DR-2199
[1-9]

RESEARCH ARTICLE Asian Journal of Dairy and Food Research

Temporal Trends and Future Projections: A Deep Dive into India's Buffalo Milk Production Through Time Series Modelling

Vaisakh Venu[1], P.K. Anjitha[2], P.R. Vipin[3], E.R. Ramdas[2], R. Senthilkumar[2], B. Sreenath[1]

10.18805/ajdfr.DR-2199

ABSTRACT

Background: Buffalo milk production in India plays a significant role in the global dairy market, with a rich history deeply intertwined with the country's economy and culture. Over six decades, the dynamics of buffalo farming have been pivotal in shaping India's dairy landscape.

Methods: This paper delves into the subject by analysing a comprehensive time series dataset spanning six decades. The focus lies on understanding the economic and cultural significance of buffalo farming, particularly in relation to milk production. Four forecasting models-ARIMA, SES, Seasonal Naive and ETS-are employed to discern temporal patterns in buffalo milk production.

Result: The study reveals that the ARIMA and ETS models outperform SES and Seasonal Naive models in capturing and elucidating data behaviour. Their superior performance underscores their efficacy in predicting buffalo milk production trends accurately. These findings offer valuable insights for policymakers and stakeholders aiming to optimize buffalo milk production and foster long-term growth in India's dairy sector.

Key words: ARIMA, Buffalo milk production, Dairy industry, ETS, Forecasting, Seasonal naïve, SES, Time series modelling.

INTRODUCTION

India is the world's top milk producer, accounting for 24 percent of the world's milk production in 2021-2022. Over the course of the last eight years, or between 2013-14 and 2021-22, India's milk output has climbed by 61%, reaching 221.1 million tonnes in 2021-22. Compared to the 2020-21 financial year, there has been a 5.29% increase in milk production. Rajasthan (15.05%), Uttar Pradesh (14.93%), Madhya Pradesh (8.6%), Gujarat (7.56%) and Andhra Pradesh (6.97%) are the top 5 states that produce milk. Collectively, they account for 53.11% of the nation's milk production. In 2022-2023 India exported 67,572.99 MT of dairy products to the world, valued at $284.65 million.

India is recognized as a rich source of buffalo germplasm resources, hosting all acknowledged high-producing breeds of this species. As the Indian dairy industry undergoes transformative changes, the buffalo sector plays a crucial role in overall milk production, contributing 66.3% of the total global volume of buffalo milk. With Indian buffaloes accounting for nearly 56% of the world's buffalo population, about 25% of global buffalo meat is produced in India. Buffalo farming has evolved into a livelihood and resource-generating enterprise for diverse segments of our farmers. It plays a major role in poverty alleviation and commercial buffalo enterprises now provide employment opportunities to rural communities.

Buffalo farming holds significant importance in our rural economy, contributing not only to milk production but also being valued for meat and draught purposes. Buffalo milk, known for its higher fat content of 7-7.5%, nearly double that of cows, commands a higher market price. Milk obtained from the mammary glands of female buffaloes, known as buffalo milk, is a highly nutritious dairy item that holds significant cultural and economic importance. Differing in composition, taste, and nutritional value from cow's milk, buffalo milk boasts higher levels of fat, protein, and minerals. Recognized for its creamy texture and robust flavour, it typically contains 7% to 8% fat, making it ideal for producing butter, ghee and other dairy products-significantly higher than cow's milk. Rich in essential nutrients like protein, calcium, phosphorus, vitamin A and B-complex vitamins, buffalo milk's elevated fat content also contributes to its energy density (Thakore and Jain, 2018).

Buffalo milk extensively utilized in creating various dairy products such as paneer, yogurt and traditional sweets,

[1]Department of Basic Engineering and Applied Sciences, Kelappaji College of Agricultural Engineering and Technology, Kerala Agricultural University, Tavanur-679 573, Kerala, India.
[2]Department of Food Processing and Engineering, Kelappaji College of Agricultural Engineering and Technology, Kerala Agricultural University, Tavanur-679 573, Kerala, India.
[3]Department of Farm Machinery and Power Engineering, Kelappaji College of Agricultural Engineering and Technology, Kerala Agricultural University, Tavanur-679 573, Kerala, India.

Corresponding Author: B. Sreenath, Department of Basic Engineering and Applied Sciences, Kelappaji College of Agricultural Engineering and Technology, Kerala Agricultural University, Tavanur-679 573, Kerala, India. Email: mathskcaet@gmail.com

How to cite this article: Venu, V., Anjitha, P.K., Vipin, P.R., Ramdas, E.R., Senthilkumar, R. and Sreenath, B. (2024). Temporal Trends and Future Projections: A Deep Dive into India's Buffalo Milk Production Through Time Series Modelling. Asian Journal of Dairy and Food Research. DOI: 10.18805/ajdfr.DR-2199.

Submitted: 24-01-2024 **Accepted:** 08-05-2024 **Online:** 23-07-2024

buffalo milk imparts a luxurious and creamy quality to these items. In regions where buffalo farming is prevalent, buffalo milk plays a pivotal role in the dairy industry, significantly contributing to overall milk production and serving as a livelihood source for many farmers. Buffalo milk indeed often contains a higher fat content compared to cow milk. This higher fat content makes buffalo milk desirable for various dairy products. The pricing of milk is commonly influenced by its fat content, among other factors. In many cases, higher fat content can lead to higher profit margins for dairy producers, as products derived from such milk, like butter and cream, are often priced at a premium. In some regions, buffalo milk plays a significant role in the dairy industry. The fact that buffalo milk accounts for the largest share (55%) of the total milk produced in the country indicates its importance in the overall dairy market. This could be due to factors such as the adaptability of buffaloes to different climates, their ability to thrive on diverse feed sources, or specific preferences for buffalo milk and its products in the local culture (Singh, 2015). Notably, countries like India are major global contributors to buffalo milk production, with the Indian subcontinent home to high-yielding buffalo breeds that substantially augment the world's buffalo milk supply. To summarize, buffalo milk stands as a valuable dairy product with distinct attributes, making it versatile for a range of culinary applications. Its higher fat content, nutritional richness and economic significance underpin its widespread consumption and utilization across different cultures worldwide.

This study explores the temporal patterns of buffalo milk production in India using various forecasting models, including ARIMA, SES, Seasonal Naive and ETS. The aim is to predict future production trends and understand the underlying mechanisms shaping buffalo milk production dynamics. The study delves into the economic and cultural significance of buffalo farming, offering nuanced insights into the multifaceted landscape of India's dairy industry. By comparing and contrasting the performance of different forecasting models, the study aims to identify the most effective approaches for modelling buffalo milk production, providing actionable insights for policymakers and stakeholders to optimize production strategies and contribute to the sustainable growth of India's dairy sector.

MATERIALS AND METHODS

Data collection and preparation

The study utilized an extensive time series dataset comprising annual measurements of raw buffalo milk production from 1961 to 2021. The dataset was meticulously curated from authoritative agricultural records and sources, ensuring its integrity and reliability. Preprocessing steps were undertaken to address missing values, ensure chronological order and aggregate the annual data into consistent time intervals.

Forecasting models and equations

Various time series models are available in literature which is to be used based on the characteristics of the data. The ability of different models to forecast the time series values is assessed by using forecast evaluation measures. Four distinct forecasting models were considered and applied to capture the temporal patterns in the milk production time series data.

Auto regressive integrated moving average (ARIMA)

The Box-Jenkins ARIMA (p, d, q) model, introduced by Box and Jenkins Box *et al*. (2015), stands as a widely employed technique for constructing univariate time series forecasting models. Developed by George Box and Gwilym Jenkins in the 1970s, the ARIMA model provides a mathematical framework for predicting processes. The Box-Jenkins modelling approach encompasses the steps of identifying a suitable ARIMA process, fitting it to the available data, and subsequently employing the established model for predictive purposes. The ARIMA model represents the time series as a function of its own past values, integrating autoregressive (AR), differencing (I), and moving average (MA) components. The general ARIMA (p, d, q) equation is expressed as:

$$Y_t = \varnothing_1 Y_{t-1} + \varnothing_2 Y_{t-2} + ... + \varnothing_p Y_{t-p} + \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Where

Y_t = Milk production,

ε_t = Independently and normally distributed with zero mean and constant variance for t=1, 2, ... , n; and $\varnothing p$ and θq are also estimated (Venu *et al*., 2023: Deshmukh Surendra Sagar and Paramasivam, 2016: Mahajan Sunali *et al*., 2020).

Exponential smoothing state space model (ETS)

The Exponential Smoothing State Space Model (ETS) is a forecasting method that combines exponential smoothing with a state space approach (Mphale and Narasimhan, 2022). ETS models are particularly useful for time series data that exhibit trends and seasonality. Developed to provide a flexible and adaptive framework for time series forecasting, ETS models have gained popularity in both academia and industry. ETS models provide a flexible framework that can handle a wide range of time series patterns, making them suitable for forecasting in diverse applications. The implementation of ETS models is available in various statistical software packages, and they can be adapted to incorporate exogenous variables for improved forecasting accuracy in ETSX models. The flexibility and adaptability of ETS make it a valuable tool for analysts and practitioners dealing with time series forecasting. The basic ETS (A, A, A) model (where "A" stands for additive) can be represented by the following equations:

$$\text{Level equation: } l_t = \propto . y_t + (1 - \propto) . (l_{t-1} + b_{t-1})$$

$$\text{Trend equation: } b_t = \beta . (l_t - l_{t-1}) + (1 - \beta) . b_{t-1})$$

$$\text{Seasonal equation: } s_t = \gamma . (y_t - l_{t-1} - b_{t-1}) + (1 - \gamma) . s_{t-s})$$

$$\text{Forecast equation: } \hat{y}_{t+h} = l_t + h . b_t + s_{t+h-s}$$

Where

y_t= Observed value at time, *t*.

$\hat{y}_{t+h}$ = Forecast for time t+h.

Where,

h = Forecasting horizon,

l_t = Level at time t.

b_t = Trend at time t.

s_t = Seasonal component at time t, α, β.

γ = Smoothing parameters for level, trend and seasonality, respectively.

s = Length of the seasonal cycle.

Simple exponential smoothing (SES)

Simple Exponential Smoothing (SES) is a basic and widely used method for time series forecasting (Ostertagova and Ostertag, 2012). It belongs to the family of exponential smoothing methods, which are designed to capture patterns and trends in time series data. SES is particularly useful when the data does not exhibit complex patterns or seasonality. Simple Exponential Smoothing is a good starting point for time series forecasting, especially when the data has a relatively stable trend and does not exhibit seasonality. It is easy to implement and computationally efficient, making it a practical choice for simple forecasting tasks. However, the appropriateness of SES depends on the specific characteristics of the time series being analysed. The SES equation is as follows:

$$\hat{y}_{t+1} = \propto . y_t + (1 - \propto) . \hat{y}_t$$

Where

$\hat{y}_{t+1}$ = Forecast for the next time period (t+1).

y_t = Actual observation at time *t*.

$\hat{y}_t$ = Forecast for time *t*.

α = Smoothing parameter.

Also known as the smoothing coefficient, and it takes values between 0 and 1. α determines the weight given to the most recent observation in the smoothing process.

Seasonal naive model

The seasonal naive model is a simple and intuitive time series forecasting method that relies on the observation that many time series exhibit seasonality, where patterns repeat over fixed intervals (Pala and Atici, 2019). The Seasonal naive model predicts future values based on the most recent observation from the same season in previous periods. While it may not capture all nuances in the data, it serves as a baseline for comparison and is particularly useful when the time series exhibits clear and consistent seasonality. If we denote the time series data as y_t.

Where,

t = Time index and the seasonal period as *s*, then the equation for the seasonal naive method is given by:

$$\hat{y}_t = y_{t-1}$$

Where,

$\hat{y}_t$ = Forecast for time t.

y_{t-s} = Observation from the same season in the previous year, as *s* is the seasonal period.

The Seasonal Naive model assumes the next value to be identical to the last observed value at the same seasonal point.

Model evaluation and fitness metrics

Each model was trained on the designated training set and utilized to generate forecasts for the test set. The performance of these models was assessed using common forecast accuracy metrics:

Mean error (ME)

Mean error is the average of all the errors in a set of observations. That is the arithmetic mean of the differences between predicted values and actual values. ME is given by:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(P_i - Q_i)$$

n = Number of data points.

P_i = Predicted value for the i[th] data point.

Q_i = Actual value of the i[th] data point.

If the mean error is close to zero, it suggests that the observations are accurate. That is, a non-zero ME indicates a bias in the observations. However, sometimes ME might not give the correct picture of accuracy, especially if there are both positive and negative errors that cancel out.

Mean absolute error (MAE)

Mean Absolute Error (MAE) (Willmott, 2005: Chai *et al*., 2014) is a metric used to evaluate the performance of a regression model. It measures the average absolute difference between the actual value and predicted values. The mean absolute error is calculated by:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|P_i - Q_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i|$$

Where,

n = Number of data points.

P_i= Predicted value for the i[th] data point.

Q_i= Actual value of the i[th] data point,

$$e_i = P_i - Q_i$$

Root mean squared error (RMSE)

Root Mean Squared Error (Willmott and Matsuura, 2005; Chai and Draxler, 2014) is commonly used to evaluate the performance of a predictive model, particularly in regression analysis. It is the quadratic mean of the differences between observed values and actual values, giving more weight to larger errors. RMSE is given by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - Q_i)^2} = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}(e_i)^2\right)}$$

n = Number of data points.

P_i = Predicted value for the i[th] data point.

Q_i = Actual value of the i[th] data point.

Mean absolute percentage error (MAPE)

MAPE (Moreno, 2013; Goodwin *et al*., 1999) is used to assess the accuracy of a forecasting or prediction model, especially in the context of time series analysis. MAPE is expressed as a percentage and measures the average absolute percentage difference between the predicted and actual values. If n is the number of samples.

P_i = Predicted value for the i^{th} data point.

Q_i= Actual value of the i^{th} data point, then MAPE is given by;

$$MAPE = \frac{1}{n}\sum_{i=1} \frac{|P_i - Q_i|}{Q_i}$$

MAPE provides a clear and intuitive way to assess the accuracy of forecasting models, especially in situations where understanding errors in percentage terms is important. Lower MAPE values indicate better accuracy, with a value of 0 indicating a perfect match between predicted and observed values. However, MAPE is sensitive to zero values in the actual data, which can result in undefined or infinite percentages. Despite such limitations, MAPE is probably the most widely used forecasting accuracy measurement. These metrics were instrumental in evaluating the accuracy and reliability of the models in predicting future milk production values.

Residual analysis

A thorough residual analysis was performed on the ARIMA model to confirm its dependability and assumptions (Revathi *et al.*, 2023: Jaiswal Priyanka and Bhattacharjee Mahua, 2022). The study included visual inspections, Q-Q plots, histograms and the Autocorrelation Function (ACF) to confirm the distribution's normality and look for any temporal trends or systematic errors. The residuals' absence of evident trends or correlations supported the model's future predictions of oilseed yield accuracy. This research confirmed that the independence, normality, and constant variance assumptions were met, which improved the model's prediction reliability.

Ljung-box test

The Ljung-Box test is a statistical test used in time series analysis to assess whether the residuals from a time series model exhibit significant autocorrelation at different lags Dare (2022). The test is particularly useful for evaluating the adequacy of the chosen model by examining whether there are any remaining patterns in the residuals that the model has not captured. The Ljung-Box test statistic is based on the sum of the squares of the sample autocorrelations of the residuals at different lags. It is calculated as follows:

$$Q = n\,(n + 2)\sum_{k=1}^{h} \frac{\hat{P}k^2}{n - k}$$

Where,

Q = Test statistic, n is the sample size.

$\hat{p}k$ = Sample autocorrelation at lag k.

h = Maximum lag considered.

RESULTS AND DISCUSSION

The study utilized an extensive time series dataset comprising annual measurements of raw buffalo milk production from 1961 to 2021 taken from FAOSTAT (Fig 1). The stationarity of buffalo milk production time series data was assessed using the Augmented Dickey-Fuller (ADF) test. The test fitted a regression model with lags and a time trend to the differenced series. The test-statistic value of 0.7954, which is very much greater than critical, indicates that the data is non-stationary. Some lagged differences had statistically significant regression coefficients, indicating the persistence of temporal patterns. Non-stationarity implications will be considered

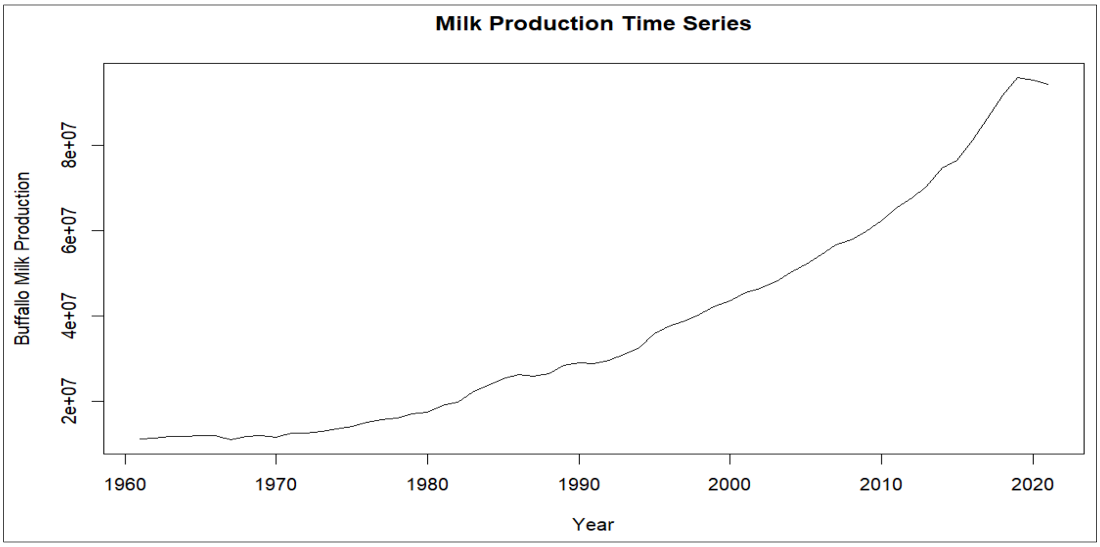

Fig 1: Buffalo milk production in India from 1961 to 2021.

for modelling.

ARIMA

Using R software, the best ARIMA model for time series data was found using the forecast package. As we already found that the data is not stationery from ADF test differencing technique is applied. For d=2 the data become stationery. ARIMA (1, 2, 1) was found to be the best model with the lowest AIC value among the several ARIMA models considered. The model equation is as follows:

$$(1 - 0.5423B)(1 - B)^2 Y_t = \varepsilon_t - 0.9117\ \varepsilon_{t-1}$$

Where,

Yt = Differenced and stationary time series,

εt = White noise error term at time t, and the autoregressive and moving average coefficients are given by 0.5423 and -0.9117, respectively.

The autoregressive coefficient of 0.5423 indicates a positive correlation between current buffalo milk production and previous values, implying that production levels will persist over time. In contrast, the moving average coefficient of -0.9117 indicates that the model rapidly adjusts to deviations from the predicted values, reflecting a strong tendency for fluctuations in buffalo milk production to revert to the long-term average. The Ljung-Box test was used to examine the residuals of the ARIMA (1, 2, 1) model for time series data on buffalo milk production. The test produced a high p-value of 0.6995, indicating that there was no significant evidence of autocorrelation at lags of up to ten. As a result, the ARIMA (1, 2, 1) model appears adequate for capturing and explaining temporal patterns in buffalo milk production time series data. The residual plots of ARIMA (1, 2, 1) is given in Fig 2, which also suggest the same conclusion as above.

Simple exponential smoothing

The Simple Exponential Smoothing (SES) model, when applied to the buffalo milk production data, reveals a notably high smoothing parameter (α = 0.9999\). This parameter signifies that the model heavily relies on recent observations to generate forecasts, with minimal weight attributed to older data points. In essence, the model prioritizes the most recent data, implying a strong preference for short-term trends over long-term patterns in buffalo milk production. The initial state (l = 11,809,985.2045\)) represents the starting level of the smoothed series, indicating that the SES model begins its forecast from this baseline value. The standard deviation of the residuals (σ=1,988,051) provides a measure of the variability of the observed data around the smoothed values, indicating the level of uncertainty inherent in the model's predictions. Fig 3 represents the residual plots of the SES model. Evaluation metrics such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) suggest that the SES model provides a reasonably good fit to the data, with lower values indicating better model performance. However, despite the relatively low AIC and BIC values (2024.055 and 2030.388 respectively), the error measures reveal a Mean Absolute Percentage Error (MAPE) of 4.07%, indicating that, on average, the model's forecasts deviate by approximately 4.07% from the actual values. Moreover, the Ljung-Box test yields a very low p-value (2.397e-12), indicating that the residuals from the SES model are not completely

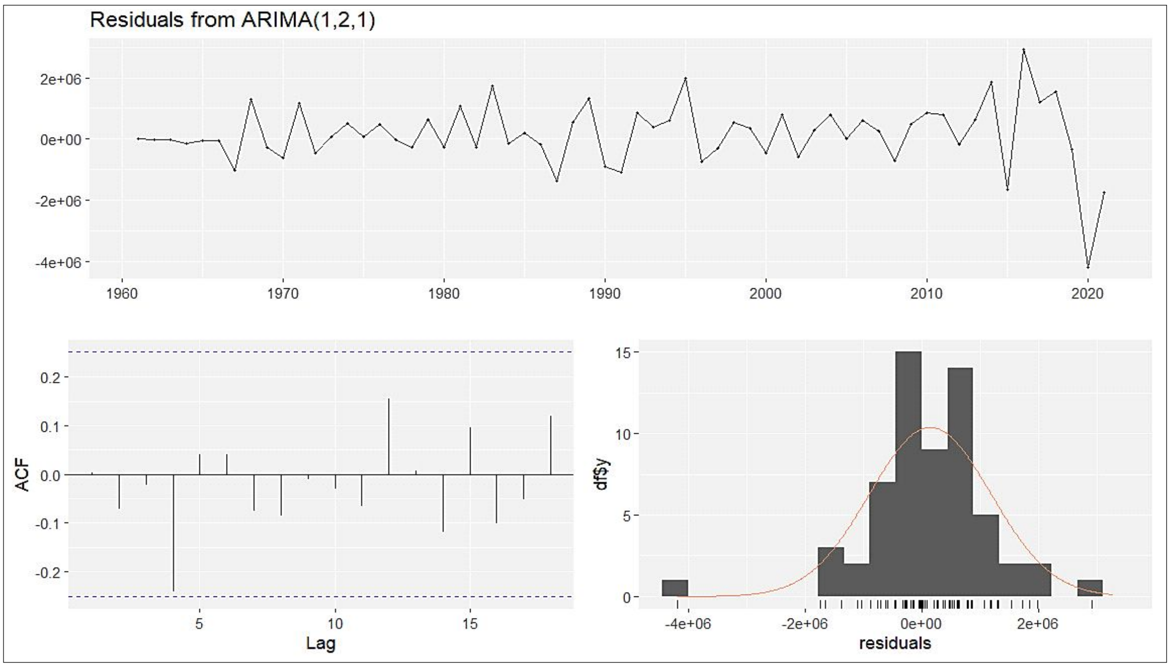

Fig 2: Residual plots of ARIMA (1,2,1) model.

free of autocorrelation, suggesting potential room for further model refinement or exploration of alternative modelling techniques to enhance predictive accuracy.

Seasonal naïve model

A seasonal model is developed for the milk production data and tested for model fitness. The Ljung-Box test provided a p value (1.939e-10) significantly smaller than the conventional significant levels which implies high autocorrelation among residuals in the model. Fig 4 depicts the residual plots which in turn agree with the above conclusion. As a result, like SES model, Seasonal Naïve Model also fails to describe the temporal patterns in the data.

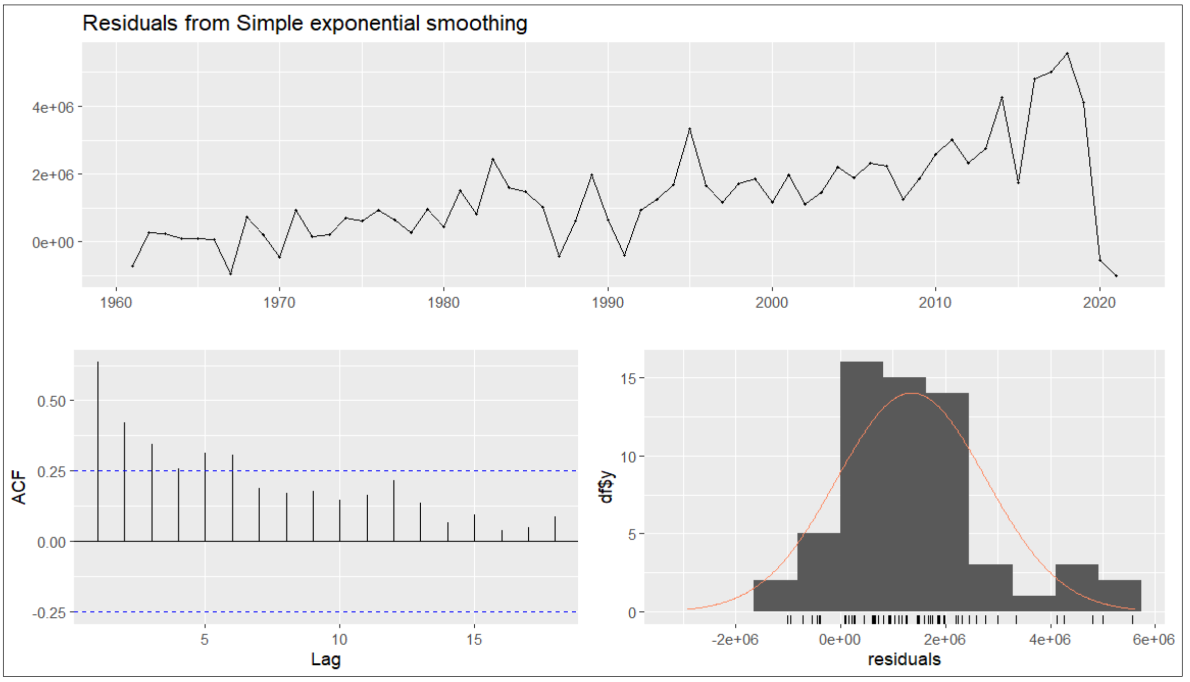

Fig 3: Residual plots of SES model.

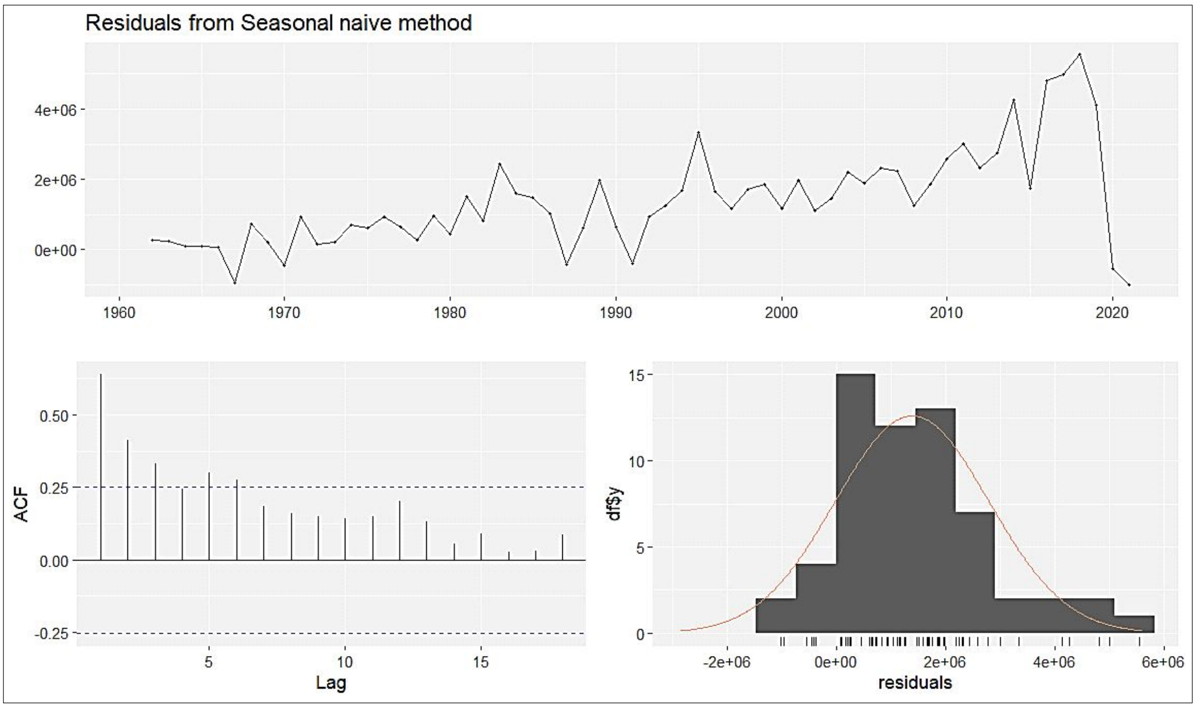

Fig 4: Residual plots of seasonal naïve method.

ETS model

ETS (A, Ad, N) model was developed for the data using R software and model equations are as follows:

$$Y_{t-1} = 11367399.6047 + 284622.4706 + \varepsilon\_t$$
$$l_t = 0.9999\ Y_t + 0.0001\ l_{t-1} + 0.8843\ \ b_{t-1}$$
$$b_t = 0.8058\ (l_t - l_{t-1}) + 0.1942 * 0.8843\ b_{t-1}$$

The model exhibits significant forecasting implications due to its emphasis on recent observations and high α parameter (0.9999), indicating adaptability to short-term fluctuations, whereas the β parameter assumes a moderate value of 0.8058, indicating a balanced consideration of recent trends and historical patterns. The introduction of damping to the trend component, as denoted by a ϕ value

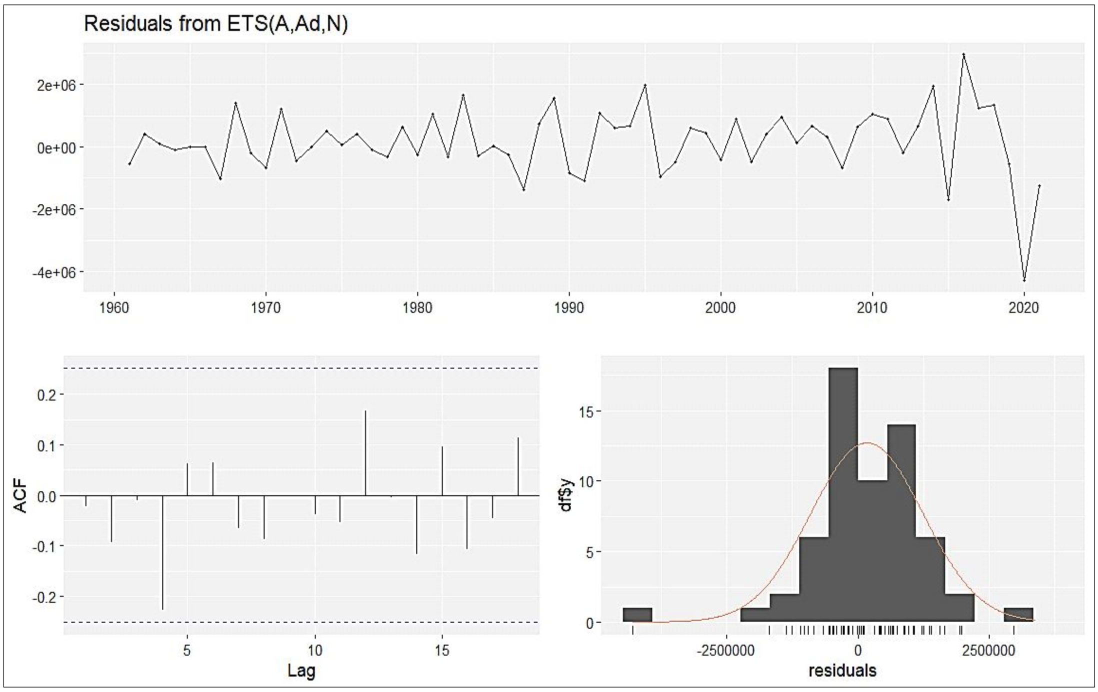

Fig 5: Residual plots of ETS model.

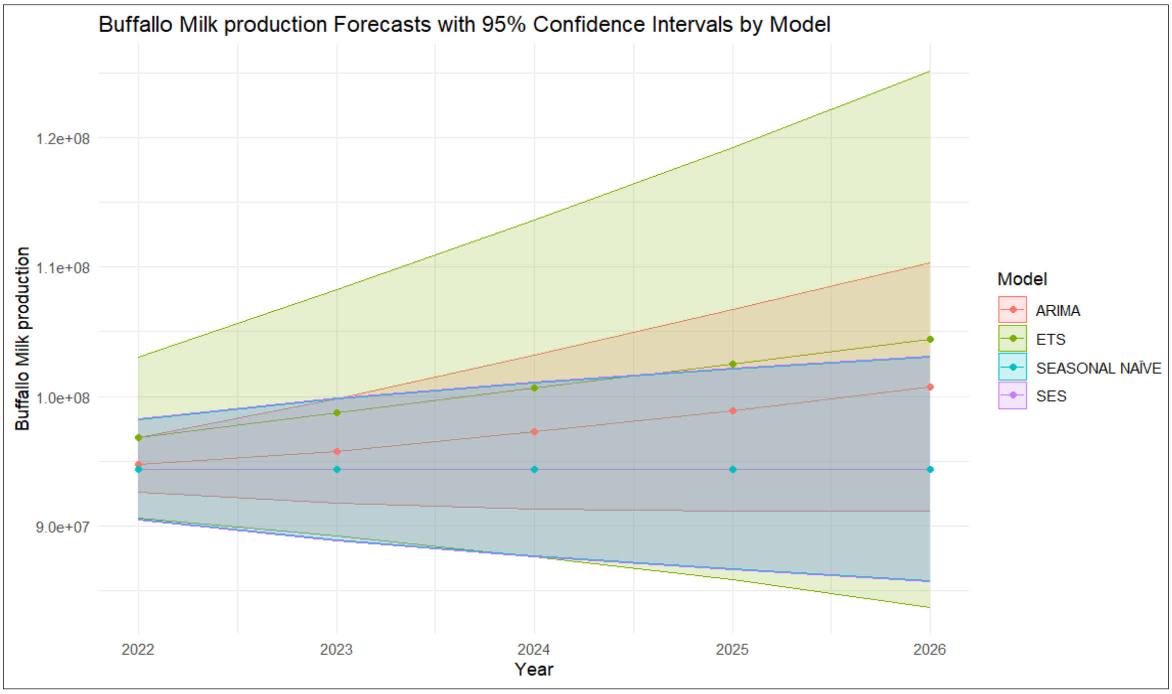

Fig 6: Forecast by different models with 95% confidence interval.

of 0.8843, suggests a controlled reduction in the amplitude of the trend over time. This nuanced approach to trend dynamics is particularly relevant in capturing both short-term variability and the broader trajectory of the time series. Fig 5 represents the residual plots of the ETS model. The fact that the ACF lags are clearly within the confidence interval lines indicates that there is no significant autocorrelation between the residuals. The Ljung-Box test produced a p value of 0.3443, which is greater than the significant levels and implies the same conclusion.

Model fitness metrics

Based on the residual analysis, we discovered that the ARIMA and ETS models explain the data's behaviour better than the SES and Seasonal naive models. We will now consider different model fitness matrices for these four models to validate these findings. Table 1 displays the ME, MAPE, MAE and RMSE values for all four models that we developed.

When the forecast accuracy of ARIMA, SES, Seasonal Naive and ETS models is compared, distinct patterns emerge. The ARIMA model performed well, with low Mean Absolute Percentage Error (MAPE) values on both the training (2.08%) and test (3.18%) sets, indicating its ability to capture underlying data patterns. While competitive, the SES model had higher MAPE values (4.12% on training and 12.26% on test), indicating potential limitations in handling more complex temporal dynamics. Seasonal Naive demonstrated competitive accuracy (4.07% on training and 8.64% on the test set), excelling in capturing seasonal patterns. The ETS model, which is like ARIMA, demonstrated strong performance (2.18% on the training set and 3.26% on the test set). Table 2 gives forecast of buffalo milk in India using SES, ARIMA, Seasonal Naïve and ETS models with 80% and 90% confidence interval and Fig 6 shows a diagrammatic representation.

Table 1: ME, MAPE, MAE and RMSE values for all four models.

Model	ME	MAE	RMSE	MAPE
ARIMA-Training set	182232.2	558123.7	786833.7	2.08%
ARIMA-Test set	1658351.9	2967820.2	3288984.6	3.18%
SES-Training set	1240408	1344112	1685114	4.12%
SES-Test set	11493470	11493470	12027516	12.26%
Seasonal naive-Training set	1275987	1355296	1697439	4.07%
Seasonal naive-Test set	7773110	7773110	8254515	8.64%
ETS-Training set	172309.8	567957.4	787098.5	2.18%
ETS-Test set	1693440.6	3033176.3	3350870.4	3.26%

Table 2: Forecast of buffalo milk production in India different models.

Model	Year	Forecast	Lo_80	Hi_80	Lo_95	Hi_95
SES	2022	94383792	91836003	96931582	90487285	98280300
	2023	94383792	90780854	97986730	88873574	99894011
	2024	94383792	89971186	98796399	87635293	101132291
	2025	94383792	89288596	99478989	86591362	102176223
	2026	94383792	88687218	100080367	85671634	103095951
ARIMA	2022	94721185	93336985	96105386	92604233	96838137
	2023	95788097	93140386	98435808	91738772	99837421
	2024	97250543	93358493	101142594	91298165	103202921
	2025	98927472	93826080	104028865	91125565	106729379
	2026	100720707	94441251	107000163	91117107	110324306
SEASONAL NAIVE	2022	94383692	91860266	96907117	90524445	98242938
	2023	94383692	90815029	97952355	88925893	99841491
	2024	94383692	90012990	98754393	87699281	101068103
	2025	94383692	89336840	99430543	86665199	102102185
	2026	94383692	88741140	100026243	85754154	103013229
ETS	2022	96808041	92738268	100877814	90583861	103032221
	2023	98716365	92498327	104934403	89206697	108226033
	2024	100624689	92133764	109115614	87638939	113610439
	2025	102533013	91606765	113459261	85822759	119243267
	2026	104441337	90907902	117974772	83743734	125138940

CONCLUSION

This study investigates the dynamics of buffalo milk production in India over six decades, focusing on data from 1961 to 2021. The data showed non-stationarity, prompting the use of different modelling techniques. Four forecasting models were used: ARIMA, SES, Seasonal Naïve and ETS. The ARIMA model was found to be most suitable, capturing both short-term fluctuations and long-term trends. The SES model, based on recent observations, showed competitive performance but higher Mean Absolute Percentage Error (MAPE) values. Seasonal Naïve and ETS models failed to capture temporal patterns. The ARIMA and ETS models demonstrated superior forecast accuracy, with lower MAPE values indicating higher forecast accuracy. The forecasts, with 80% and 90% confidence intervals, provide valuable insights for policymakers and stakeholders in optimizing production strategies and fostering long-term growth in the dairy industry. The study emphasizes the importance of robust forecasting models in strategic decision-making and advancing the sustainability and resilience of the dairy sector. Further research and refinement of modelling techniques are needed to address the complexities and uncertainties in dairy production dynamics.

Conflict of interest

The authors declare that they have no conflict of interest or competing interests.

REFERENCES

Box, G.E.P., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. (2015). Time Series Analysis: Forecasting and control. John Wiley and Sons.

Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis: Forecasting and Control, rev. ed. San Francisco: Holden-Day. 1976.

Chai, T. and Draxler, R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)-Arguments against avoiding RMSE in the literature. Geoscientific Model Development. 7(3): 1247-1250.

Dare, J., Patrick, A.O. and Oyewola, D.O. (2022). Comparison of stationarity on ljung box test statistics for forecasting. Earthline Journal of Mathematical Sciences. 8(2): 325-336.

Deshmukh Surendra Sagar, Paramasivam R. (2016). Forecasting of milk production in India with ARIMA andVAR time series models. Asian Journal of Dairy and Food Research. 35(1): 17-22. doi: 10.18805/ajdfr.v35i1.9246.

Goodwin, P. and Lawton, R. (1999). On the asymmetry of the symmetric MAPE. International Journal of Forecasting, 15(4): 405-408.

Jaiswal, P., Bhattacharjee, M. (2022). Understanding the potential of livestock market with special reference to the export of swine meat from India: A study of time-series analysis using ARIMA-based forecasting method. Asian Journal of Dairy and Food Research. 41(3): 293-297. doi: 10.18805/ajdfr.DR-1797.

Moreno, J.J.M., Pol, A.P., Abad, A.S. and Blasco, B.C. (2013). Using the R-MAPE index as a resistant measure of forecast accuracy. Psicothema. 25(4): 500-506.

Mahajan, S., Sharma, M., Gupta, A. (2020). ARIMA modelling for forecasting of rice production: A case study of India. Agricultural Science Digest. 40(4): 404-407. doi: 10.18805/ag.D-5029.

Mphale, O. and Narasimhan, V.L. (2022). Comparative Forecasts of Confirmed COVID-19 Cases in Botswana Using Box-Jenkin's ARIMA and Exponential Smoothing State-Space Models. In Recurrent Neural Networks. CRC Press. (pp. 355-381).

Revathi, N., Gourshetty, D., Haritha, V., Razak, N., Sugathan, A., Sankar, M., Thomas, A., Ragh, A., Dinkar, H.M. and Venu, V. (2023). From historical data to future predictions: Analysing and forecasting oilseed yield trends in India using timeseries methods. The Pharma Innovation Journal. SP-12(12): 1145-1157.

Ostertagova, E. and Ostertag, O. (2012). Forecasting using simple exponential smoothing method. Acta Electrotechnica et Informatica. 12(3): 62-66.

Pala, Z. and Atici, R. (2019). Forecasting sunspot time series using deep learning methods. Solar Physics. 294(5): 50. doi: 10.1007/s11207-019-1434-6.

Singh, D. (2015). Role of dairy farming in social and economic change of India. IOSR Journal of Agriculture and Veterinary Science. 8(2): 84-85.

Thakore, V. and Jain, N.K. (2018). Protein and fat examination from the raw milk of different mammalian species (Cow, buffalo, goat and human) with successive lactation days. The Pharma Innovation Journal. 7(10): 506-510.

Venu, V., Vipin, P.R. and Prajitha, N.K. (2023). A comparative analytical study of many regression model approaches, arima model and a hybrid model for forecasting area, production and productivity of coconut in Kerala, India. Current Journal of Applied Science and Technology. 42(40): 37-49.

Willmott, C.J. and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research. 30(1): 79-82.