Section A -Research paper
ISSN 2063-5346

FORECASTING INFANT MORTALITY RATE IN PAURI-GARHWAL REGION: AN APPLICATION OF AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

Nidhi Gairola[1*], O.K. Belwal[1]

[1]Department of Statistics, H.N.B. Garhwal University (A Central University), Srinagar-Garhwal (UK) India.

Corresponding author E-mail: nidi123gairola@gmail.com

Article History: Received: 06.04.2023 Revised: 29.05.2023 Accepted: 19.06.2023

Abstract:

Infant mortality is a major indicator of public health because it affects many things. Globally, infant mortality appears to have decreased by an average of about 23 per 1,000 live births and infant mortality by an average of about 25 per 1,000 live births. In 2000, when the United Nations adopted the 8th Millennium Development Goal (MDG), India's infant mortality rate (IMR) was 25 times that of developed countries, and it was even higher in rural India. In NFHS-3, 4 and 5, the infant mortality rate in Uttarakhand was 42, 40 and 39.1 respectively. Infant mortality is considered an important indicator of the quality of life and socio-economic status of a country. One of the Sustainable Development Goals is to reduce the mortality rate to 25/1,000 live births by 2028. One of the most well-known ways to predict this is the automated ARIMA model introduced by Box and Jenkins (Box-Jenkins model) in the 1960s. IMR prediction can help implement interventions to reduce infant mortality within the target range. Projections for the model period (2017-2022) showed the accuracy of the chosen ARIMA method (2, 1, 2). ARIMA post-model projections (2, 1, 2) show a downward trend in the IMR (2023). When actual data is available for the next year (2023), the model can be validated and a more accurate forecast can be made.

Keywords: ARIMA, Infant Mortality Rate, Akaike information criterion (AIC), Bayesian information criterion (BIC), socioeconomic

INTRODUCTION

Worldwide, infant mortality appears to have declined on average by about 23/1000 live births and mortality of older infants by about 25/1000 live births (Rutstein, 1984). In 2000, when the UN adopted eight-millennium development goals (MDG), the infant mortality rate (IMR) in India was more than 25 times, with rural India being higher than rural areas in the developed countries. In total, more than 5.0 million children under age 5, including 2.3 million newborns, along with 2.1 million children and youth aged 5 to 24 years – 43 per cent of whom are adolescents – died in 2021 (UNICEF, 2022). IMR reflects the socioeconomic development of a nation. India was identified as a high disparity country on absolute and relative scales on infant mortality Chao et al., 2018). According to nationally representative data NFHS-5 (2019–2021), there is a considerable variation in the mortality rate in different states of India. The highest mortality rate per thousand is observed in Uttar Pradesh (60); Chhattisgarh (50); Madhya Pradesh (49); Jharkhand (45); Odisha (41); Rajasthan (37) and

whereas the low mortality rate states are Tamil Nadu (22); Kerala (5). In India, the caste differential in mortality is found to be minimal whereas, the mortality among the poorest wealth quintile (59) is thrice higher than that of the richest wealth quintile (20). United Nations Millennium Development Goals (2000) aimed to reduce mortality by two-thirds (MDG 4) by 2015, but many poor countries could not achieve the target and were labeled as 'off-track', 'insufficient progress', or 'no progress' (Cha, 2017). Poor countries faced many challenges to achieving this target, such as lack of health care infrastructure and health professionals, lack of resources and technology, vaccines, lower literacy rate, low household wealth, whereas rich countries increase household wealth and mother literacy, which help to reduce mortality (Lomazzi et al., 2014). The call for SDGs with the motto 'No one left behind' attempted to reduce inequalities across gender, region, class, and caste. The proposed SDG target for mortality aims to reduce at least as low as 25 deaths per 1000 live births by 2030. Prior studies suggest that India had the largest number of deaths of all countries in 2015, with substantial sub national disparities and the enormous absolute disparities (Liu et al., 2019). While India achieved MDG on child mortality at the national level, many states, regions and some socioeconomic groups lagged behind to achieve it.

The rural Indians always experience higher mortality than urban Indians. There are considerable rural-urban infant mortality differentials existing at the national and state levels irrespective of the level of the mortality. It was found that wide disparity in socioeconomic and community-level factors was the reason for rural-urban gap in mortality (Saikia et al., 2013). It should be noted that there are continuous efforts to reduce mortality particularly in rural India through the intervention of different programs like the National Health Mission and other initiatives such as Janani Shishu Suraksha Karyakram (JSSK), Rashtriya Bal Swasthya Karyakram (RBSK), Mother and child health wings (MCH Wings), District hospital and knowledge centre (DHKC), National Iron+ Initiative, National ambulance services, National Mobile Medical Units (NMMUs), Poshan Abhiyan etc. Yet, a quick glance at the rural-urban gap in mortality shows that rural-urban gap still exists in high and low mortality states such as Chhattisgarh (27), Uttar Pradesh (13), Madhya Pradesh (14), Orrisa (11), Rajasthan (6), Jharkhand (28) and Tamil Nādu (9). It is imperative to investigate whether "place of residence" creates any mortality divide and if so, it is crucial to find out the factors behind such a gap to fulfill the DG motto-No one left behind. Forecasting demographic characteristics like fertility, morbidity, mortality, etc., is an important facet for the socio-economic planners as it facilitates them to analyse and regulate policies for the betterment of the human population. To forecast such characteristics, require an appropriate model building so that a reliable result can be obtained. IMR fluctuates according to the health status of the country, which is a dynamic process. In this regard, if the forecast IMR is available, intervention could be planned and implemented effectively at the right time. One of the most common methods of forecasting is the autoregressive integrated moving average (ARIMA) model, widely used in the field of health and agricultural sciences, being simple and easy to use (Mandal, 2005). As infant mortality is one of the indicators of the health status of the country, an attempt was made, in the current study, to forecast the IMR of Pauri division for the year of 2023 with a detailed stepwise explanation of ARIMA model.

MATERIALS AND METHODS

The current study was based on the secondary data analysis of IMR of Pauri between 2017 and 2022, collected from the PHC, CHC and Govt data. The data available to the public are not individually identifiable (the data referred to in the present study were the annual mortality rates and not the individual data). Ethical approval was not obtained for this study as there was no direct involvement of any human subject. For the analysis (forecast), the ARIMA model introduced by Box and Jekins (also called Box–Jenkins model) in the 1960s for forecasting a time series variable was used.

The ARIMA method is an extrapolation method for forecasting data/variables, and like other methods of forecasting, it requires only long-time series data. To forecast any event, the best ARIMA model has to be selected/prepared as per the following three steps (i.e., model identification, parameter estimation, and model validation). Once the model is validated, the procedure could be carried out with the available data. In the current study, the data for IMR for the period of 2017-2022 (6 years) were used to fulfill the need for a time series data set for analysis using the ARIMA model. The expression for the ARIMA (p, d, q) model is $Y_t = b_0 + \Phi_1 Y_{t-1} + \ldots + \Phi_p Y_{t-p} + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q} + e_t$, where Y_t and e_t are actual and random error at period t, respectively, and the other variables involved are the order of autoregressive (AR) part (p), the degree of differencing involved (d), and the order of the moving average (MA) part (q).

The first step (model identification) in forecasting any time series data using ARIMA model is to check how stationary the available data is because the ARIMA model is used when the data are stationary. A stationary data series is those whose values vary over time around a constant mean and variance. Hence, to apply the ARIMA model to forecast data that are nonstationary, it should be transformed into a stationary data series. There are several ways to ascertain stationarity of data. The most common method to check stationarity of data is by examining its graph/time plot (Upadhyay, 2015).

The transformation of data from nonstationary to stationary can be done by considering appropriate differencing of the data. The newly constructed variable after considering the difference of order (d) is "Xt" which can be examined for stationarity. For each difference of order (d), the final hypothetical time series data should be checked for stationarity. The stationarity of the new time series hypothetical data can be checked by ACF and PACF. In the current study, the difference of order (d)1 was sufficient to achieve stationarity in mean. The stationarity of the hypothetical time series data after difference of order 1 was checked by ACF and PACF of IMR data.

After transforming the data to a stationary set, the next step (parameter estimation) is to identify the values of p and q, where p is the order of AR component and q is the order of MA component. When AR and MA are combined, it is known as ARMA. However, in time series data, two other components besides AR and MA are the "trend (long-term variation)" and "seasonality (short-term variation)." When the ARMA model is integrated with the trend and seasonality component, it is known as the ARIMA model. For each ARIMA model to

forecast the data under consideration, a set of values, i.e., (p, d, q) is required. The value of "d" is already estimated as 1 for the current forecast. Using different values (0, 1, 2) for p and q, with permutations and combinations, many models can be suggested. The criteria for the best model of forecasting for the available data are the minimum Akaike information criterion (AIC) and Bayesian information criterion (BIC). For the current study, to consider different values for *p* and *q*, the AIC and BIC values were estimated and the model (ARIMA 2,1, 2) with lowest values was selected for the forecast.

The last step (model validation) in ARIMA model estimation is the validation of the selected model. There are many methods of validating the selected model, but the most common (goodness of fit) is by examining the ACF and PACF of the residuals of IMR.

Table 1: Model identification for infant mortality rate forecasting

Model	**AIC**	**BIC**
ARIMA (0,1,0)	252.98	254.29
ARIMA (0,1,1)	251.11	252.16
ARIMA (1,1,0)	249.13	255.23
ARIMA (1,1,1)	251.43	256.45
ARIMA (2,1,1)	243.21	250.41
ARIMA (2,1,2)	241.56	248.15

AIC=Akaike information criterion, BIC=Bayesian information criterion, ARIMA=Autoregressive integrated moving average

RESULTS

The data are nonstationary as there is a downward trend and the values did not vary over time around a constant mean and variance. Another method of checking for stationarity of data is the autocorrelation function (ACF) and partial ACF (PACF) of the IMR time series data. The values are not within the upper and lower confidence limits, which signify that the data set is nonstationary. The ACF and PACF of residuals of IMR at different lag are within the confidence level. This proves that the selected ARIMA model is appropriate for the forecast of the available data or data under consideration.

Using the selected ARIMA (2, 1, 2) model, the IMR was forecasted for post-sample period of 2023. The results showed a declining trend of IMR in the predicted period. The predicted infant mortality in Pauri division is going to decline from current rate. Figure 1 and 2 shows autocorrelation function (ACF) and partial autocorrelation function (PACF) of infant data considering upper and lower limit. Figure 3 and 4 shows ACF and PACF for infant mortality

rate in Pauri division for selection of appropriate model. Figure 5 shows comparative analysis of ACF and PACF of infant forecasting of selected area while figure 6 shows the prediction of infant mortality for 2023 based on previous data since 2017. The autocorrelation series of Box-Ljung depicted in table 1 along with std. error. Table 2 shows PACF standard error. Model summary of good fit of forecasting is given in table 3. Table 4 shows residual ACF summary considering mean, minimum, maximum and percentile for trend identification given in table 5. The model fit statistics of Ljung-Box test is listed in table 6 and it shows Normalized BIC is 3.258 which shows significant decline in infant mortality in selected period. The Ljung-Box is a type of statistical test of whether any of a group of autocorrelations of a time series are different from zero calculated based on formula Q (m) = N (N+2) ∑ h = 1 m ρ ^ h 2 N – h and it examines whether there is significant evidence for non-zero correlations at given lags.

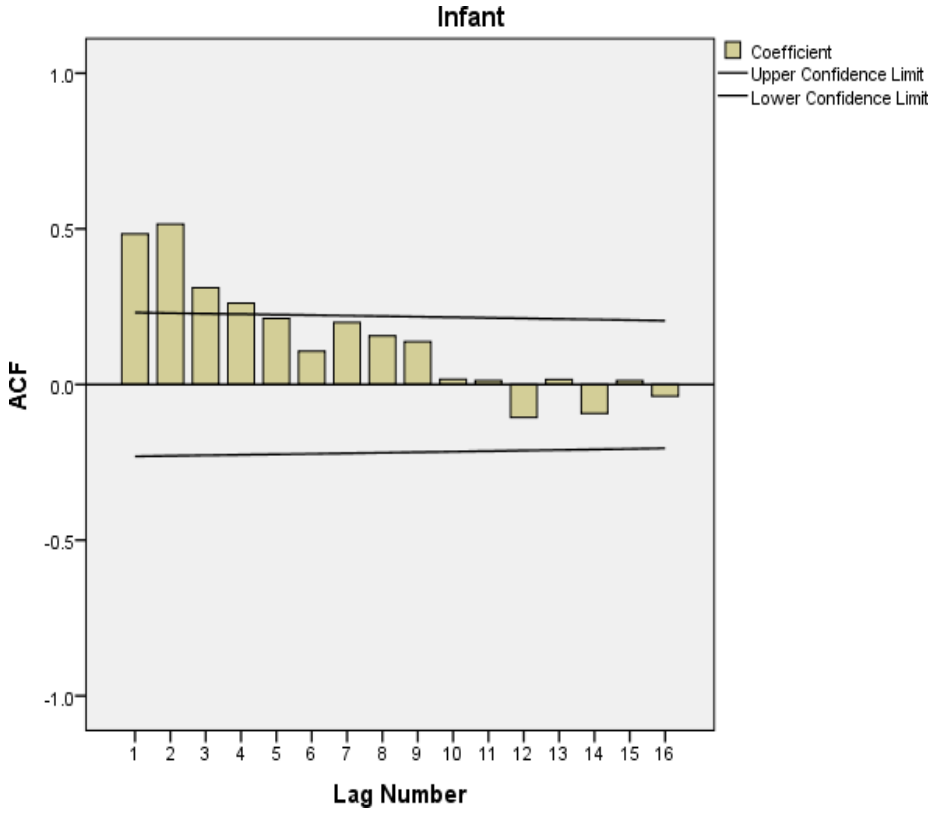

Figure 1. Autocorrelation function of Infant data on lower and upper limit

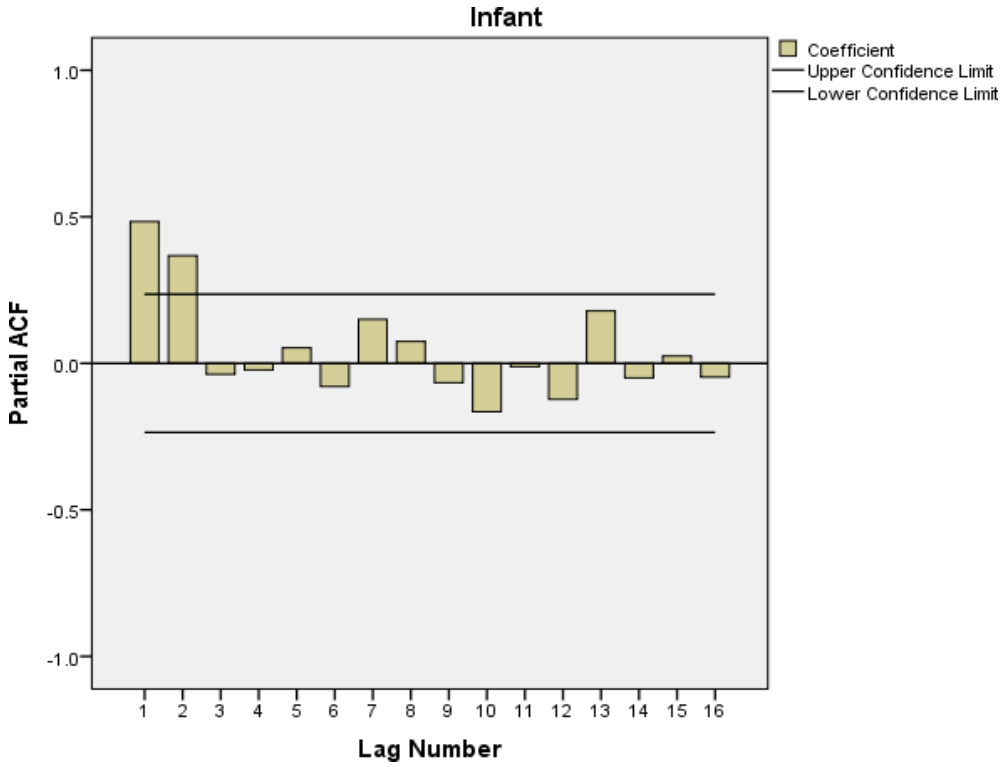

Figure 2. Partial autocorrelation function of Infant data on lower and upper limit

Table 1. Autocorrelations Series

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic		
			Value	df	Sig.[b]
1	-.563	.116	23.452	1	.000
2	.243	.115	27.874	2	.000
3	-.121	.115	28.991	3	.000
4	-.013	.114	29.004	4	.000
5	.030	.113	29.075	5	.000
6	-.137	.112	30.577	6	.000
7	.140	.111	32.154	7	.000
8	-.018	.110	32.180	8	.000
9	.046	.109	32.359	9	.000
10	-.080	.108	32.907	10	.000
11	.084	.108	33.521	11	.000
12	-.233	.107	38.291	12	.000
13	.227	.106	42.886	13	.000
14	-.198	.105	46.458	14	.000
15	.144	.104	48.373	15	.000
16	-.038	.103	48.508	16	.000

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

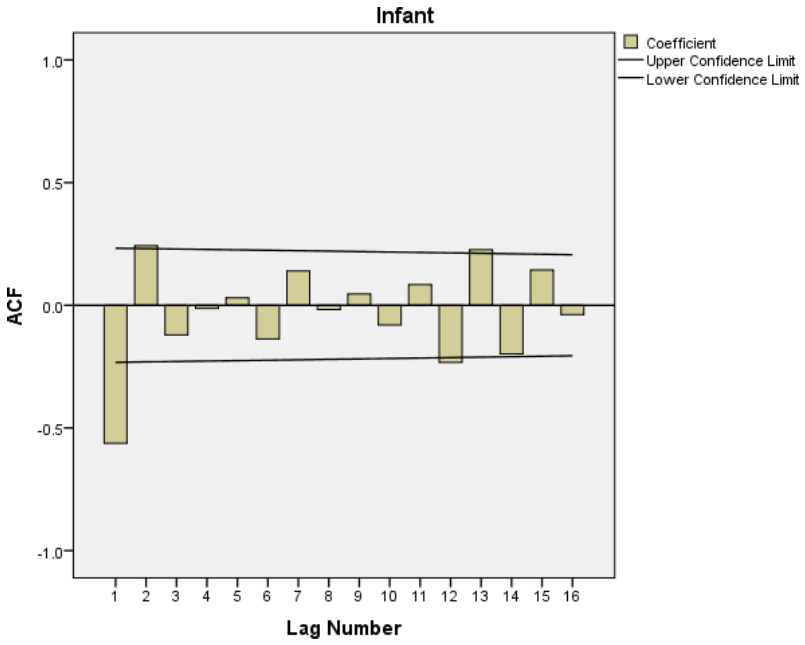

Figure 3. Autocorrelation function of infant birth in Pauri Division for selection of model

Table 2. Partial Autocorrelations

Series: Infant		
Lag	Partial Autocorrelation	Std. Error
1	-.563	.119
2	-.108	.119
3	-.046	.119
4	-.121	.119
5	-.053	.119
6	-.193	.119
7	-.054	.119
8	.082	.119
9	.099	.119
10	-.052	.119
11	.026	.119
12	-.259	.119
13	-.010	.119
14	-.062	.119
15	-.037	.119
16	-.022	.119

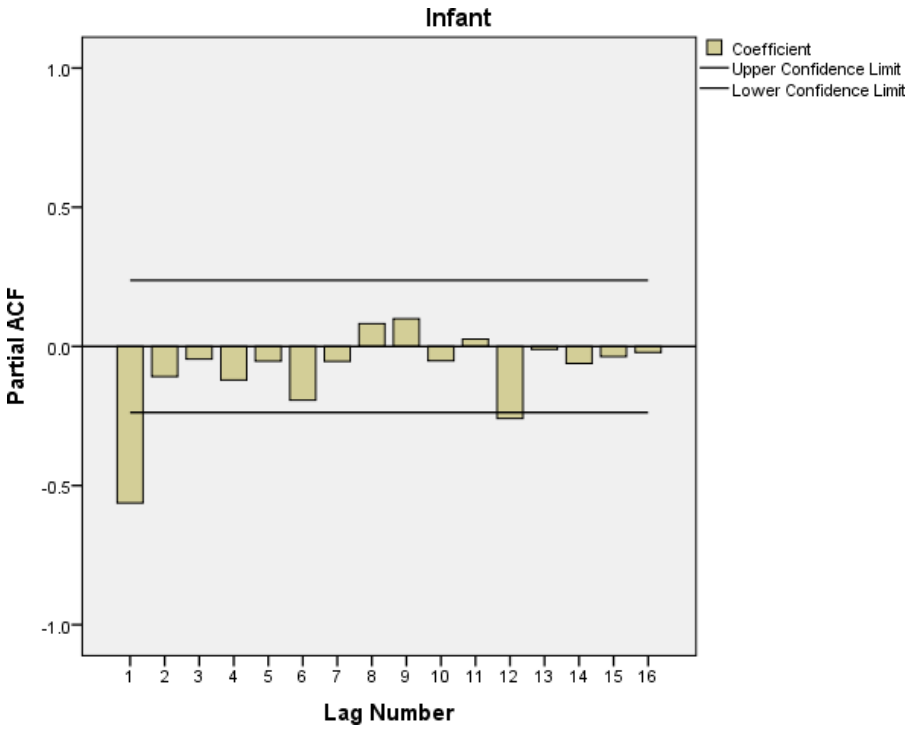

Figure 4. Partial autocorrelation function of infant birth in Pauri Division for selection of model

Table 3a. Model Summary for good fit of forecasting

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile
					5
Stationary R-squared	-1.002E-013	.	-1.002E-013	-1.002E-013	-1.002E-013
R-squared	-1.002E-013	.	-1.002E-013	-1.002E-013	-1.002E-013
RMSE	4.949	.	4.949	4.949	4.949
MAPE	97.718	.	97.718	97.718	97.718
MaxAPE	619.444	.	619.444	619.444	619.444
MAE	4.150	.	4.150	4.150	4.150
MaxAE	13.806	.	13.806	13.806	13.806
Normalized BIC	3.258	.	3.258	3.258	3.258

SE= Std. Error

Table 3b. Model Summary for good fit of forecasting

Fit Statistic	Percentile				
	10	25	50	75	90
Stationary R-squared	-1.002E-013	-1.002E-013	-1.002E-013	-1.002E-013	-1.002E-013
R-squared	-1.002E-013	-1.002E-013	-1.002E-013	-1.002E-013	-1.002E-013
RMSE	4.949	4.949	4.949	4.949	4.949
MAPE	97.718	97.718	97.718	97.718	97.718
MaxAPE	619.444	619.444	619.444	619.444	619.444
MAE	4.150	4.150	4.150	4.150	4.150
MaxAE	13.806	13.806	13.806	13.806	13.806
Normalized BIC	3.258	3.258	3.258	3.258	3.258

Table 3c. Model Summary for good fit of forecasting

Fit Statistic	Percentile
	95
Stationary R-squared	-1.002E-013
R-squared	-1.002E-013
RMSE	4.949
MAPE	97.718
MaxAPE	619.444
MAE	4.150
MaxAE	13.806
Normalized BIC	3.258

Table 4a- Residual ACF Summary considering mean, minimum, maximum and percentile for trend identification

Lag	Mean	SE	Minimum	Maximum	Percentile			
					5	10	25	50
Lag 1	.484	.	.484	.484	.484	.484	.484	.484
Lag 2	.515	.	.515	.515	.515	.515	.515	.515
Lag 3	.310	.	.310	.310	.310	.310	.310	.310
Lag 4	.261	.	.261	.261	.261	.261	.261	.261
Lag 5	.212	.	.212	.212	.212	.212	.212	.212
Lag 6	.107	.	.107	.107	.107	.107	.107	.107
Lag 7	.199	.	.199	.199	.199	.199	.199	.199
Lag 8	.156	.	.156	.156	.156	.156	.156	.156
Lag 9	.137	.	.137	.137	.137	.137	.137	.137
Lag 10	.016	.	.016	.016	.016	.016	.016	.016
Lag 11	.012	.	.012	.012	.012	.012	.012	.012
Lag 12	-.106	.	-.106	-.106	-.106	-.106	-.106	-.106
Lag 13	.016	.	.016	.016	.016	.016	.016	.016
Lag 14	-.093	.	-.093	-.093	-.093	-.093	-.093	-.093
Lag 15	.013	.	.013	.013	.013	.013	.013	.013
Lag 16	-.037	.	-.037	-.037	-.037	-.037	-.037	-.037
Lag 17	-.078	.	-.078	-.078	-.078	-.078	-.078	-.078
Lag 18	-.191	.	-.191	-.191	-.191	-.191	-.191	-.191
Lag 19	-.255	.	-.255	-.255	-.255	-.255	-.255	-.255
Lag 20	-.256	.	-.256	-.256	-.256	-.256	-.256	-.256
Lag 21	-.357	.	-.357	-.357	-.357	-.357	-.357	-.357
Lag 22	-.304	.	-.304	-.304	-.304	-.304	-.304	-.304
Lag 23	-.297	.	-.297	-.297	-.297	-.297	-.297	-.297
Lag 24	-.265	.	-.265	-.265	-.265	-.265	-.265	-.265

Table 4b- Residual ACF percentile for trend identification

Lag	Percentile		
	75	90	95
Lag 1	.484	.484	.484
Lag 2	.515	.515	.515
Lag 3	.310	.310	.310
Lag 4	.261	.261	.261
Lag 5	.212	.212	.212
Lag 6	.107	.107	.107
Lag 7	.199	.199	.199
Lag 8	.156	.156	.156
Lag 9	.137	.137	.137
Lag 10	.016	.016	.016
Lag 11	.012	.012	.012
Lag 12	-.106	-.106	-.106
Lag 13	.016	.016	.016
Lag 14	-.093	-.093	-.093
Lag 15	.013	.013	.013
Lag 16	-.037	-.037	-.037
Lag 17	-.078	-.078	-.078
Lag 18	-.191	-.191	-.191
Lag 19	-.255	-.255	-.255
Lag 20	-.256	-.256	-.256
Lag 21	-.357	-.357	-.357
Lag 22	-.304	-.304	-.304
Lag 23	-.297	-.297	-.297
Lag 24	-.265	-.265	-.265

Table 5a- Residual PACF Summary considering mean, minimum, maximum and percentile for trend identification

Lag	Mean	SE	Minimum	Maximum	Percentile			
					5	10	25	50
Lag 1	.484	.	.484	.484	.484	.484	.484	.484
Lag 2	.367	.	.367	.367	.367	.367	.367	.367
Lag 3	-.037	.	-.037	-.037	-.037	-.037	-.037	-.037
Lag 4	-.023	.	-.023	-.023	-.023	-.023	-.023	-.023
Lag 5	.053	.	.053	.053	.053	.053	.053	.053
Lag 6	-.079	.	-.079	-.079	-.079	-.079	-.079	-.079
Lag 7	.150	.	.150	.150	.150	.150	.150	.150
Lag 8	.075	.	.075	.075	.075	.075	.075	.075
Lag 9	-.066	.	-.066	-.066	-.066	-.066	-.066	-.066
Lag 10	-.165	.	-.165	-.165	-.165	-.165	-.165	-.165
Lag 11	-.012	.	-.012	-.012	-.012	-.012	-.012	-.012
Lag 12	-.123	.	-.123	-.123	-.123	-.123	-.123	-.123
Lag 13	.179	.	.179	.179	.179	.179	.179	.179
Lag 14	-.050	.	-.050	-.050	-.050	-.050	-.050	-.050

Lag 15	.025	.	.025	.025	.025	.025	.025	.025
Lag 16	-.047	.	-.047	-.047	-.047	-.047	-.047	-.047
Lag 17	-.106	.	-.106	-.106	-.106	-.106	-.106	-.106
Lag 18	-.225	.	-.225	-.225	-.225	-.225	-.225	-.225
Lag 19	-.037	.	-.037	-.037	-.037	-.037	-.037	-.037
Lag 20	-.030	.	-.030	-.030	-.030	-.030	-.030	-.030
Lag 21	-.155	.	-.155	-.155	-.155	-.155	-.155	-.155
Lag 22	-.063	.	-.063	-.063	-.063	-.063	-.063	-.063
Lag 23	.021	.	.021	.021	.021	.021	.021	.021
Lag 24	-.114	.	-.114	-.114	-.114	-.114	-.114	-.114

Table 5b- Residual PACF percentile for trend identification

Lag	Percentile		
	75	**90**	**95**
Lag 1	.484	.484	.484
Lag 2	.367	.367	.367
Lag 3	-.037	-.037	-.037
Lag 4	-.023	-.023	-.023
Lag 5	.053	.053	.053
Lag 6	-.079	-.079	-.079
Lag 7	.150	.150	.150
Lag 8	.075	.075	.075
Lag 9	-.066	-.066	-.066
Lag 10	-.165	-.165	-.165
Lag 11	-.012	-.012	-.012
Lag 12	-.123	-.123	-.123
Lag 13	.179	.179	.179
Lag 14	-.050	-.050	-.050
Lag 15	.025	.025	.025
Lag 16	-.047	-.047	-.047
Lag 17	-.106	-.106	-.106
Lag 18	-.225	-.225	-.225
Lag 19	-.037	-.037	-.037
Lag 20	-.030	-.030	-.030
Lag 21	-.155	-.155	-.155
Lag 22	-.063	-.063	-.063
Lag 23	.021	.021	.021
Lag 24	-.114	-.114	-.114

Table 6- Model Fit Statistics for Ljung- Box test

Model	Number of Predictors	Model Fit statistics			Ljung-Box Q(18)	
		Stationary R-squared	R-squared	Normalized BIC	Statistics	DF
Infant-Model_1	0	-1.002E-013	-1.002E-013	3.258	68.077	18

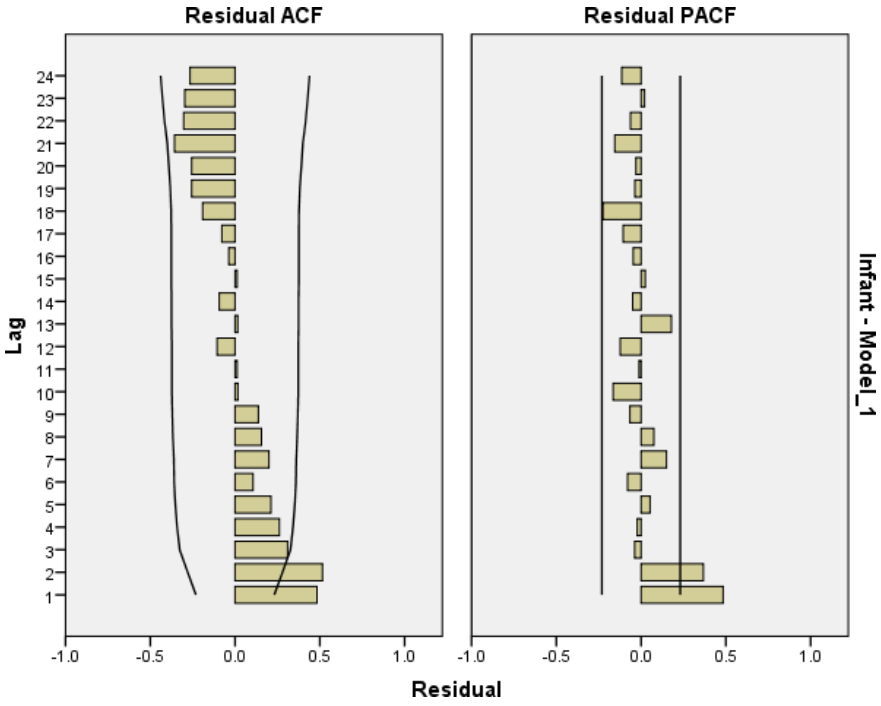

Figure 5. Comparative analysis of ACF and PACF of infant forecasting of Pauri Division

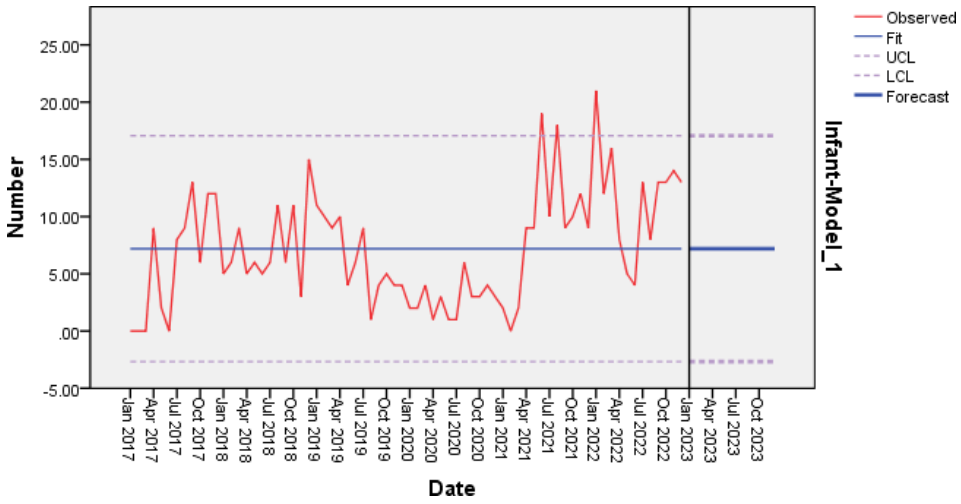

Figure 6. Infant Forecasting for 2023 based on time series data

DISCUSSION

According to the available literature, forecasting for a specific period is the fourth and final step (Kasyoki, 2015) in the ARIMA model. There are two kinds of forecasts: sample period (the period for which data are available) and post-sample period (actual forecasting period) forecast. The sample period forecast gives confidence to the model as the gap between actual data and forecast data could be discovered. This is an indirect indicator of the accuracy of the model. The post-sample forecast generates a genuine forecast for a specific period for use in

planning and intervention. The selected ARIMA model is used in the current study (2, 1, 2), and the IMR has been forecasted for both sample and post-sample periods.

The gap between the actual data (red line) and forecast data (blue line) is visible but negligible and gives confidence to the selected model for further forecasts. The post-sample forecast has been plotted with 95% confidence interval. As the period of forecast increases, the confidence range becomes broader. In the current study, the post-sample forecast for 2023 and the change in the confidence interval range becomes noticeable. There is a decreasing trend of IMR in the period of forecast (2023).

It is obvious that many factors directly or indirectly impact on infant mortality, thereby reflecting in the rates. An analysis of these annual rates indirectly takes into account the factors impacting on them. One of the limitations of the current study tool is that direct consideration of factors is not possible in the use of the ARIMA model for analysis.

CONCLUSION

The ARIMA model is the most commonly used method of forecast because of its simplicity and its usefulness for any time series data, the only requirement being the long-time series data. The forecast is not 100% perfect; however, if the current data are available and the best model has been selected, the accuracy of forecasting of any variable is improved. In the current study, time series IMR data had been used to forecast the IMR. The predicted data show that IMR will decline. When the data for another year (2023) are available, the model can be checked for validity, and the forecast can be done more accurately. The current study reveals the application of a statistical tool for forecasting an event to help in the proper planning of intervention. Similarly, other health events could be forecasted so that proper intervention can be planned at the appropriate time.

REFERENCES

Cha S (2017). The impact of the worldwide millennium development goals campaign on maternal and under-five child mortality reduction: 'Where did the worldwide campaign work most effectively?'. Glob Health Action.10(1):1267961.

Chao F, You D, Pedersen J, Hug L, Alkema L (2018). National and regional under-5 mortality rate by economic status for low-income and middle-income countries: a systematic assessment. Lancet Glob Health.6(5):e535–47.

Kasyoki A (2015). Simple steps for fitting ARIMA model to time series data for forecasting using R. *Int J Sci Res.*4(3):318–321.

Liu L, Chu Y, Oza S, Hogan D, Perin J, Bassani DG (2019). National, regional, and state-level all-cause and cause-specific under-5 mortality in India in 2000-15: a systematic analysis with implications for the sustainable development goals. Lancet Glob Health;7(6):e721–34.

Lomazzi M, Borisch B, Laaser U. (2014). The millennium development goals: expe- riences, achievements and what's next. Glob Health Action. 7:23695.

Mandal BN (2005). Forecasting Sugarcane Production in India with ARIMA Model. Statistics 1-13.

National Family Health Survey-4 (2015–16) *Ministry of Health and Family Welfare, Government of India*. Available from: http://www.rchiips.org/NFHS/NFHS-4Reports/India.pdf .

National Family Health Survey-5 (2019-20). *Ministry of Health and Family Welfare, Government of India*. Available from: https://main.mohfw.gov.in/sites/default/files/NFHS-5_Phase-II_0.pdf .

Rutstein Shea Oscar (1984). Infant and Child Mortality: levels, Trends and Demographic Differentials. Comparative Studies. CROSS-NATIONAL SUMMARIES NUMBER 43. WORLD FERTILITY SURVEY, Netherlands

Saikia N, Singh A, Jasilionis D, Ram F (2013). Explaining the rural–urban gap in infant mortality in India. Demogr Res. 29:473–506.

UNICEF (2022). Levels and Trends in Child Mortality. *UNICEF*. Available at https://data.unicef.org/resources/levels-and-trends-in-child-mortality/

Upadhyay R (2015). *Step-by-Step Graphic Guide to Forecasting through ARIMA Modeling using R – Manufacturing Case Study*. Available at http://ucanalytics.com/blogs/step-by-step-graphic-guide-to-forecasting-through-arima-modeling-in-r-manufacturing-case-study-example/