ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 11| Issue: 4| April 2025|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2025: 8.691 || ISI Value: 1.188

EXPLORING THE TRENDS OF INFANT MORTALITY IN GUJARAT: USING AUTOREGRESSIVE MOVING AVERAGE AND NEURAL NETWORK AUTOREGRESSION MODELING

Pankaj Kumar D. Parmar[1], Dr. Sanjay Patel[2]
[1]*Research Scholar, Dept. of Statistics, VNSGU*
[2]*Asst. Prof. HOD, Department of Statistics, Udhana Citizen Commerce College, Surat*

ABSTRACT

Over the last thirty years, there has been progress in reducing the infant mortality rate in Gujarat; nevertheless, the state is anticipated to not meet the Sustainable Development Goal (SDG) target for infant mortality. This necessitates increased efforts to tackle this challenge. This research used infant mortality data obtained from the Sample Registration System (SRS), the World Bank, and various reports to evaluate the situation in Gujarat. The study identified trends in infant mortality and developed predictive models to provide an accurate forecast of the mortality rate in the state.

Two time series models were examined and assessed: the Autoregressive Integrated Moving Average (ARIMA) and Neural Network Autoregression (NNAR). The comparison of these models was conducted based on five different metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Percentage Error (MPE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Both models revealed similar trends and were considered appropriate for forecasting; however, the NNAR model surpassed the ARIMA model in all five-evaluation metrics. The analysis shows a declining trend in the infant mortality rate in Gujarat, with the NNAR model projecting a reduction of roughly 5-6% in the rate over the next decade.

KEYWORDS – *Autoregressive Integrated Moving Average (ARIMA), Neural Network autoregression (NNAR), Infant mortality (IMR), Time series, Data modelling*.

I. INTRODUCTION

Child mortality serves as a crucial measure of a nation's socio-economic progress as well as the overall quality of life for its people, particularly mothers. The survival rate of children is predominantly threatened between the ages of 0 and 5, with the highest risk occurring during the first year of life. In 2019, it was reported that worldwide, 70 percent of fatalities among children and youth under 25 occurred within the group of children under 5 years old (UN IGME, 2020). Even more alarming is that a significant portion of these deaths could have been avoided. Infant mortality is defined as the death of a child prior to reaching their first birthday. Child mortality is categorized into three groups: neonatal mortality (0–27 days), infant mortality (1–11 months), and child mortality (12–59 months). According to a report from the World Health Organization, 5.3 million children under the age of five lost their lives in 2018 alone. The likelihood of a child passing away before their first birthday is particularly elevated, especially in Africa, where the rate is 52 per 1000 live births, compared to just 7 per 1000 live births in Europe (World Health Organization, 2018).

The global infant mortality rate (IMR) has significantly declined from 65 deaths per 1,000 live births in 1990 to 29 in 2018, reflecting improvements in socioeconomic conditions. According to the National Family Health Survey-4, the IMR fell by 28% from 57 to 41 per 1,000 live births between 2015 and 2016. The Indian Government aims to reduce the IMR to 28 per 1,000 live births by 2019. Forecasting the IMR using the autoregressive integrated moving average (ARIMA) model can help implement effective strategies to further reduce infant mortality. The annual number of infant deaths decreased from 8.7 million in 1990 to 4.0 million in 2018. In Gujarat, the IMR dropped from approximately 145 deaths per 1,000 live births in 1971 to 23 in 2020. Despite this progress, a significant gap exists between rural and urban IMRs, with the rural rate being 22 deaths higher per 1,000 live births, surpassing rates in most major Indian states, except Assam.

One of the objectives of The Sustainable Development Goals (SDGs) is to eliminate preventable deaths of newborns and children under 5 years old by 2030, with all nations striving to decrease neonatal mortality to as low as 12 per 1,000 live births and under-5 mortality to a maximum of 25 per 1,000 live births. While several states within countries have achieved this goal, Gujarat state is included among them.

Infant mortality notably affects the overall loss of years of human life as it occurs at an early age and is more prevalent than in other age brackets. Infants are particularly susceptible to environmental influences as well as their mothers' health during and after pregnancy, especially in economically disadvantaged communities (UNFPA, 2007). Elevated infant mortality rates can drive families, especially those with low income, to have more children in an attempt to offset the risk of not having any surviving children, which in turn contributes to population growth. This study seeks to model the infant mortality rate in Gujarat by employing Autoregressive Moving Average and Neural Network Autoregression, comparing the models to determine the most effective method.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 11| Issue: 4| April 2025|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2025: 8.691 || ISI Value: 1.188

II. LITERATURE REVIEW

Mishra, Sahanaa, and Manikandan (2019) conducted a study on forecasting India's Infant Mortality Rate (IMR) using the Autoregressive Integrated Moving Average (ARIMA) model. Their primary objective was to apply ARIMA modeling to predict IMR trends for the period 2017 to 2025. Based on their analysis, the ARIMA (2,1,1) model was identified as the most suitable, demonstrating a continuous decline in IMR over the forecasted years. The study projected a decrease in IMR from 33 per 1,000 live births in 2017 to 15 per 1,000 live births by 2025, emphasizing the ongoing improvements in child healthcare initiatives in India.

Agarwal, Tripathi, and Pareek (2021) compared traditional time series forecasting methods with Bayesian approaches to predict India's infant mortality rates. They first ensured data stationarity through logarithmic transformations and second-order differencing before selecting the most appropriate model using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The researchers applied maximum likelihood estimation for the traditional model, whereas for the Bayesian model, they employed fuzzy priors and Markov Chain Monte Carlo (MCMC) simulations. Their findings indicated that the Bayesian approach provided higher accuracy in IMR predictions for both past and future years (2019–2023), suggesting its potential advantages over traditional statistical techniques.

Bhatia et al. (2019) examined infant and under-five mortality trends in India from 1992 to 2016, analyzing regional variations using three rounds of nationally representative household surveys. Their study focused on both high-performing states (such as Kerala, Maharashtra, Punjab, Tamil Nadu, and Haryana) and low-performing states (including Bihar, Madhya Pradesh, Uttar Pradesh, Chhattisgarh, and Rajasthan). The results confirmed a steady decline in IMR and child mortality across India, but also highlighted significant disparities between states, emphasizing the need for region-specific health interventions to bridge these gaps.

Gorr (1994) emphasized the need for further research to identify conditions under which Artificial Neural Networks (ANNs) or traditional statistical models outperform one another. His study highlighted the ongoing debate among researchers regarding the effectiveness of ANN-based forecasting versus traditional time series methods like ARIMA. While some studies favor traditional approaches due to their interpretability and statistical rigor, others advocate for ANNs' ability to capture complex nonlinear patterns, urging future studies to explore their respective strengths and weaknesses.

III. OBJECTIVE

The trend of infant mortality was identified, and models were developed to accurately predict the mortality rate in Gujarat. To assess and compare the performance of the ARIMA and NNAR models, evaluation was conducted using five criteria: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Percentage Error (MPE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). This study aimed to determine which model performed better.

IV. MATERIALS AND METHODS

Data Source

This study was conducted using secondary data obtained from the World Bank www.worldbank.org and SRS data (Simple Random Sampling). The data contains the mortality for infants in Gujarat State between 1971 and 2020.

Data Analysis

Two time series models, namely the Autoregressive Integrated Moving Average (ARIMA) and Neural Network Autoregression (NNAR), were utilized to predict annual infant mortality rates; their forecasting accuracies were compared, and the superior model was employed to anticipate infant mortality rates for the years 2022, 2023, 2024, and 2030. The historical mortality data covers the period from 1971 to 2021, resulting in a total of 51 observations. To prevent underfitting or overfitting the model due to inadequate dataset training, the models were constructed with the dataset divided into two parts. Approximately 80% of the time-series data formed the training dataset, while the remaining portion was set aside as a test/validation dataset to assess model accuracy. The effectiveness of these methods was evaluated by analyzing their error metrics, which included the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Percentage Error (MPE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). All analyses were conducted using the R-Studio development software and SPSS-25. The criteria can be respectively derived by using;

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 \ldots\ldots\ldots\ldots \quad (1)$$

$$RSME = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \ldots\ldots\ldots\ldots (2)$$

$$MPE = \frac{100\%}{n}\sum_{t=1}^{n}\frac{y_t - \hat{y}_t}{y_t} \ldots\ldots\ldots\ldots (3)$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{|y_t - \hat{y}_t|}{\frac{1}{n-1}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}\right) \ldots\ldots\ldots\ldots \quad (4)$$

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n}\left(\frac{|y_t - \hat{y}_t|}{y_t}\right) \ldots\ldots\ldots\ldots (5)$$

where n represents the number of observations, y_t denotes the actual observation value at time t, and $\hat{y}_t$ stands for the predicted observation value. The RMSE, MSE, and MAE are particularly significant because minimizing the first two metrics enhances the accuracy of predictions, while minimizing the third measure leads to the prediction of the median (Perone, 2021). Both MAPE and MPE are also crucial as they are not scale-dependent. Among these, MAPE is the most widely utilized measure (Goodwin & Lawton, 1999; Kim & Kim, 2016), whereas MPE has the additional benefit of serving as a bias measure since both positive and negative forecast errors can offset one another.

V. RESULTS, FINDINGS AND DISCUSSION

The time series graph depicting the infant mortality rate is show in Figure 1. This figure reveals a consistent downward trend, indicating that Gujarat's infant mortality rate has steadily declined from 145 in 1971 to 23 in 2021. During this period, there was a relatively stable horizontal movement from 1975 to 1993, suggesting that the infant mortality rate did not exhibit significant increases or decreases but instead fluctuated between 106 and 109.

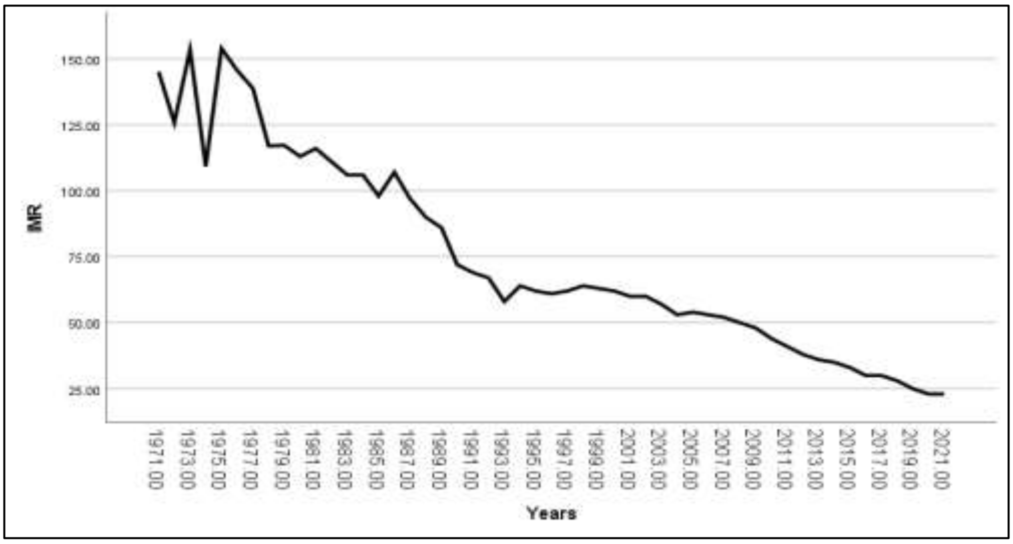

Figure-1: The time series plot of Infant mortality in Gujarat

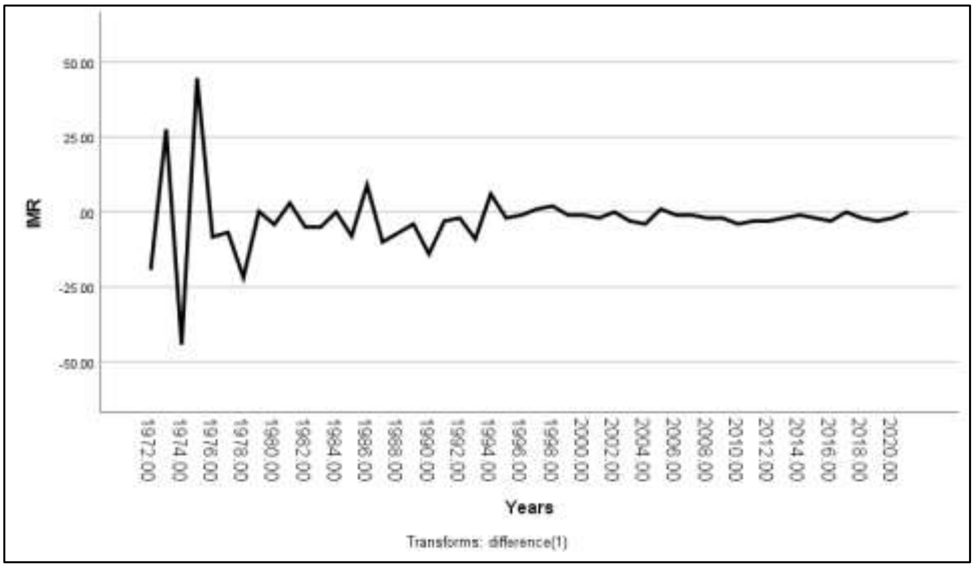

Figure-2: The time series plot after first differencing of Infant mortality dataset in Gujarat

The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test is utilized to check for the stationarity of a time series. This step is essential before proceeding with appropriate modeling. A significance level of 0.05 is applied when testing the stationarity of the series. When the series is stationary, it suggests that the data does not exhibit a unit root. The test's hypotheses are formulated as follows:

Hypothesis

- **Null hypothesis (H0):** Series is **stationary**.
- **Alternative hypothesis (H1):** Series is **non-stationary**.

Table-1 shows the result of the KPSS test. From the table, it can be inferred that the series is **non-stationary** as the test statistic and p-value of the KPSS test are obtained as **1.0685** and **0.01** respectively, which is less than the level of significance. The plot of the stationary dataset obtained through the first differencing is shown in **Figure-2**.

Table -1 Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test

KPSS Test	t-Statistic	Prob.*
KPSS Test Statistics	1.0685	0.001

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 11| Issue: 4| April 2025|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2025: 8.691 || ISI Value: 1.188

One approach to convert a stationary time series into a non-stationary one is by applying the lag. Generating the lag of the data yields a stationary series following the first differencing. Once a stationary series is achieved, modeling is performed using both ARIMA and NNAR methods.

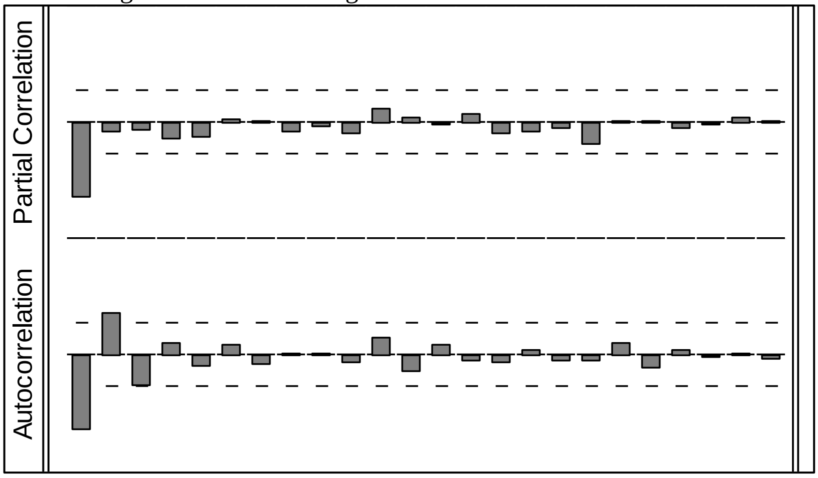

Figure-3 ACF and PCF Graph

The given **Figure-3** shows the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), essential tools in time series analysis. In the ACF plot, a strong positive correlation at lag 1 followed by values within the confidence limits suggests minimal autocorrelation beyond the first lag, indicating a possible MA (1) process. The PACF plot shows a significant negative correlation at lag 1 with no notable values afterward, suggesting a potential AR (1) process. This pattern aligns with characteristics of a stationary series, possibly achieved through differencing. Such behavior is common in ARIMA models, particularly AR (1), MA (1), or ARMA (1,1) structures.

5.1 Modelling the data with Autoregressive Integrated Moving Average (ARIMA)

The stationary series was analyzed using the Autoregressive Integrated Moving Average (ARIMA) method. Model selection was conducted using the Akaike Information Criterion (AIC), which helped determine the appropriate Autoregressive (AR) and Moving Average (MA) components. Additionally, the "Auto-ARIMA" function from the "time-series" package in R software was used to assist in selecting the best-fitting model. Based on this criterion, the ARIMA (1,1,1) model was identified as the most suitable. The model estimates are provided in Table 2.

Table-2 ARIMA (1,1,1) model estimates for Infant mortality rate in Gujarat

Model		AR1	MA1	Sigma	Log-Likelihood	AIC
ARIMA (1, 1, 1)	Co-efficient	-0.8332	0.3413	79.4	-180.63	367.26
	St. Error	0.1778	0.2809			

Table-3 Ljung-Box Q Test

Model	Statistics	DF	Sign.
ARIMA (1, 1, 1)	5.4099	8	0.8622

The ARIMA (1,1,1) model was identified as the best fit for analyzing the Infant Mortality Rate (IMR) in Gujarat, based on the Akaike Information Criterion (AIC) and model selection techniques. The estimated AR (1) coefficient of **-0.8332** indicates a strong negative relationship between current and past IMR values, meaning that an increase in one year is likely to be followed by a decrease in the next. The MA (1) coefficient of **0.3413** suggests that past forecasting errors have a minor influence on current values. The model's AIC of **367.26** confirms its efficiency compared to alternative models.

From Table 3, we do not reject the null hypothesis, indicating that the Ljung-Box Q test is not statistically significant at the 5% level. This conclusion is reinforced by the model's p-value of 0.8622, which is greater than 0.05, confirming that the residuals are independently distributed and do not exhibit serial correlation. Consequently, the ARIMA (1,1,1) model is recommended for forecasting. Other models considered were found to be less suitable due to higher AIC values. These include: ARIMA (1,1,0) with an AIC of 366.87, ARIMA (1,1,2) with an AIC of 369.19, ARIMA (2,1,1) with an AIC of 369.16 and ARIMA (2,1,2) with an AIC of 371.16. This comparison further supports the selection of ARIMA (1,1,1) as the most appropriate model for analyzing the Infant Mortality Rate in Gujarat. This validates the reliability of the ARIMA (1,1,1) model for forecasting IMR trends. Overall, the selected model provides a statistically sound and effective approach to predicting infant mortality rates in Gujarat.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 11| Issue: 4| April 2025|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2025: 8.691 || ISI Value: 1.188

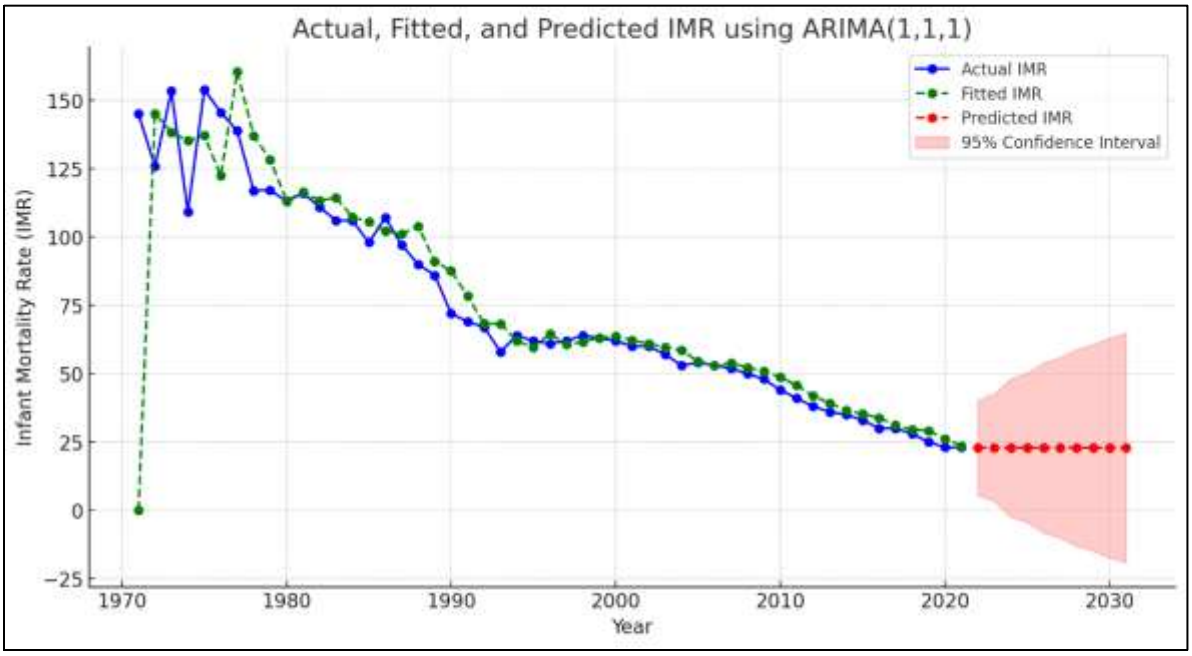

Figure- 4 The Forecasted from ARIMA (1, 1, 1)

Figure-4 illustrates the forecasted values obtained from the ARIMA (1,1,1) model. The model was trained and validated using infant mortality rate data from 1971 to 2021. As shown in the graph, the predicted IMR values generally follow the observed trend but tend to be slightly lower than the actual recorded figures.

5.2 Modelling the data Neural Network Autoregression (NNAR)

The Feed-Forward neural network core algorithm was utilized to implement the time series artificial neural network. For analyzing the dataset on infant mortality rates in Gujarat, the model employed is NNAR (1, 1), averaging 20 networks, where each network is a 1-1-1 configuration with four weight options for linear output units and the estimated sigma value is 0.007693.

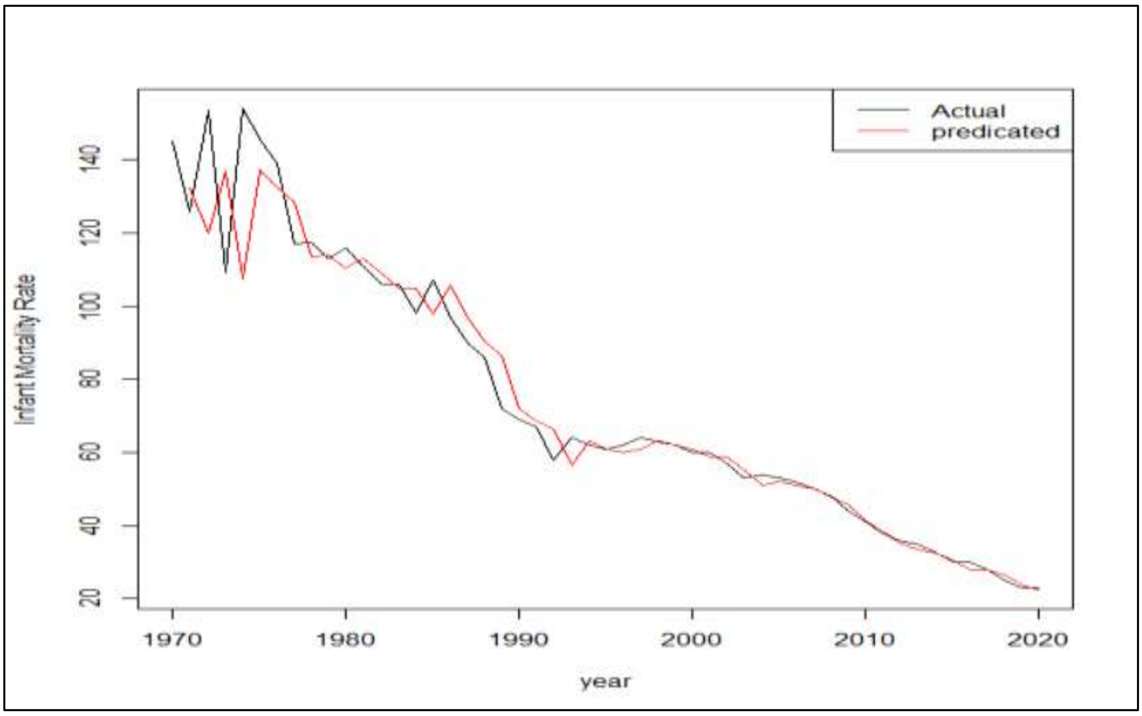

Figure- 5 The plot of the Actual and Predicted values using NNAR (1,1)

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 11| Issue: 4| April 2025|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2025: 8.691 || ISI Value: 1.188

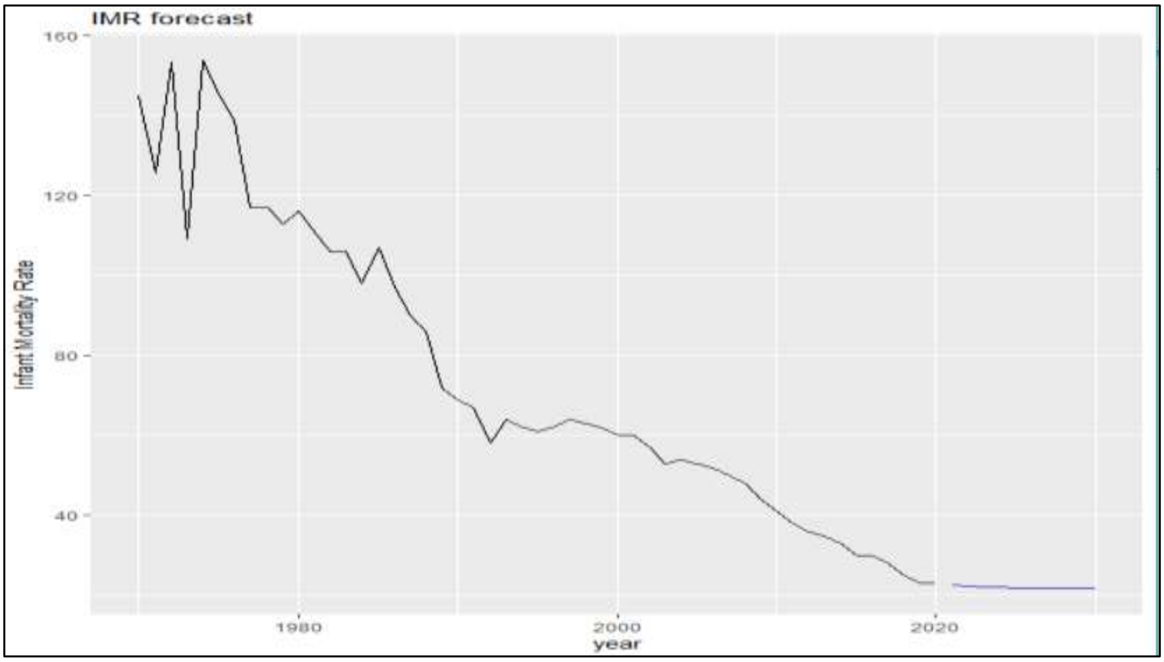

Figure- 6 The plot of the Forecasted values using NNAR (1,1)

Figures 5 and 6 present the actual values, predicted values, and forecasted values. We performed a validation of the NNAR model in a manner similar to that of the ARIMA (1,1,1) model depicted in Figure 4, which is shown in Figure 6. To train the model, we utilized the expected future occurrences from 2022 to 2030, ensuring these data were not included in the model's development. The model's predictions initially approximate the actual values closely, and they become increasingly accurate as the years increase.

5.3 Model Comparison and Selection

In general, both models produce values that differ slightly from the actual results, indicating that each model is effective in representing the data. However, it is always advisable to select the optimal model when presented with alternatives. To assess the performance of these models, five (5) criteria were employed. The criteria used for evaluation include the Mean Squared Error (MSE), the Root Mean Squared Error (RMSE), the Mean Percentage Error (MPE), the Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE).

Table -4 Comparison of ARIMA and NNAR Model

Criteria	ARIMA (1,1,1)	NNAR (1,1)
MSE	3.2222	0.4142
RSME	8.8227	10.0457
MPE	5.4028	-0.3725
MAE	5.8840	5.1073
MAPE	7.5171	5.4425

Table -4 shows the performances of the two models based on the five criteria considered. Based on the value of all five criteria, the NNAR can be considered a better alternative to the ARIMA model when modelling the time series data of infant mortality rate in Gujarat state. Figure-7 further shows the validity of the inference. The predicted value of the NNAR is much closer to the actual values through the years considered.

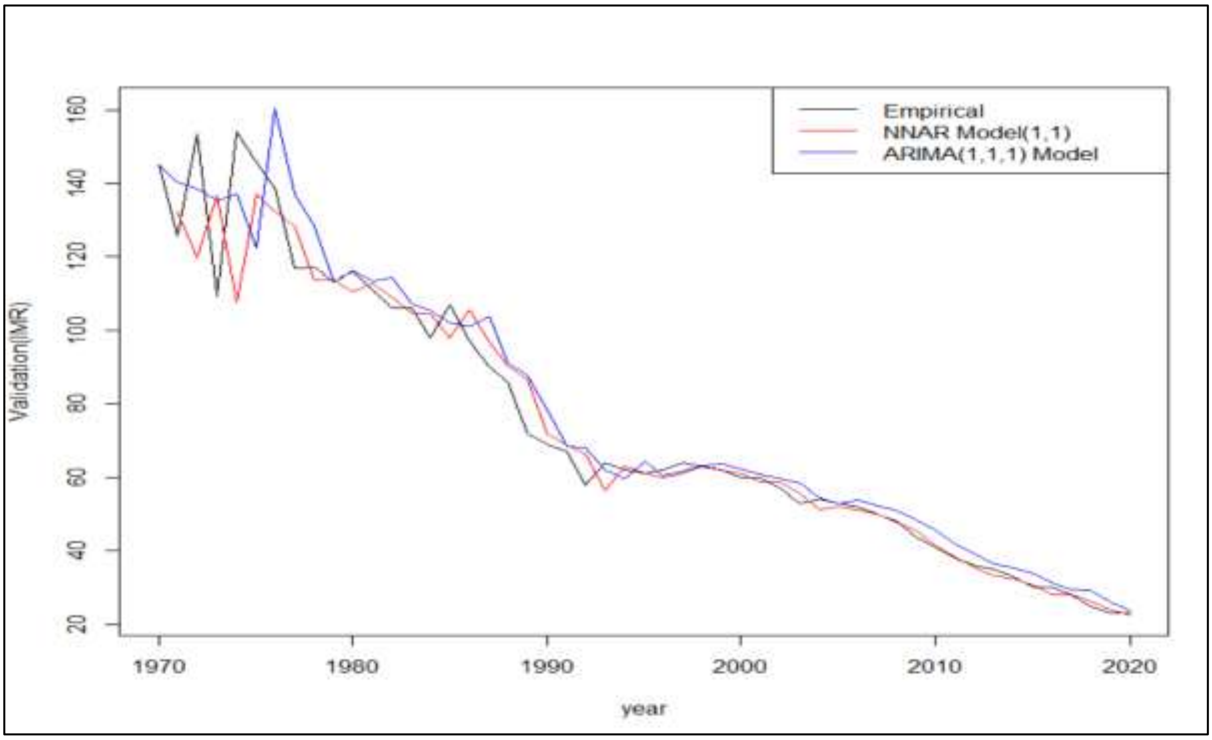

Figure-7 Plot of the empirical data with the estimated values based on ARIMA and NNAR

Based on the data presented in Table-4 and Figure-6, predictions made using the NNAR model indicate that the infant mortality rate in Gujarat is projected to be 22.6120 by the end of 2021, 22.6394 by the end of 2022, 22.1498 by the end of 2023, and 21.7640 by the end of 2030. These figures suggest a continuing reduction in the infant mortality rate in the coming years, with an anticipated decrease of approximately 5% to 6% by the end of 2030.

VI. CONCLUSION

The analysis of Infant Mortality Rate (IMR) trends in Gujarat over the past several decades indicates a consistent and significant decline. In the 1970s, Gujarat's IMR was over 150 deaths per 1,000 live births, but by 2020, it had dropped below 30 per 1,000 live births. However, after 2020, the rate of decline has stabilized, suggesting that while significant progress has been made, additional efforts are required to achieve further reductions. This plateau may be attributed to socioeconomic disparities, healthcare accessibility, and maternal health factors that require targeted interventions.

In comparing forecasting models, the Neural Network Autoregressive (NNAR) model demonstrated superior predictive accuracy over the traditional ARIMA (1,1,1) model. The ARIMA model, selected based on the Akaike Information Criterion (AIC = 367.26) and Ljung-Box Q test (p-value = 0.8622), captured the overall trend but lacked the ability to model complex nonlinear relationships. In contrast, the NNAR model produced lower forecasting errors and better aligned with historical trends, highlighting the effectiveness of Artificial Neural Networks (ANNs) in capturing nonlinear dependencies in IMR data.

Based on the forecasting results, Gujarat's IMR is expected to decline by 5% to 6% in the coming years. If current trends persist, the state may reach an IMR of approximately 25 per 1,000 live births by 2030, though this remains above the Sustainable Development Goal (SDG) target of 12 per 1,000. To further accelerate this progress, policy measures should focus on improving maternal healthcare, increasing institutional deliveries, enhancing neonatal care facilities, and addressing rural-urban disparities. Integrating machine learning models like NNAR for real-time monitoring and decision-making can also help optimize health strategies.

In conclusion, while Gujarat has made substantial progress in reducing infant mortality, achieving further improvements will require sustained policy efforts, technological advancements, and targeted healthcare interventions. These findings emphasize the importance of data-driven decision-making in shaping future health policies and achieving long-term reductions in IMR.

VII. REFERENCES

1. *World Health Organization (2018), World health statistics 2018: monitoring health for the SDGs, sustainable development goals. World Health Organization, Cop.*
2. *WHO. (2020), December. Health in 2015: From MDGs to SDGs. https://www.who.int/ data/gho/publications/mdgs-sdgs*
3. *Nurhasniza Idham Abu Hasan, Azlan Abdul Aziz (2022), Projection of Infant Mortality Rate in Malaysia using R, Journal Sains Kesihatan Malaysia 20 (2) 2022: 23-36*
4. *Anupama Singh, Prof. K.N.S. Yadava and Km. Priti (2015), Projection of Infant Mortality Rate: Application of Life Expectancy, Global Journal of Multidisciplinary studies, ISSN: - 2348-0459, Volume-4, issue-4 March 2015.*
5. *Bamidele Mustapha Oseni1 and Olalekan Joseph Igboroodowo (2022), Rattanakosin Journal of Science and Technology: RJST 2022, Volume 4 Issue 2: 1-9, ISSN: 2673-0553.*
6. *Bongaarts, J. (2005). Long-range trends in adult mortality: Models and projection methods. Demography 42(1): 23-49.*

7. *Perone, G. (2021), Comparison of ARIMA, ETS, NNAR, TBATS, and hybrid models to forecast the second wave of COVID-19 hospitalizations in Italy, The European Journal of Health Economics. doi:10.1007/s10198-021-01347-4*
8. *Ewere, F. & Eke, D. O. 2010. Time Series Analysis and Forecast of Infant Mortality Rate in Nigeria: An Arima Modeling Approach. Canadian Journal of Pure and Applied Sciences 14(2):5049-5059*
9. *SRS Bulletin, December 2011, Volume. 46. No.1. Registrar General of India.*
10. *National Health Profile 2010, Registrar General of India.*
11. *District Heath Survey, Gujarat.*
12. *https://www.dhsprogram.com/pubs/pdf/FR374/FR374_Gujarat*
13. *Indirect Estimates of District wise IMR and Under 5 Mortality using Census 2011 data.*