthly birth rates in Scotland. Their study combined non-linear ML techniques with traditional models in a one-step-ahead prediction exercise. The results highlighted the accuracy of machine learning in fertility forecasting. The model aids policy planning and healthcare decisions. It also supports early identification of pregnancy risks and treatment planning. Kim and Shin (2023) explored fertility prediction using individual-level micro data from Korea’s Census 2% sample. Their model emphasized completed fertility rates for improved cohort-based forecasting. Results showed younger cohorts delay childbirth more than earlier ones.

Education and childbirth history were key predictors. The study offers a refined lens for fertility trend analysis. Karunanidhi and Sasikala (2023) analyzed birth rate trends in Tamil Nadu using ARIMA models. They applied the Box-Jenkins methodology to identify the most suitable model for forecasting. Their results pointed to a strong model fit and predicted a decline in future birth rates. The study showcased the reliability of ARIMA models in demographic forecasting. Their findings emphasize the importance of data-driven approaches in shaping effective public welfare policies.

Adekanmbi and Akinyemi (2017) compared the forecasting power of ARIMA and ARIMAX models for predicting live births in Nigeria. They found that both models performed well, but ARIMAX had a slight edge in accuracy. The ARIMAX model included the population of women of childbearing age as an exogenous variable. This improved its ability to explain variations in birth trends. Their study supports the use of exogenous factors to enhance demographic forecasts.

Parmar and Patel (2024) used the Box-Jenkins method to forecast India’s birth rate through ARIMA modeling. Their analysis identified the best-fitting model based on AIC and BIC values. Findings indicated a consistent decline in birth rate over time, highlighting a negative relationship between birth rate and year. The study demonstrates the usefulness of ARIMA models in long-term demographic forecasting and policy planning.

Wang (2023) conducted a time series analysis of China’s declining birth rate. The study used historical fertility data to examine recent trends. ARIMA and ETS models were applied to forecast short-term changes. ARIMA was found to provide better predictive performance. Results revealed a steady annual decline in the birth rate. The study highlights a looming fertility crisis without policy action. Since different models can yield significantly different results, a comparative analysis is essential to determine the most suitable forecasting technique. This study aim to evaluate and compare the performance of univariate forecasting models in predicting fertility rates, with a focus on addressing the variability in forecasting accuracy across methods.

Materials and Methods

This study utilizes secondary data on annual live births (male and female) from 2005 to 2023, collected from the State Hospital in Ilaro, Ogun State, Nigeria. It applies Univariate Time Series Analysis to model and forecast fertility rates over the period using three methods: Auto-Regressive Integrated Moving Average (ARIMA), Single Exponential Smoothing (SES), and Holt’s Damped Trend Method.

Model Specification and Estimation

ARIMA Model

The ARIMA model, developed by Box and Jenkins, is a widely used method for time series forecasting that combines Auto-Regressive (AR), Integrated (I), and Moving Average (MA) components. Its general form includes lagged values of the series and past error terms:

$$Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t. \qquad (1)$$

Where:
Y_t is the actual value of the series at time t, c is a constant (intercept term), $\phi_1, \phi_2, \ldots, \phi_p$ are the parameters of the AR (AutoRegressive) part, capturing the influence of past values, $\theta_1, \theta_2, \ldots, \theta_q$ are the parameters of the MA (Moving Average) part, reflecting the impact of past forecast errors, ε_t is the random error (white noise) at time t, p is the number of AR terms, q is the number of MA terms, and d (from the "Integrated" part) is the number of times the data has been differenced to achieve stationarity. The modeling process involves three key stages: identification, where checking for stationarity and selecting model order is predominant, estimation, that is; using least squares or maximum likelihood, and diagnostic checking (residual analysis). ARIMA assumes the series is stationary (achieved through differencing if needed), has a linear structure based on past values, and that errors are normally distributed.

Single Exponential Smoothing (SES)

Single Exponential Smoothing (SES) forecasts data without trend or seasonality by applying exponentially decreasing weights using the formula:

$$S_t = \alpha Y_t + (1-\alpha)S_{t-1}. \quad (2)$$

Holt's Linear Exponential Smoothing

An extension of SES for data with trends, Holt's model involves:

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + b_{t-1}) \quad (3)$$

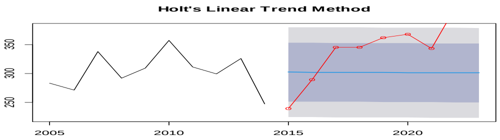
$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1} \quad (4)$$

$$\hat{Y}_{t+m} = L_t + m b_t \quad (5)$$

Where:

Equation (3) deduced the level equation, Equation (4) deduced the trend equation, while equation (5) is the forecast equation. L_t: Estimated level (smoothed value) of the series at time t, Y_t: Actual observed value at time t, L_{t-1}: Estimated level at time $t-1$, b_{t-1}: Estimated trend (slope) at time $t-1$, α: Smoothing parameter for the level ($0 < \alpha < 1$), b_t: Updated estimate of the trend at time t, $L_t - L_{t-1}$: Change in level between periods, β: Smoothing parameter for the trend ($0 < \beta < 1$), $\hat{Y}_{t+m}$: Forecasted value, m: Number of steps ahead to forecast, L_t: Last estimated level, and b_t: Last estimated trend (slope).

Results and Discussion



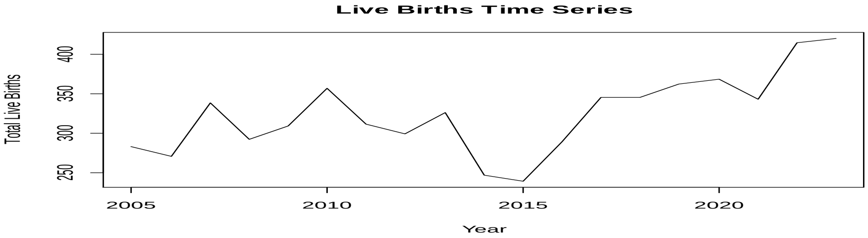


Figure 1: Live-Births Graphical Plot

Figure1 deduced an overall upward trend in live births from 2005 to 2021, with fluctuations and a notable dip around 2015, followed by a sharp increase after 2020.

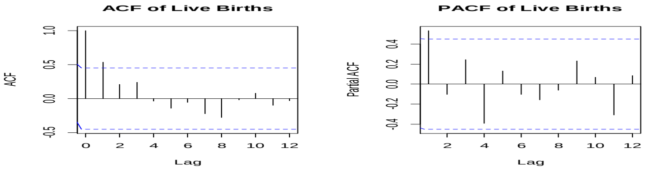


Figure 2: Autocorrelation and Partial Autocorrelation of Live Births

Figure 2 shows the ACF and PACF plots which indicate a significant correlation at lag 1, with values decreasing for higher lags. Most subsequent lags fall within the confidence intervals, suggesting a likely AR (1) process in the live births time series data.

Table 1: KPSS Stationarity Test

Statistic	Value
KPSS Level	0.4262
Truncation Lag Parameter	2
p-value	0.0659

Table 1 indicates that the p-value (0.06588) is slightly higher than the standard 0.05 significance level. As a result, we fail to reject the null hypothesis of stationarity at the 5% level, suggesting that the live birth time series data is likely stationary around a constant level.

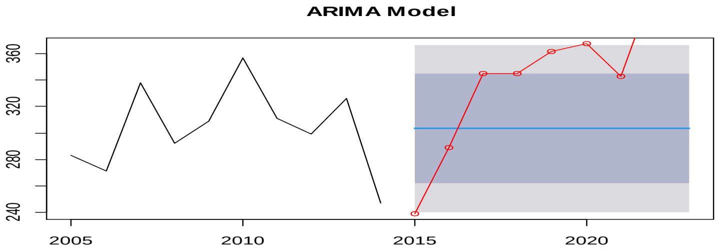


Figure 3: Live Births Forecast Using Holt's Linear Trend Method

Figure 3, displays Holt's linear trend method applied to live birth data. The observed values (black) show variability from 2005 to around 2015, followed by a forecasted increase (red) beyond 2015. The confidence intervals (blue shaded region) widen as the forecast progresses, indicating increased uncertainty. The forecast suggests continued growth in live births, although the growing width of the confidence intervals reflects diminishing accuracy over time. The model effectively captures the observed trend but shows less confidence in future predictions. Further, the Holt's method forecast suggests a stable trend with a slight decline over time, starting with an initial level of 304.29 and a minor negative trend of -0.18. The small smoothing parameters (alpha and beta = 0.0001) indicate that the model adapts slowly to new data. The forecast error, as reflected by the sigma (39.72), suggests moderate variability, with error measures such as RMSE (30.77) and MAPE (8.41%) indicating reasonable accuracy. The model shows no significant bias (ME = -0.0077) or residual autocorrelation (ACF1 = -0.1667), performing well compared to a naive forecast (MASE = 0.6354). However, the confidence intervals widen as the forecast extends into the future, highlighting growing uncertainty and the need for caution in long-term predictions beyond 2023.

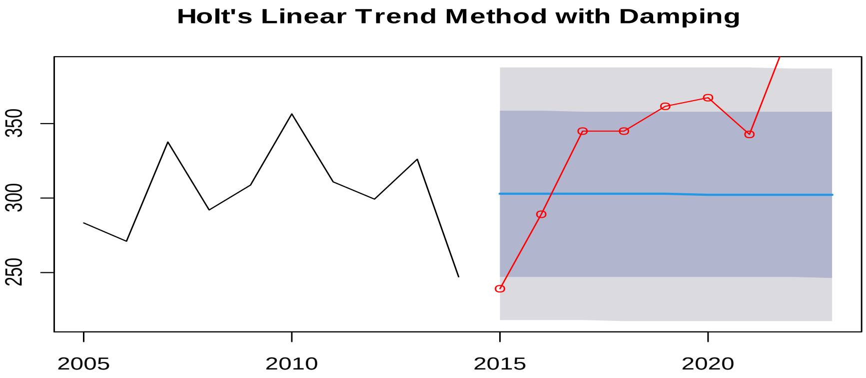


Figure 4: Holt's Linear Trend Method with Damping

Figure 4, illustrates Holt's linear trend method with damping applied to the live births data. The observed points (black) show fluctuations over time, while the forecasted values (red) suggest a moderate increasing trend. The blue shaded areas represent confidence intervals that expand as the forecast progresses, indicating higher uncertainty for future predictions. The model effectively captures the upward trend seen in recent years but suggests a slowing or damping effect on future growth. So, the Damped Holt’s method forecast reveals a stable trend, beginning with an initial level of 304.64 and a slight negative trend of -0.18, supplemented by a damping factor (phi = 0.9615) that indicates a gradual decrease in the trend's influence over time. The small smoothing parameters (alpha and beta = 0.0001) suggest slow adaptation to changes in the data, while the sigma value (43.54) reflects a moderate level of forecast error variability. The model's performance is assessed with error measures showing a mean error (ME) of -0.50, root mean square error (RMSE) of 30.79, and mean absolute percentage error (MAPE) of 8.42%, indicating reasonable accuracy. The forecasts for the years 2015 to 2023 maintain a slight decline, with confidence intervals widening over time, underscoring increasing uncertainty in future predictions.

Figure 5: ARIMA (Auto-Regressive Integrated Moving Average)

The figure 5 demonstrates an ARIMA model applied to a time series, showing historical data with irregular fluctuations and forecasted values. The forecast indicates an initial sharp rise followed by stabilization, with increasing uncertainty over time. The results of the ARIMA model indicate that the fitted model is an ARIMA (0,0,0) with a non-zero mean, suggesting that the data is essentially modeled as a constant average of 303.30, with a standard error of 9.73. The log likelihood of the model is -48.45, and the model selection criteria show an AIC of 100.91, an AICc of 102.62, and a BIC of 101.51, which are used to evaluate the model's fit. Training set error measures reveal a mean error (ME) very close to zero (-2.28e-10), indicating an unbiased forecast, while the root mean square error (RMSE) is 30.77 and the mean absolute error (MAE) is 24.90, suggesting reasonable accuracy. The MAPE value of 8.38% signifies that the average forecast error is about 8.38% of the observed values.

Table 2: **Model Comparison Using AIC, BIC, RMSE, and MAPE**

Model	AIC	BIC	RMSE	MAPE
Holt’s	101.56	103.07	30.77	8.4
Damped_Holt’s	103.57	105.38	30.79	8.42
ARIMA	100.91	101.51	30.77	8.38

Among the three models compared in table 2, ARIMA shows the best overall performance with the lowest AIC (100.91) and BIC (101.51), reflecting a superior fit while penalizing model complexity. Its RMSE (~30.77) is nearly identical to that of Holt’s method, which also reports an AIC of 101.56 and BIC of 103.07, making Holt’s slightly less optimal. Damped Holt’s method performs the worst, with the highest AIC (103.57), BIC (105.38), and a marginally higher RMSE (30.79), indicating lower predictive accuracy. Regarding MAPE, ARIMA again performs best at 8.38%, narrowly beating Holt’s (8.41%) and Damped Holt’s (8.42%). Overall, ARIMA strikes the best balance between goodness-of-fit and forecast accuracy, while Damped Holt’s adds complexity with no significant benefit.

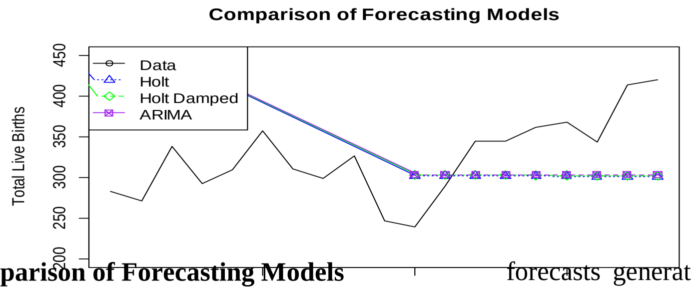


Figure 6: Comparison of Forecasting Models

Figure 6, illustrates the total live births from 2005 to 2023, comparing the actual historical data with forecasts generated by three distinct models: Holt's method, Damped Holt's method, and the ARIMA model. The observed data is represented by a solid black line, while forecasts from Holt's and Damped Holt's methods are depicted with triangle and diamond markers, respectively; ARIMA forecasts are illustrated with purple squares. Although these forecasts indicate a slight upward trend towards the series' conclusion, they also demonstrate significant variability and divergence among the models, particularly noting the relative stability of the ARIMA model. Considering all the factors, the ARIMA model emerges as the best forecasting method in this analysis due to its unbiased forecasts, lower AIC/BIC values, and lower variability in forecast errors. It provides a stable and reliable forecasting option, especially in a case where maintaining a constant average is critical.

Conclusion

This comparative analysis of Holt’s linear trend, Damped Holt’s method, and ARIMA models for forecasting live births identifies ARIMA as the most reliable option, offering stable forecasts with unbiased predictions, lower AIC/BIC values, and reduced error variability. ARIMA’s strength lies in its ability to capture consistent patterns, making it well-suited for long-term projections. Although Holt’s and Damped Holt’s methods effectively modeled the trends, their higher error variability and slight biases limit their reliability. Given the uncertainty in forecasting, ARIMA is recommended, particularly for mid- to long-term planning. To maintain accuracy, the model should be updated regularly with new data. Continuous monitoring and periodic re-assessment will further enhance forecast precision and support effective resource planning. An adaptive approach will ensure that the forecasting framework evolves with emerging trends and fluctuations, improving the overall reliability of predictions and enabling better decision-making for live birth projections.

References

Adekanmbi, V. T., & Akinyemi, J. O. (2017). *Modeling and forecasting live births in Nigeria: ARIMA vs. ARIMAX models*. African Population Studies, 31(1), 3628–3640.

Bermudez, L. G., Lutalo, T., Nalugoda, F., & Gray, R. H. (2022). *Understanding fertility transitions in low- and middle-income countries: Trends and determinants*. Population Research and Policy Review, 41(4), 1505–1524.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts.

Karunanidhi, G., & Sasikala, R. (2023). *Application of ARIMA model to forecast the birth rate in Tamil Nadu*. Journal of Statistical and Econometric Methods, 12(1), 25–35.

Kim, S., & Shin, Y. (2023). *Forecasting cohort-based fertility trends in Korea using micro-level census data*. Journal of Population Economics, 36(1), 101–128.

Lam, T., & Wang, H. (2021). *Demographic forecasting with Gaussian process regression*. Demographic Research, 45, 1155–1190.

Maulida, U., Kurnia, A., & Rahmah, N. (2023). *Determinants of fertility rate in Indonesia: A spatial regression approach*. Journal of Population Sciences, 14(2), 45–61.

Olatoregun, H., Akinyemi, J. O., & Odimegwu, C. (2014). *Fertility transition in Nigeria: Exploring the role of desired family size and education*. African Population Studies, 28(2), 1095–1106.

Parmar, R., & Patel, H. (2024). *ARIMA modeling for forecasting birth rates in India: A time series approach*. International Journal of Statistics and Applied Mathematics, 9(1), 17–23.

Tzitiridou-Chatzopoulou, M., Georgiadis, T., & Kolovos, A. (2024). *Fertility forecasting using adaptive machine learning: A case study of Scotland’s monthly birth rates*. Applied Artificial Intelligence,38(1),1–18.

Wang, J. (2023). *Forecasting China's declining fertility: A time series approach*. Journal of Population and Social Studies, 31(2), 145–160