# Comparative Analysis of Holt's model with ARIMA model for the future prediction of Birth Rate in India

**Minakshi Mishra[1], Anuj Kumar*[2] and Ratan Kumar Thakur[3]**
Department of Statistics, School of Physical and Decision Sciences,
Babasaheb Bhimrao Ambedkar University (A Central University),
Lucknow-226025, INDIA

***Abstract:***

The present study aimed at the comparison of ARIMA model with Holt's model and ARIMA model. In this study, time series data concerning about birth rate was collected for the period from 1995-96 to 2018-19 (24 years). Two separate models were used to find the best fitting model for the future forecast of India's birth rate: Holt's and Autoregressive integrated moving average (ARIMA). In terms of low accuracy measure error, Holt's models and the ARIMA model are found to be more accurate than the other analysed models. Forecasting errors namely Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Mean Percentage Error (MPE). The ARIMA model performance and is more accurate than the Holts model, which has the lowest MSE = 0.012227, RMSE= 0.110577, MAE = 0.087999, MPE= 0.088056, and MAPE=0.366742.

***Keywords**:* Birth Rate, Holt's, ARIMA.

## 1.Introduction

Fertility is one of a population's most important demographic characteristics (Bhende *et al*., 1997; Pathak *et al*., 2016; Singh *et al*., 2022). The level of childbearing in a population is measured by fertility rates. They are crucial in defining a population's growth rate as well as its age structure (the proportion of the population who is young against the proportion who is old) (Bhende *et al*., 1997; Tiwari *et al*., 2021). Some fertility measurements are true "rates," in the sense that they represent the "risk" of having a child during a specific time frame. The crude birth rate, the general fertility rate, and age-specific fertility rates are examples of such measurements. Other metrics of fertility, such as cohort completed fertility and the total fertility rate, are not true rates but do capture the level of childbearing in a community (Bhende *et al*., 1997; Bongaarts, 2020). Fertility prediction plays a vital role due to its relevance in predicting and controlling population growth and dynamics (Pathak *et al*., 2016; Radkar, 2020).

This study investigates various causes of declining birth rates. We start by looking at birth rates by age, education, race and ethnicity, marital status, and birth parity for various demographic groups (Kearney *et al*., 2022). We wanted to see how a gonadotropin releasing hormone agonist trigger, alone or in combination with several low dosage human chorionic gonadotropin regimens, affected rates of oocyte maturation and cumulative live birth in high responders who were on gonadotropin releasing hormone antagonist cycles (He *et al.,* 2022). We characterise monthly variations in preterm birth by method of delivery adjusted for seasonality and trend using data from the US Census of Births from 2010 to2020 (Dench *et al*., 2022). The researchers wanted to examine how well five different intrauterine adhesion evaluation techniques predicted the live birth rate following transcervical adhesion (Cao *et al.,* 2021). The birth rate is a crude measure of a population's fertility and is an important factor in population expansion. It calculates the number of live births per thousand population in a particular year and location.

Many researchers (Shang *et al*., 2020) have considered model averaging techniques, (Singh *et al.,* 2022) considered Bongaarts' proximate determinants model, (Fagbamigbe *et al*., 2014; Devkota, 2018, Arbu *et al*., 2021, Singh *et al*., 2022) considered multiple regression techniques, (Elhag *et al*., 2020) considered exponential smoothing, (Adeyinka *et al*., 2020) used Holt's winter technique, (Tripathi *et al*., 2018; Adeyinka *et al*., 2020, Elhag *et al*., 2020). It presents a basic econometric model for COVID-2019 spread prediction. We utilised an ARIMA model predicted from Johns Hopkins epidemiological data to estimate the epidemiological trajectory of COVID-2019 prevalence and incidence (Benvenuto *et al.* 2020). The ARIMA model, which incorporates demographic characteristics, social economic circumstances, transportation accessibility, and environmental resources into four regression models, is utilised in this work to forecast other time series models. In mainland Spain and California, they employed the ARIMA technique to establish a strategy for estimating next-day electricity costs for both spot and long-term contracts (Contreras *et al.* 2003).

The number of requests for automobiles from PT. Suzuki Indomobil Motor from 2017 to 2019 is data with seasonal trends that may be smoothed with ARIMA and Holt-Winters exponential smoothing (Pujiharta *et al*., 2022). Rainfall is forecasted using the ARIMA and Holt Winters models, with the ARIMA and Holt Winters models providing the best accuracy (Joshi and Tyagi 2021). The-Sutte Indicator approach is more accurate than ARIMA (2,2,2) and Holt Winters in predicting data of Infant Mortality Rate in China, according to the findings and discussion. The-Sutte Indicator method's MSE and MAPE values are less than ARIMA (2,2,2) and Holt Winters, as can be observed from the MSE and MAPE values (Kurniasih et al., 2018)

## 2.Materials and Methods

The current investigation was based on secondary data (time series data) on India's birth rate, which was acquired from www.indiastat.com. during the years 1995-96 to 2018-19 (24 years). The data analysis has been performed using the R software.

### 2.1 Double Exponential Smoothing (Holt's) Model

Holt improved simple exponential smoothing to allow data with a trend to be forecasted (Fig.1) and it is also known as Double exponential smoothing method (Holt, 2004, Nazim *et al*., 2014). Holt's technique uses two smoothing parameters to update the level and trend component and have three equations which perform different function: (1) level smoothing equation ($l_t$) (2) trend smoothing equation ($b_t$) and (3) forecasting equation ($h_{t+m}$). The mathematical equation for Holt's method is given as follows:

Level smoothing equation: $l_t = \alpha\, y_t + (1-\alpha)\,(l_{t-1} + b_{t-1}$ (1)

Trend smoothing equation: $b_t = \beta\,(l_t - l_{t-1}) + (1-\beta)\, b_{t-1}$ (2)

Forecasting equation : $h_{t+m} = l_t + m\, b_t$ (3)

$\alpha$=smoothing constant for the data ($0<\alpha<1$)

$\beta$= smoothing constant for the trend estimate ($0<\beta<1$)

m=number of periods ahead to be forecast

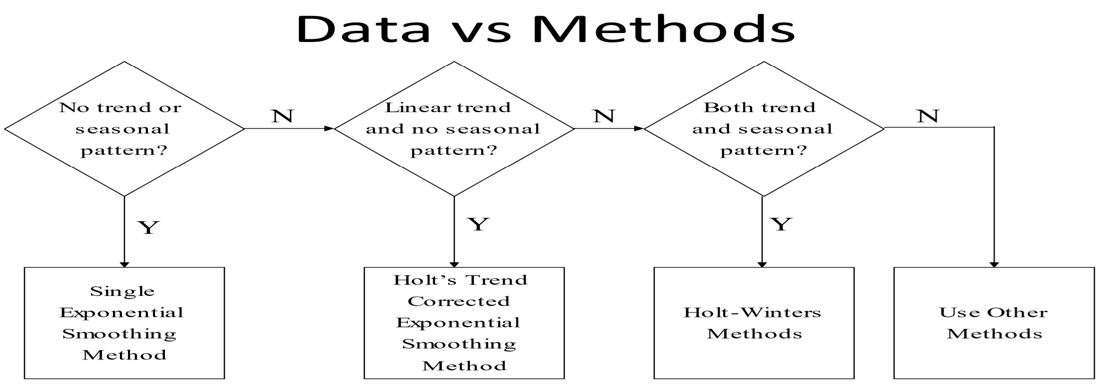



**Fig. 1** Selection Procedure for Double Exponential Smoothing

### 2.3 Autoregressive Integrated Moving Average (ARIMA)

The Autoregressive Integrated Moving Average (ARIMA) model was created by Box and Jenkins and is widely used in time series forecasting (Tripathi *et al*., 2018). The fundamental goal of fitting this ARIMA model is to identify the time series' stochastic process and properly predict future values (Asteriou *et al*., 2011; Liu *et al*, 2021). ARIMA can be used in situations when there is a signal of non-stationarity in the data. The architecture of Box-Jenkins Methodology is shown in (Fig. 2),

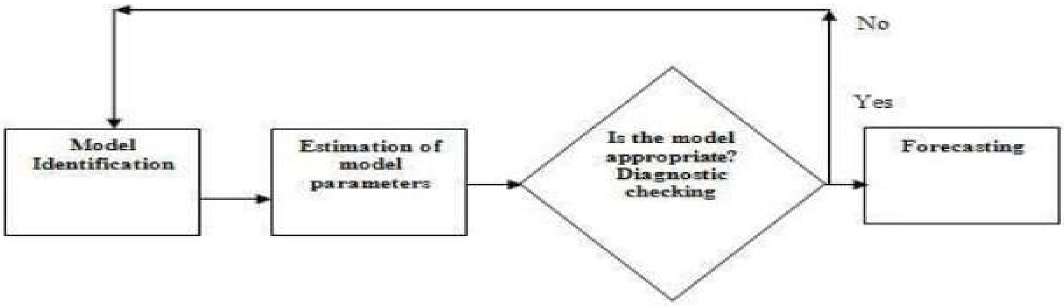



**Fig. 2** Architecture of Box-Jenkins Methodology

An Autoregressive Integrated Moving Average (ARIMA) is used to describe a time series $X_t$ of order (p, d, q) if d[th] difference of $X_t$ i.e., $Y_t = \nabla^d X_t$ is a stationary ARMA (p, q) process. As $Y_t \sim$ ARMA (p, q), it has the following form:

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots\dots\dots + \Phi_p Y_{t-p} + \epsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots\dots\dots + \theta_q \varepsilon_{t-q} \dots\dots (5)$$

Where $\Phi_p \neq 0$ and $\Phi_q \neq 0$, $\epsilon_t \sim$ white noise $(0, \sigma^2)$, Cov $(\epsilon_t , Y_{t-1}) = 0 \; \forall \; j > 0$.

The ARIMA model is a composite of several models rather than a single model (Elhag *et al*., 2020). As a result, the ARIMA equation is stated in the form of the orders of these distinct models designated by the letters p, d, and q and it is explained in (Fig.3).

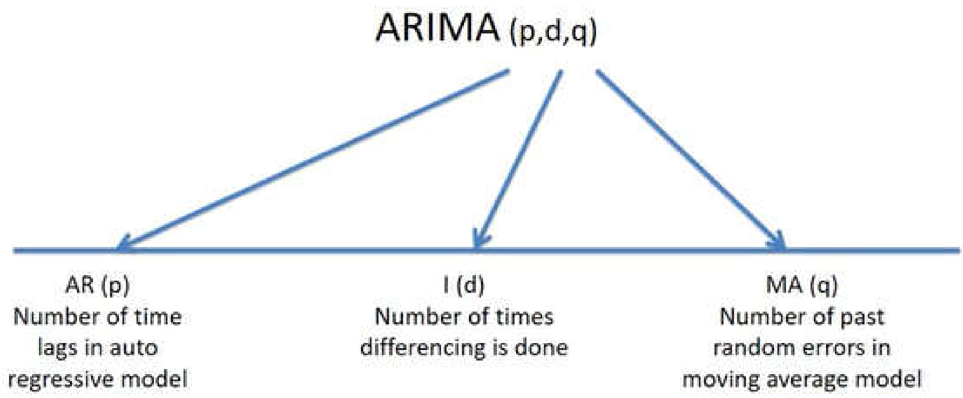



**Fig. 3** Architecture of ARIMA Model

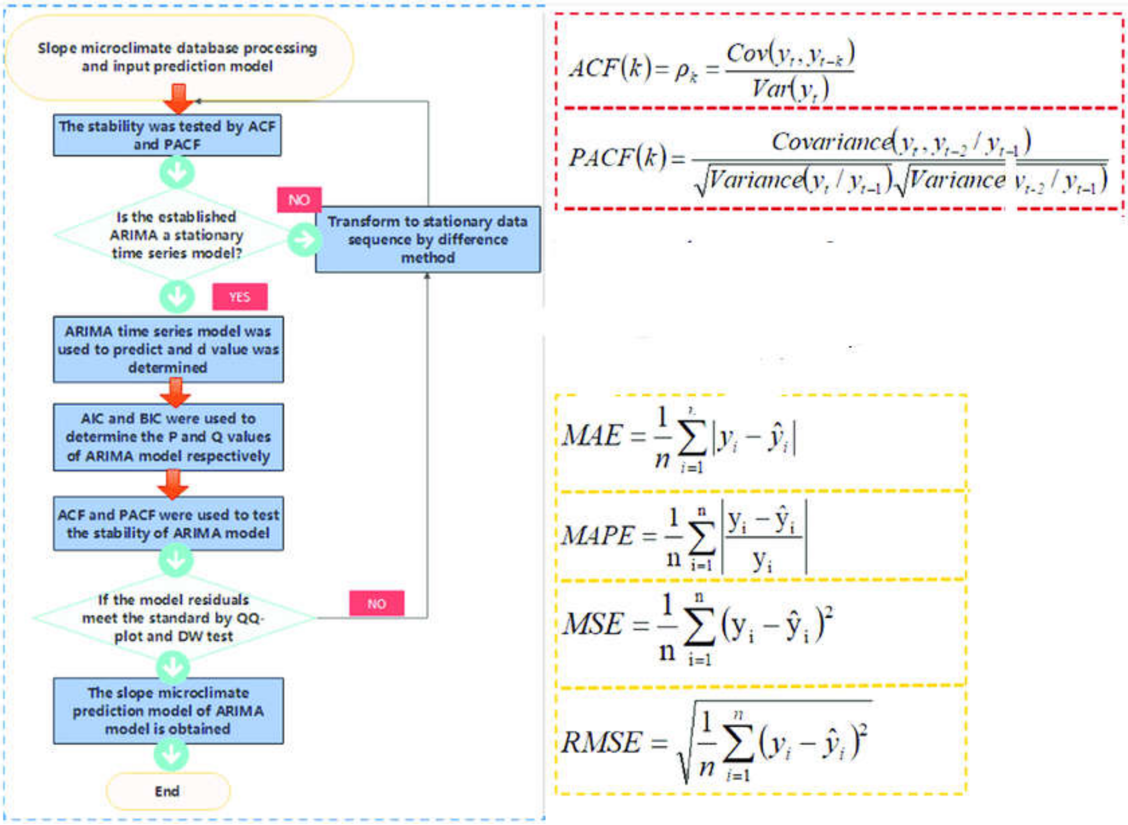


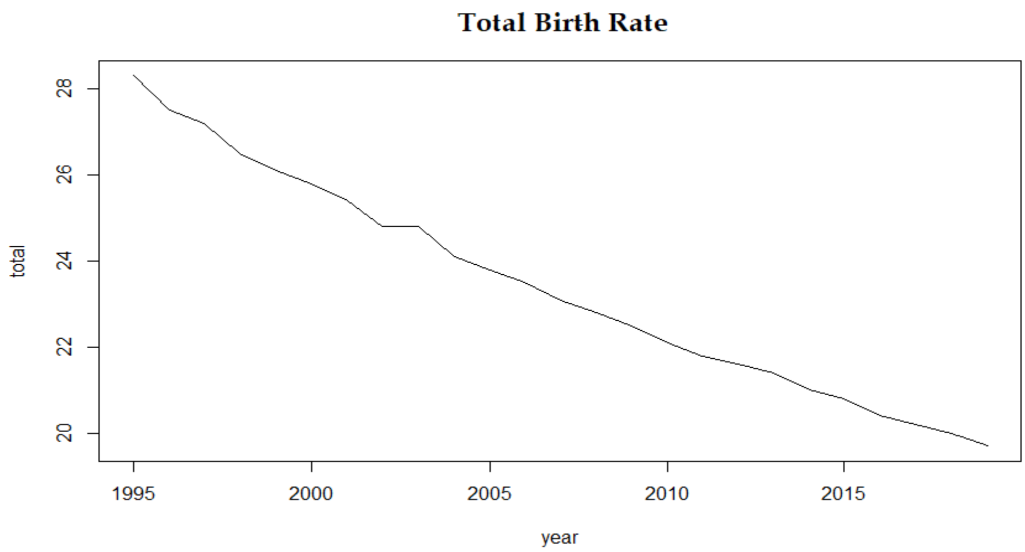

**Fig. 4** Process of ARIMA analysis and accuracy measurements

## 3.Results and Discussion

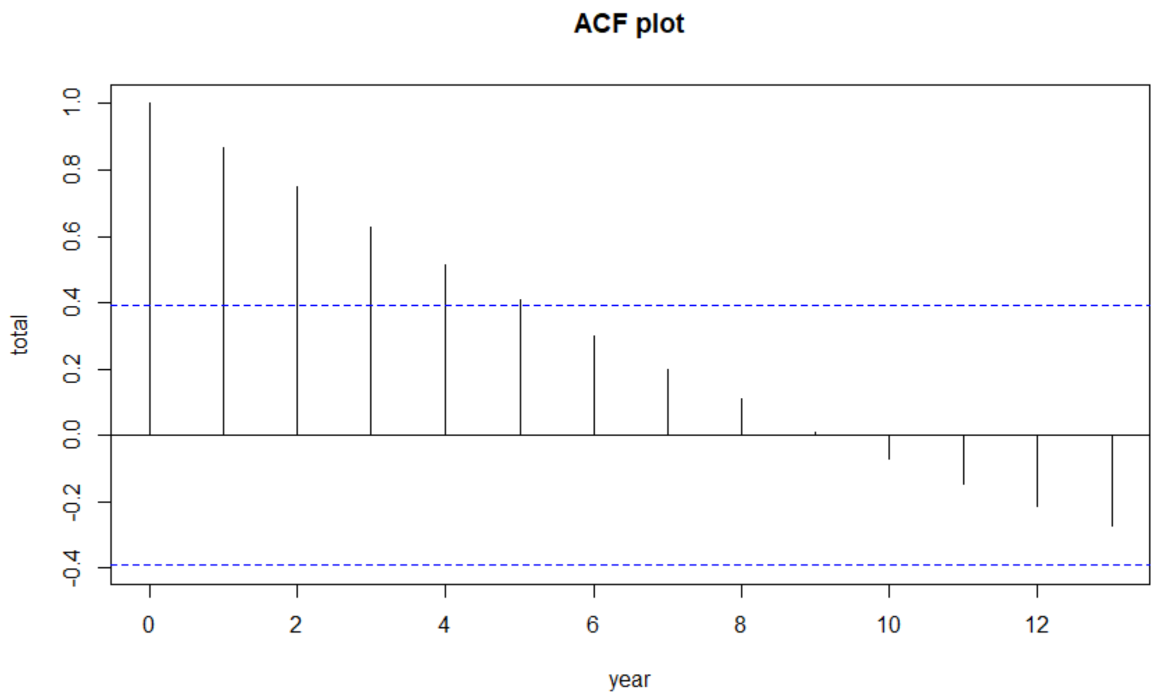
In this section, a comparative of the prediction accuracy of the ARIMA model to Holt's model for future prediction of India's birth rate and it is shown in Table 1. The outcome revealed that ARIMA Model is significantly better than the other considered model i.e., Holt's model and it is observed by the low values of accuracy measures namely mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), mean percentage error (MPE), mean absolute percentage error (MAPE).

**Table 1 Comparison of Various Model on the basis of Model Evaluation Criteria**

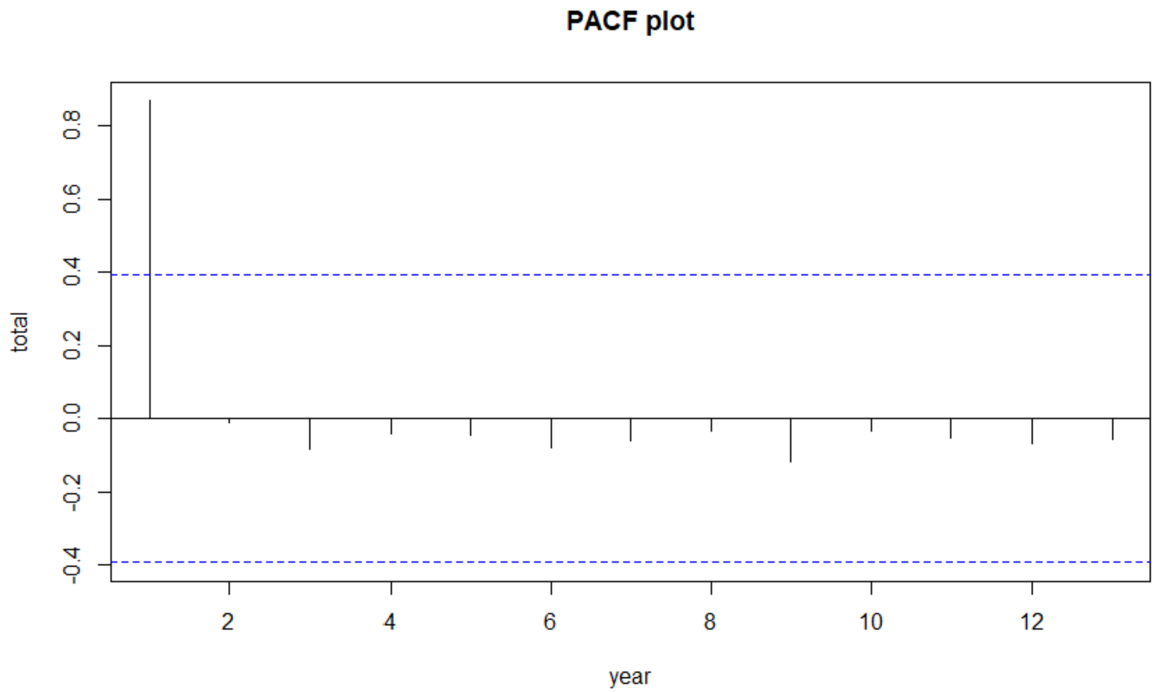
| Model Evaluation Criteria | ARIMA | HOLT's |
|---|---|---|
| MSE | 0.012227 | 0.018578 |
| RMSE | 0.110577 | 0.136302 |
| MAE | 0.087999 | 0.094424 |
| MPE | 0.088056 | 0.212323 |
| MAPE | 0.366742 | 0.389555 |



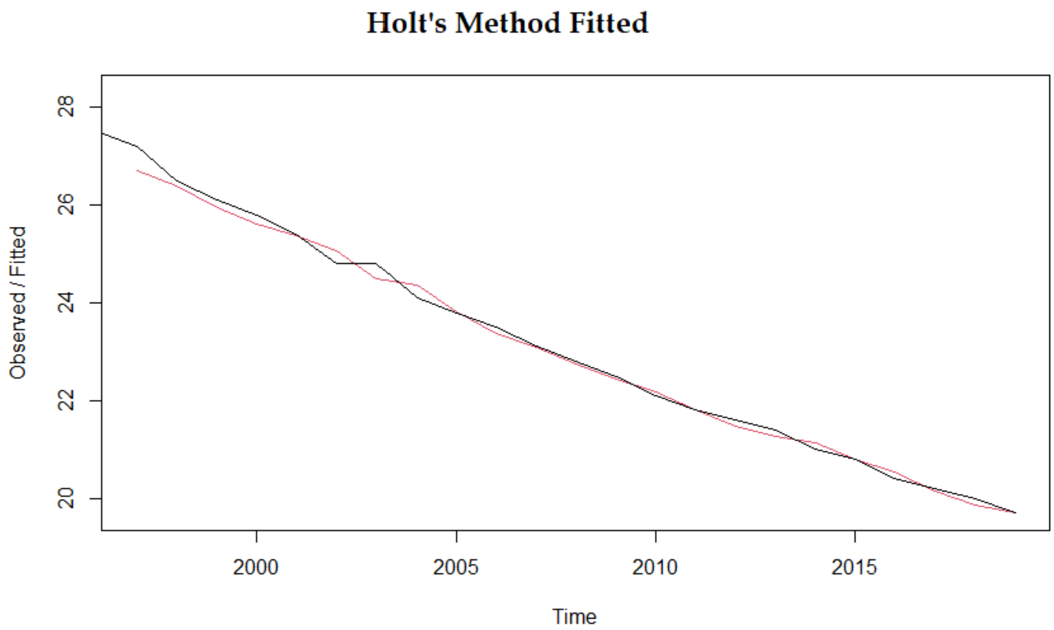

The above **Fig. 5** line plot shows total actual birth rate data in India from 1995 to 2019.



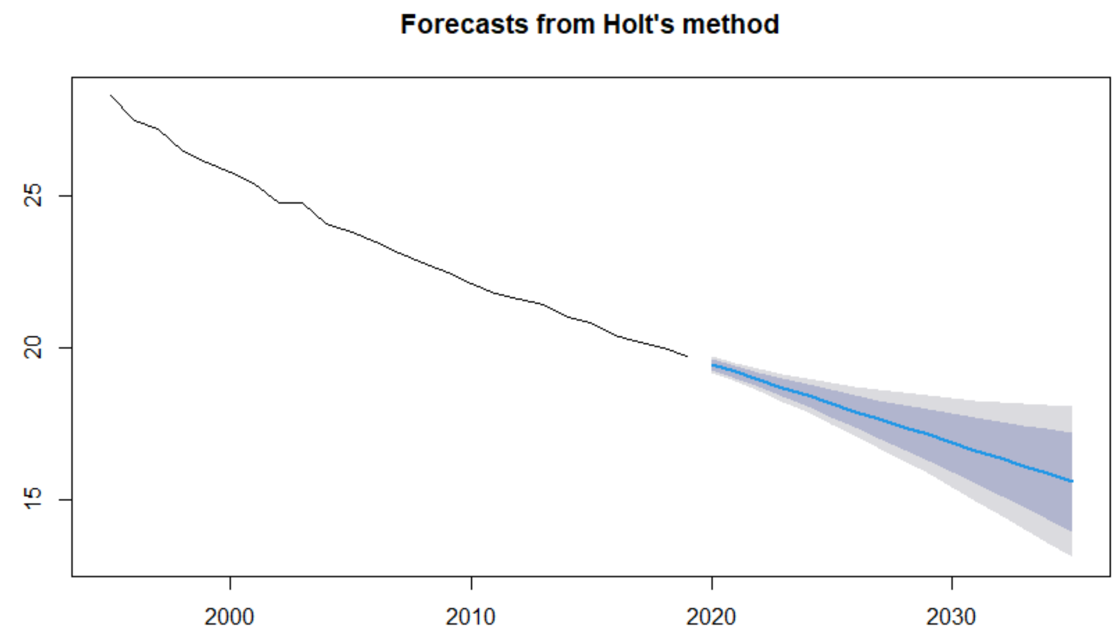

The above **Fig. 6** The autocorrelation function was used to actual data plot in India from 1995 to 2019.



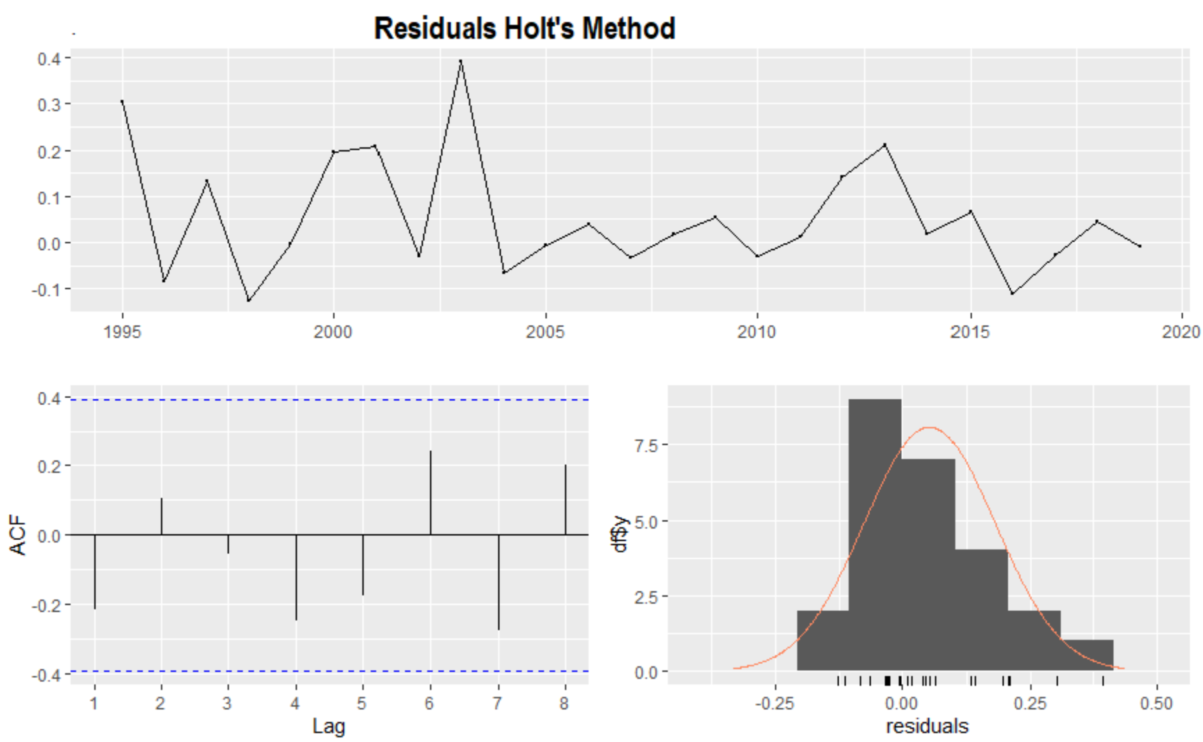

The above **Fig. 7** Actual data from 1995 to 2019 in India was shown using the parital autocorrelation function.



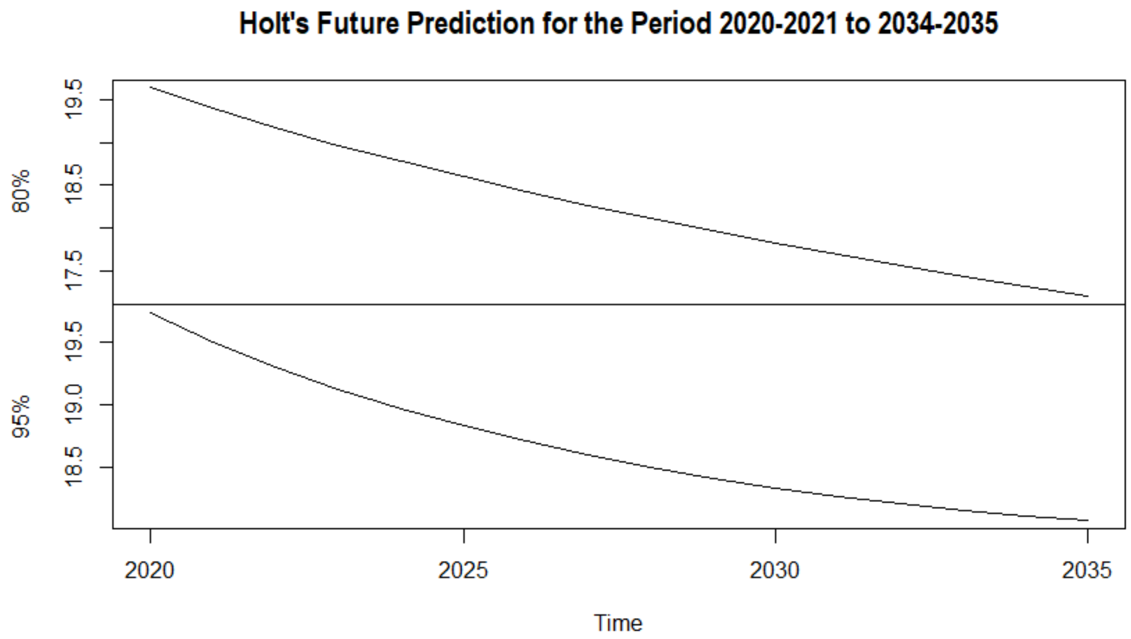

The above **Fig. 8** Holt's method actual and fitted plot are used in India (1995-2019).



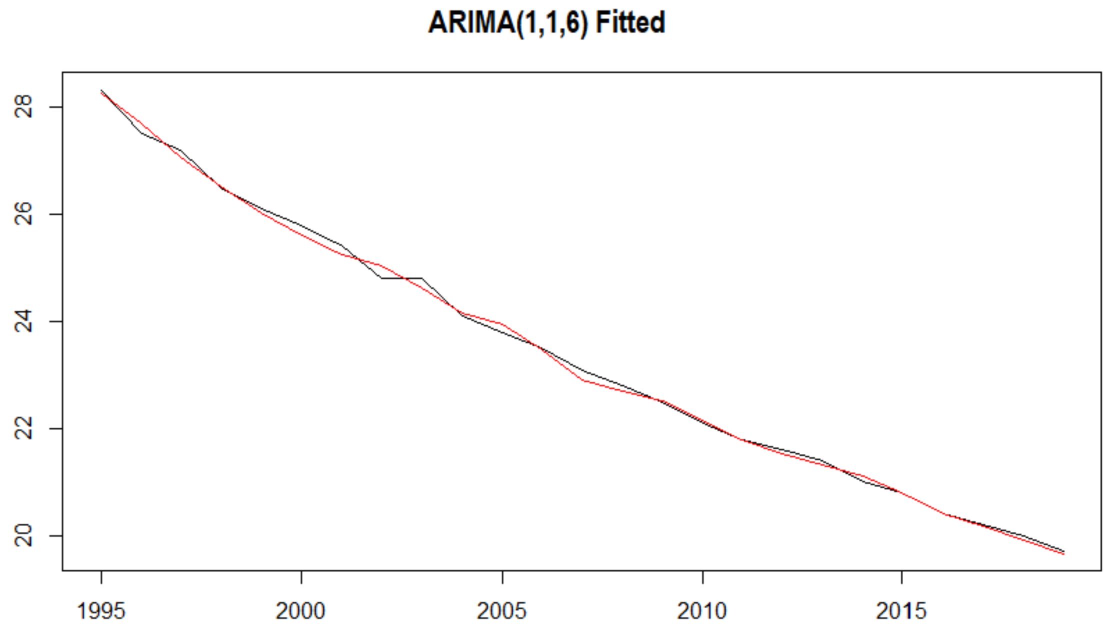

The above **Fig. 9** birth rate fit and forecast plots using Holt's method in India (1995-2035).



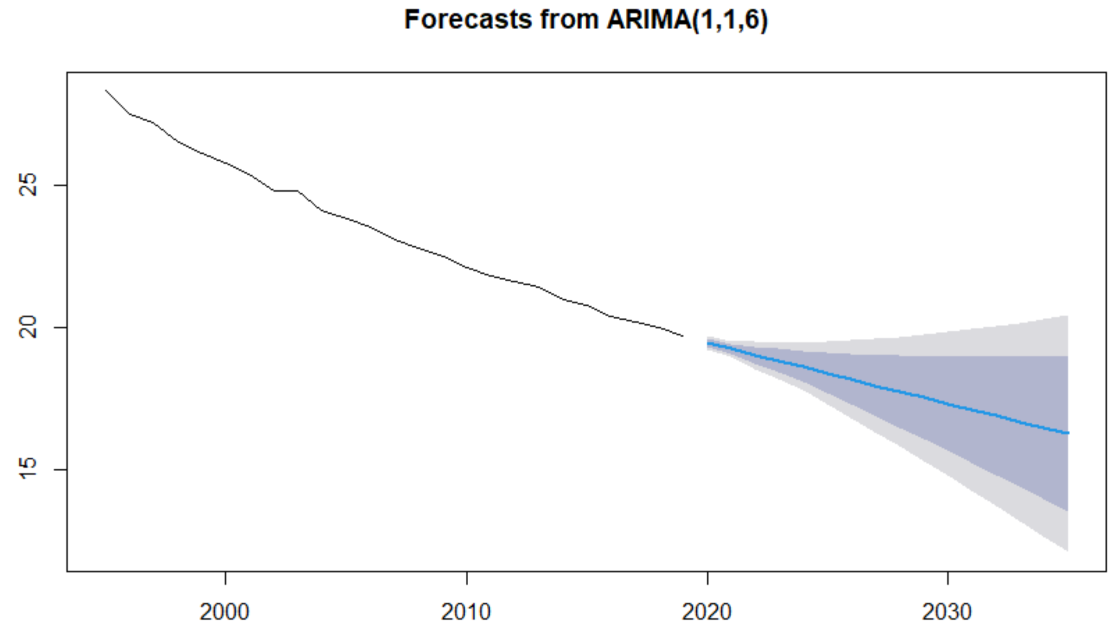

The above **Fig. 10** Holt's method Residuals plots.



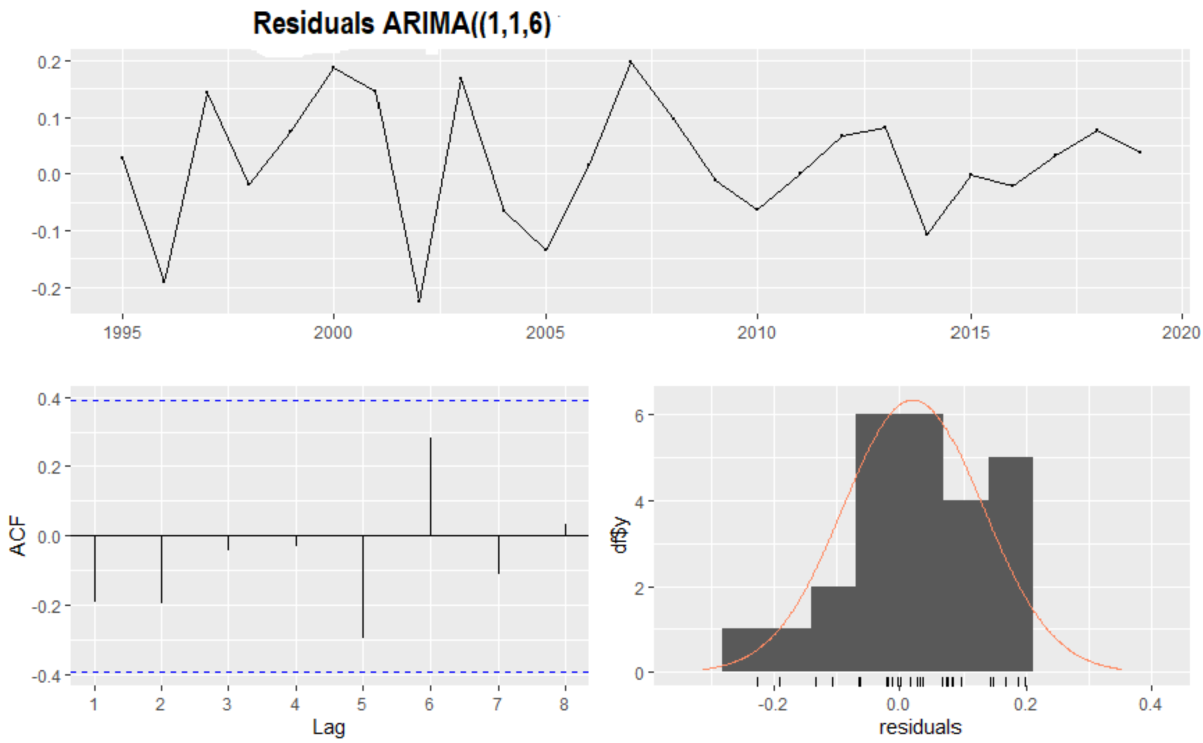

The above **Fig. 11** Holt's method during the 2020-2021 to 2034-2035 period, the compersion upper future prediction is 80% and 95%. The best future forecast has a 95% higher confidence level.



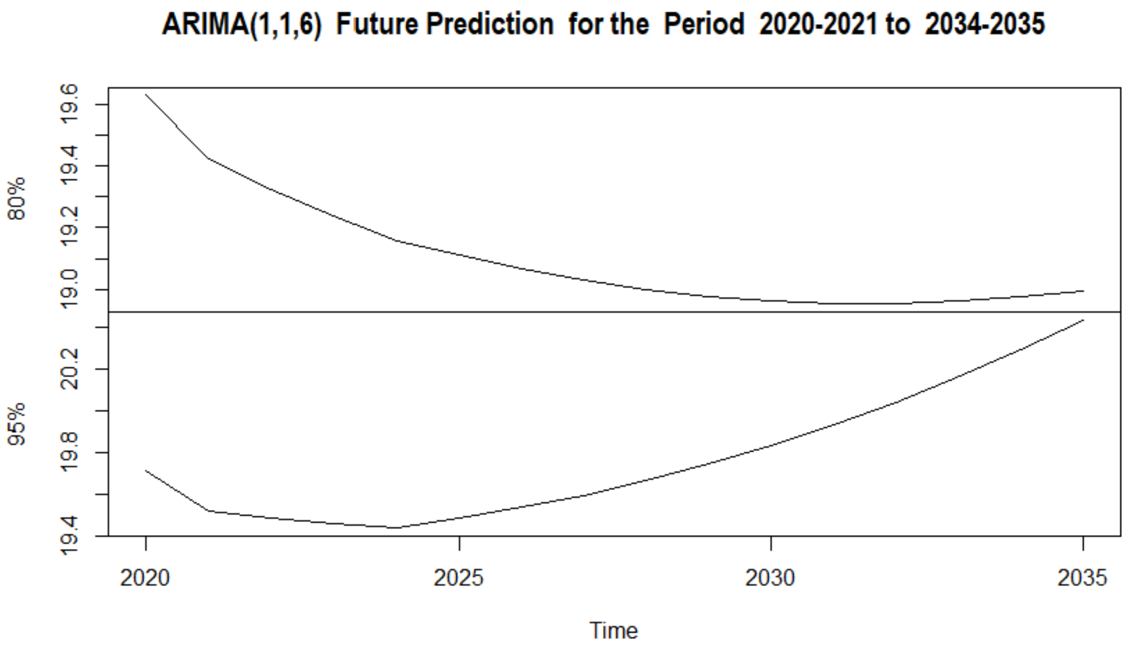

The above **Fig. 12** ARIMA (1,1,6) actual and fitted plot are used in India (1995-2019).




The above **Fig. 13** birth rate fit and forecast plots using ARIMA(1,1,6) in India (1995-2035).




The above **Fig. 14** ARIMA(1,1,6) Residuals plots.




The above **Fig. 15** ARIMA(1,1,6) during the 2020-2021 to 2034-2035 period, the compersion upper future prediction is 80% and 95%. The best future forecast has a 95% higher confidence level.

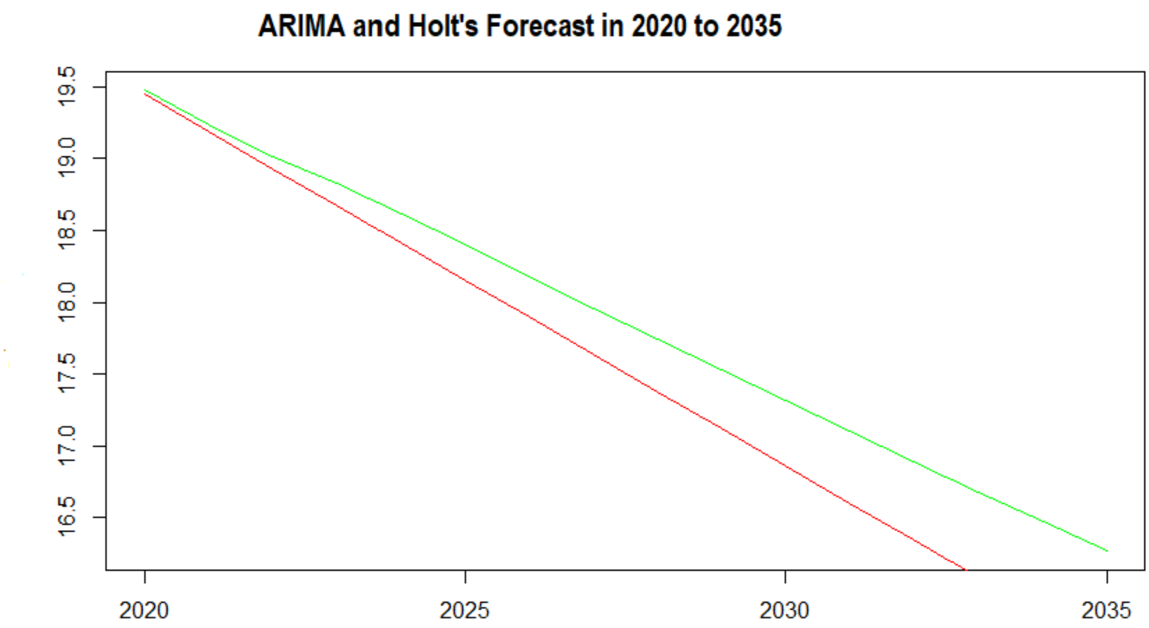


**Fig. 16** shows ARIMA and Holt's models in future forecasting (2020-2035). In comparison to Holt's future prediction line in red, the ARIMA future prediction line in green provides the best prediction.

The birth rate was calculated annually from 1995 to 2019, and it clearly shows that the birth rate has been steadily dropping from 1995 to 2019. From the predict of birth rate for 1995 -2019 from ARIMA, and Holt's model. The out-of-sample future predictions for each model from 2020 to 2035, on the other hand, were not similar. ARIMA's MSE (0.012227) was lower than Holt's method (0.018578), and ARIMA's comparable RMSE (0.110577) was also lower than Holt's method (0.136302). When compared to Holt's method, ARIMA has the lowest MAE (0.087999). (0.094424), ARIMA has the lowest MPE when compared to Holt's method (0.088056). (0.212323), and ARIMA has the lowest MAPE when compared to Holt's method (0.3667742). (0.389555). In this study, the efficacy of time series to predict outcomes was relative to conventional statistical methods. Furthermore, the efficiency of computer hardware such as the central processing unit (CPU) and random-access memory (RAM) has a significant impact on calculation time (RAM). To design treatments for enhancing child survival programmes in India, we prioritised the accuracy of procedures over time.

## 4.Conclusion

The ARIMA model was shown to be more appropriate than other commonly used techniques for estimating future estimations of India's birth rate in this study. Making accurate birth rate forecasts will allow policymakers to take more informed decisions about population growth, family planning programmes, and health-care delivery. It is also critical for government policy, planning, and decision-making in multiple sectors, including maternal and child health, childcare, education, and housing.

## Acknowledgement

Authors are highly thankful to the Editor-in-Chief for his immense support on time and helped in revising the manuscript with efficient reviewers' comments.

## 5.References

1. A. K. Tiwari and S. Mishra, "Fertility pattern in India and estimation of Total Fertility Rate", *International Journal of Agricultural and Statistical Sciences,* vol. 17, no. 1, **(2021),** pp. 2279-2289.
2. A. A. Elhag and H. Abu-Zinadah, "Forecasting under applying Machine Learning and Statistical Models", *Thermal Science*, 24 (Supplement 1): S131-S137, **(2020)**.
3. A. Radkar, "Indian Fertility Transition", *Journal of Health Management*, vol. 22, no. 3, **(2020),** pp. 413–423.
4. A. Nazim and A. Afthanorhan, "A comparison between single exponential smoothing (SES), double exponential smoothing (DES), holt's (brown) and adaptive response rate exponential smoothing (ARRES) techniques in forecasting Malaysia population", *Global Journal of Mathematical Analysis,* vol. 2, no. 4, **(2014),** pp. 276-280.
5. A. F. Fagbamigbe and A. S. Adebowale, "Current and Predicted Fertility using Poisson Regression Model: Evidence from 2008 Nigerian Demographic Health Survey", African Journal of Reproductive Health, vol. 18, no. 1, **(2014)**, pp. 71-83.

6. A. A. Bhende and T. Kanitkar, "Principles of Population Studies. Himalaya Publishing House", *New Delhi, India*, **(1997)**.
7. B. P. Singh, A. K. Tiwari and V. Patel, "Estimating total fertility rate using information on proportion of females of different reproductive age groups", *Bulletin of Mathematics and Statistics Research*, vol. 10, no. 1, **(2022),** pp. 1-8.
8. B. M. Devkota, "Estimation of total fertility rate and birth averted due to contraception: Regression Approach"*, Nepal Population Journal*", vol. 18, no. 17, **(2018)**, pp. 105-112.
9. C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages", *International Journal of Forecasting, Elsevier*, vol. 20, no. 1, **(2004)**, pp. 5-10.
10. D. Dench, T. Joyce and H. Minkoff, "United States Preterm Birth Rate and COVID-19", *RESEARCH BRIEFS*, doi.org/10.1542/peds.2021-055495, **(2021)**.
11. D. A. Adeyinka and N. Muhajarine, "Time series prediction of under-five mortality rates for Nigeria: comparative analysis of artificial neural networks, Holt-Winters exponential smoothing and autoregressive integrated moving average models", *BMC Medical Research Methodology,* vol. 20, no. 1, **(2020)**, pp. 1-11.
12. D. Benvenuto, M. Giovanetti, L. Vassallo, S. Angeletti and M. Ciccozzi, "Application of the ARIMA model on the COVID-2019 epidemic dataset*", Data in Brief*, vol. 29, **(2020)**.
13. D. Asteriou and S. G. Hall, "ARIMA models and Box-Jenkins Methodology: Applied Econometrics", *New York: Palgrave Macmillan,* **(2011)**.
14. H. L. Shang and H. Booth, "Synergy in fertility forecasting: improving forecast accuracy through model averaging", *Genus*, vol. 76, no. 1, **(2020),** pp. 1-23.
15. H. Joshi and D Tyagi, "Forecasting and Modeling Monthly Rainfall in Bengaluru, India: An Application of Time Series Models", *International Journal of Scientific Research in Mathematical and Statistical Sciences*, vol. 8, no. 1, **(2021),** pp. 39-46.
16. J. Bongaarts, "Trends in fertility and fertility preferences in sub-Saharan Africa: the roles of education and family planning programs", *Genus*, vol. 76, no. 32, **(2020)**, pp. 1-15.
17. J. Contreras, R. Espinola, F. J. Nogales and A. J. Conejo, "ARIMA Models to Predict Next-day Electricity Prices*", IEEE Transactions on Power Systems*, vol. 18, no. 3, **(2003)**, pp. 1014-1020.
18. K. B. Pathak and F. Ram, "Techniques of Demographic Analysis. Himalaya Publishing House*", New Delhi, India*, **(2016)**.
19. M. Cao, Y. Pan, Q. Zhang, D. You, S. Feng and Z. Liu, "Predictive value of live birth rate based on different intrauterine adhesion evaluation systems following TCRA", *Reproductive Biology and Endocrinology*, vol. 19, no. 13, **(2021)**.
20. N. Arbu, A. Lim and C. Me-ead, "Forecasting the Patterns and Trends Age-Specific Fertility Rate in South Asia", *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 8, **(2021)**, pp. 2595-2600.
21. N. Kurniasih, A. S. Ahmar, D. R. Hidayat, H. Agustin and E. Rizal, "Forecasting Infant Mortality Rate for China: A Comparison Between α-Sutte Indicator, ARIMA, and Holt-Winters", *Journal of Physics*, vol. 1028, no. 1, **(2018),** pp. 012195
22. P. Pujiharta, Y. D. Darma, N. R. Wiyanti and A. Gunawan, "Comparative Analysis of Arima Model and Exponential Smoothing in Predicting Inventory in Automotive Companies", *Budapest International Research and Critics Institute*, vol. 5, no. 1, **(2022).**
23. P. K. Tripathi, R. K. Mishra and S. K. Upadhyay, "Bayes and Classical Prediction of Total Fertility Rate of India Using Autoregressive Integrated Moving Average Model", *Journal of Statistics Applications and Probability*, vol. 7, no. 2, **(2018),** pp. 233-244.
24. S. Singh, C. Shekhar, A. Bankole, R. Acharya, S. Audam and T. Akinade, "Key drivers of fertility levels and differentials in India, at the national, state and population subgroup levels, 2015–2016: An application of Bongaarts' proximate determinants model", *Plos*

*One*, vol. 17, no. 2, **(2022),** pp. 1-39.
25. S. M. Kearney, B. P. Levine and L. Pardue, The Puzzle of Falling US Birth Rates since the Great Recession. Journal of Economic Perspectives, vol. 36, no. 1, **(2022),** pp. 151-176.
26. Y. He, Y. Tang, S. Chen, J. Liu and H. Liu, "Efect of GnRH agonist alone or combined with diferent low-dose hCG on cumulative live birth rate for high responders in GnRH antagonist cycles: a retrospective study", *BMC Pregnancy and Childbirth*, vol. 22, no. 172, **(2022)**.
27. Z. Liu, Z. Zhu, J. Gao and C. Xu, "Forecast methods for Time Series Data: A Survey". *IEEE Access*, 9: 91896-91912, **(2021).**