



# Trend Analysis of Milk Production in India

Dipriya R. Lyngkhai<sup>1</sup>, S. Basanta Singh<sup>2</sup>, Ram Singh<sup>1</sup>, Hehlangki Tyngkan<sup>1</sup>

10.18805/ajdfr.DR-1789

## ABSTRACT

**Background:** Dairy sector is a thriving enterprise in Indian agriculture showing colossal growth responsible for placing the country at the top position worldwide in milk production. This sector has seen a rising growth; however, there seems to be demand supply gap over the next few years. Under this backdrop, the study has been taken with the twin objectives of studying the trend and growth rate of production and per capita availability of milk and to forecast the production of milk through ARIMA modelling.

**Methods:** The study is based on secondary data of milk production from 1989 to 2019. The growth was examined by estimating compound annual growth rate and the Autoregressive integrated moving average (ARIMA) methodology was applied for modeling and forecasting of milk production of India. Augmented dickey-fuller (ADF) test was used to determine the stationarity of the model after which it was used to forecast the future production.

**Result:** The CAGR of milk production was higher in comparison to per capita availability (4.34 and 2.71 per cent per annum respectively). The forecast from the fitted ARIMA model show that the milk production is expected to be 244.7 million tonnes in 2024.

**Key words:** Akaike information criteria (AIC), ARIMA, Augmented dickey-fuller (ADF) test, Forecast, Growth rate.

## INTRODUCTION

Dairy sector is a thriving enterprise in Indian agriculture showing colossal growth responsible for placing the country at the top position worldwide in milk production. India was ranked top in the world for milk production, with 196.18 million tonnes produced (FAO, 2019).

It constitutes for 21% of the world milk production (GoI, 2018). Milk contributes close to a third of the gross income of rural households and in the case of those without land, nearly half of their gross income. It also accounts for 67 per cent of the value of output from the livestock subsector in agriculture (GOI, 2019).

This sector has seen a rising growth with annual growth rate of milk production greater than 6 per cent per annum since the past five years *i.e.*, from 2014-15 to 2018-19 owing to a large dairy herd. The contribution of milk from buffaloes is the highest accounting for 49 per cent followed by cows and goats at 48 and 3 per cent respectively. The top ten milk producing states are Uttar Pradesh, Rajasthan, Madhya Pradesh, Andhra Pradesh, Gujarat, Punjab, Maharashtra, Haryana, Bihar and Tamil Nadu. Together they contribute 81.4 per cent to the total milk production in the country (GOI, 2019).

However, there seems to be demand supply gap over the next few years due to increase in population, rise in per capita income, changing lifestyle, food habits, export opportunities *etc.* The per capita availability of milk in India during 2017-18 was 375 gm/day and by 2023-24, it is estimated to increase to 592 gm/day (GOI, 2018).

Milk production will play a vital role in helping the farmers realize the mission of the Union Government with respect to doubling farmers' income and for making India self-reliant or *Atmanirbhar Bharat*. The government has taken impressive steps to improve the animal husbandry infrastructure. However, due to the dynamic nature of the

<sup>1</sup>School of Social Sciences, College of Post Graduate Studies in Agricultural Sciences, Central Agricultural University, Umiam-793 103, Meghalaya, India.

<sup>2</sup>Department of Agricultural Economics, Central Agricultural University, Manipur-795 004, Imphal, India.

**Corresponding Author:** Dipriya R. Lyngkhai, School of Social Sciences, College of Post Graduate Studies in Agricultural Sciences, Central Agricultural University, Umiam-793 103, Meghalaya, India. Email: dipriya216@gmail.com

**How to cite this article:** Lyngkhai, D.R., Singh, S.B., Singh, R. and Tyngkan, H. (2022). Trend Analysis of Milk Production in India. Asian Journal of Dairy and Food Research. 41(2): 183-187. DOI: 10.18805/ajdfr.DR-1789.

**Submitted:** 21-07-2021 **Accepted:** 05-01-2022 **Online:** 08-03-2022

society along with the ever changing economic and natural conditions; quantities and prices of milk are subject to variations. Hence, the need for accurate and reliable forecast is indispensable and cannot be exaggerated. It will not only be beneficial for farmers but also the policy makers in planning for future milk demand. Though farmers continue to make forecasts based on their past experiences, statistical analysis of historical data is believed to provide more accurate and precise predictions. A wide number of time series forecasting models are available, be that as it may the ARIMA (Auto regressive integrated moving average) model has emerged as one of the most popular models. Paul and Das (2010) used Box-Jenkins models to forecast fishery dynamics, Lohano and Soomro (2006) and Taye *et al.* (2020) conducted similar investigations using autoregressive models in Pakistan and Ethiopia respectively, Sankar and Vijayalakshmi (2017) also used the univariate ARIMA model for analyzing ghee production in Tamil Nadu. Forecasting of milk production has been a popular topic and will continue

to remain so due to its role in social, economic and nutritional security. ARIMA model has been used extensively by many researchers for forecasting the production of milk in India (Paul *et al.*, 2014, Deshmukh and Paramasivan, 2016, Kaur and Rakshit, 2018, Mishra *et al.*, 2020).

Under this backdrop, the study has been taken with the twin objectives which are.

- (1) To study the trend and growth rate of production and per capita availability of milk and
- (2) To forecast the production of milk through ARIMA modeling.

## MATERIALS AND METHODS

### Data

The present study is based completely on secondary data on milk production (Million tonnes) and per capita availability of milk(gram/day) which are collected from the Ministry of Fisheries, Animal Husbandry and Dairying (2019) for the years 1989-90 to 2018-19 *i.e.*, a period of 30 years. The R statistical software (4.0.2) package has been used for analyzing the data and prediction of the future values.

### Methodology

#### Compound annual growth rate

The performance of milk production and the growth rate of per capita availability of milk during the period of 1990-2019 are examined by estimating compound annual growth rate using the given formula:

$$Y_t = a b^t$$

Where,

$Y_t$  = Milk production/Per capita availability of milk.

$a$  = Constant.

$b$  = Regression coefficient.

$t$  = Time variable.

After log transformation and estimation of the above function as.

$$\ln Y_t = \ln a + t \ln b$$

Compound annual growth rate has been estimated as:

$$r = [\text{antilog}(\ln b) - 1] \times 100$$

### ARIMA

ARIMA popularly called Box-Jenkins (BJ) methodology is a statistical tool in which future values explained by past values of the variable itself and random error terms (Gujarati and Porter, 2009). ARIMA stands for Autoregressive Integrated Moving Average. There are two basic assumptions which need to be fulfilled before applying this model *viz.*, linearity and stationarity. A general ARIMA model is represented as ARIMA (p,d,q).

Where,

$p$  = Number of autoregressive terms.

$d$  = Number of times the series has to be differenced before it becomes stationary.

$q$  = Number of moving average terms.

A generalized ARIMA model is of the form.

$$\Phi(B) y_t = (1-B)^d \Psi(B) e_t$$

Where

$B$  = Backward shift operator such that  $By_t = y_{t-1}$ .

$y_t$  = Actual value at time  $t$ .

$e_t$  = Random error at time  $t$ .

$\Phi_i(i=1,2,\dots,p)$  and  $\Psi_j(j=1,2,\dots,q)$  are the model parameters

The method consists of four steps:

#### Identification

The tentative values of  $p$ ,  $d$  and  $q$  are worked out with the help of Augmented Dickey-Fuller test and correlograms of autocorrelation (ACF) and partial autocorrelation (PACF) functions.

#### Estimation

At this stage, the tentatively chosen parameters of the ARIMA model are estimated based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC) values.

$$AIC = T \log(\sigma^2) + 2(p+q+1)$$

$$BIC = T \log(\sigma^2) + (p+q+1) \log T$$

Where,

$T$  = Number of observations used for estimation of parameters.

$\sigma^2$  = Mean square error.

#### Diagnostic checking

For a model to be a reasonable fit to the data, the residuals from the fitted model are examined using the Shapiro-Wilk normality test, given by:

$$W = \frac{[\sum_{i=1}^n a_i x_{(i)}]^2}{(\sum_{i=1}^n x_i - \bar{x})^2}$$

Where,

$x_i$  = Ordered random sample values.

$a_i$  = Constants generated from the covariances, variances and means of a normally distributed sample.

#### Forecasting

In this stage, to evaluate the accuracy of the fitted model, the one step ahead forecasting for the years 2015-2019 is done after which the total milk production has been forecasted for a period of five years from 2020-2024. Having being satisfied with the performance of the selected model, the forecasting is done up to a period of five years in order to minimize the errors that may arise by increasing the forecasting period.

## RESULTS AND DISCUSSION

### Trend and growth rates in milk production and per capita availability of milk in India

The trend in milk production and per capita availability of milk in India for the past thirty years have been depicted in Fig 1. It is evident from the figure that milk production has increased steadily over the three decades which results in higher availability of milk and milk products for India's growing population. The per capita availability of milk showed a positive sustained growth from 171 gm/day in 1990 to 214 gm/day in 2000 to the current 394 gm/day (GOI, 2019). This can be attributed to the increase in herd size as well as the

higher productivity of milch animals due to breeding programmes as well as continuous selection of productive animals over the years.

The compound annual growth rates of total milk production in India over the period 1990-2019 and over the decade is displayed in Table 1. A perusal of the table shows that the compound annual growth rate of milk production is 4.36 per cent per annum during the three decades from 1990 increasing consistently to 2019. The decadal growth rate shows a steady increase in the yield of milk and the availability of milk has seen a healthy growth in the recent decade at 4.17 per cent. However, the increase in milk yield has been attributed to the increase in animals rather than the yield itself hence education of the milch livestock owners and farmers about the importance of animal breeding program and health care practices is much required (Mishra *et al.*, 2021). The demand for milk and milk products has seen a considerable rise due change in per capita income, population boom food habits and export potential. The positive growth rate and the instability in milk production over different states displays the potential to raise average yields substantially through improved breeding practices in a number of major producing states over time.

## ARIMA

### Model identification

From Fig 1 it is seen that the milk production shows a linear trend and the stationarity of the data is determined by using the augmented Dickey-Fuller (ADF) test which is estimate by regression of the given equation.

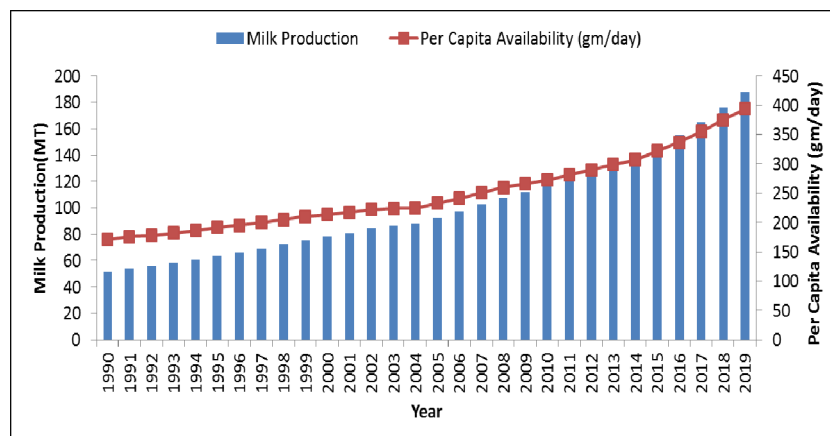
$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + \epsilon_t$$

Where,

$\epsilon_t$  = White error term

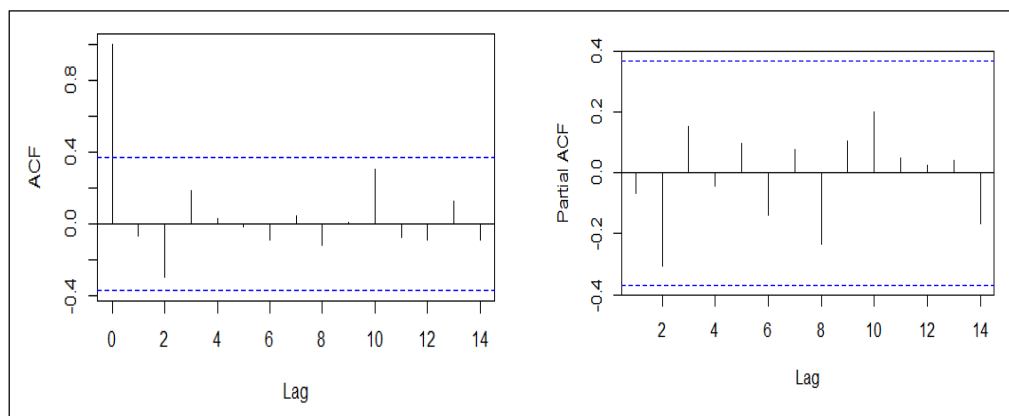
$$\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2}), -\Delta Y_{t-2} = (Y_{t-2} - Y_{t-3}), \text{ etc}$$

Where the Null hypothesis:  $H_0: \delta=0$  and Alternative hypothesis:  $H_1: \delta<0$  (*i.e.*, the time series is stationary). The current scenario shows 2.25 (Table 2) which in absolute terms is much less than even the 10 per cent critical  $\tau$  value of -3.1378 indicating the presence of a unit root, hence the time series is non-stationary and requires differencing. After differencing the series two times, the Dickey-Fuller Statistic ( $\tau$ ) is significant at the 10 per cent level with -3.30 which is greater than the critical value. After verification of the stationarity of the series, the ACF and PACF correlograms are examined to determine the order of the model.



**Fig 1:** Milk production and per capita availability of milk over the year.

Source: Ministry of Fisheries, Animal Husbandry and Dairying, 2019.



**Fig 2:** ACF and PACF of double differenced data.

**Table 1:** Compound annual growth rate of milk production and per capita availability during 1990-2019.

Period	Milk production	R <sup>2</sup>	Per capita availability	R <sup>2</sup>
1990-99	4.34***	0.99	2.29***	0.99
1999-10	4.12***	0.98	2.51***	0.95
2010-19	5.43***	0.99	4.17***	0.98
Overall	4.36***	0.99	2.71***	0.98

\*\*\*Significant at 1 per cent level of significance.

Source: Computed by author based on the secondary data obtained from Basic Animal Husbandry Statistics, 2019.

**Table 2:** ADF test for testing stationarity of the model.

Order of the model	Dickey-fuller statistic ( $\tau$ )	p-value
0	2.25	0.99
1	0.61	0.96
2	-3.30	0.09

Source: Computed by author using R (4.0.2) software.

**Table 3:** Estimation and diagnosis of ARIMA model.

ARIMA (p, d, q)	AIC	BIC
(0, 2, 0)	87.89	89.22
(1, 2, 0)	89.88	92.54
(0, 2, 1)	89.87	92.53
(1, 2, 1)	87.89	91.88

#### Shapiro-wilk test

W	0.95
p-value	0.19

Source: Computed by author using R (4.0.2) software.

**Table 4:** Forecast of milk production for the fitted ARIMA (0,2,0) model in million tonnes.

Year	Actual production	Predicted production
2015	146.3	143.0
2016	155.5	154.9
2017	165.4	164.7
2018	176.3	175.3
2020	-	199.1
2021	-	210.5
2022	-	221.9
2023	-	233.3
2024	-	244.7

Source: Computed by author using R (4.0.2) software.

We observe in Fig 2 that none of the lags seems to be statistically different from zero for both the autocorrelations and partial autocorrelations. Hence, it is implied that there is no need for adding AR (p) or MA (q).

#### Estimation

Four models are chosen and tested to find a model with good fit. This is done to eliminate the error of choosing the wrong model. The models included for testing include ARIMA (0,2,0), ARIMA (0,2,1), ARIMA (1,2,0) and ARIMA (1,2,1). All models are within the limit which indicates good fit. The appropriate model is chosen based on the minimum values

of Akaike information criterion (AIC) and Bayesian information criterion (BIC) which are provided by the results from R. It can be seen from Table 3 that ARIMA (0,2,0) and ARIMA (1,2,1) show same values for AIC. However, the BIC value of the former model is lesser compared to the latter hence the ARIMA (0,2,0) is selected for the purpose of modelling and subsequent forecasting of milk production.

#### Diagnostic checking

To determine the validity of the ARIMA (0,2,0) model, the residuals are tested for normality with the help of Shapiro-Wilk test. From Table 3, it can be seen that the p-value is 0.19 which is greater than 0.05, hence it can be concluded that the residuals are normally distributed and the model is suitable for predicting the production of milk.

#### Forecasting

ARIMA (0,2,0) model is used for comparing the actual and predicted values of milk production from the year 2015-2019 for further validation of the model. From Table 4 it can be noted that most of the predicted values lie in close proximity with the actual values. Finally, the table also shows the results from the forecasts for five years along with the lower and upper confidence interval at 80 and 95 per cent *i.e.* from 2020 to 2024. The milk production is expected to be 199.1, 210.5, 221.9, 233.3 and 244.7 Million tonnes (MT) for 2020, 2021, 2022, 2023 and 2024 respectively. A similar trend in increasing production of milk in India was reported by Mishra *et al.* (2020). This is expected to facilitate growth in the other sectors of the economy as well. According to the conclusions of the another study on milk production by ARIMA methodology, India is expected to be at the forefront in milk production among South Asian countries in the next five years (Mishra *et al.*, 2021).

#### CONCLUSION

The CAGR of milk production is higher in comparison to per capita availability which shows that the availability of milk is relatively lower and to ensure food and nutritional security of the growing population, there is a need to increase the production of milk from its present level. To do this, optimal allocation of resources is critical and forecasting is an important exercise to determine the direction of the present action. Accordingly, the present investigation forecasts that milk production will increase in the coming years and expected to be 244.7 MT 2023-24. However, it is far below the target of 300 MT made by the National Action Plan on

Dairy Development Vision. Hence, there is a need to boost the production of milk and this prediction can be beneficial for policy makers in formulating strategies for augmenting and sustaining production of milk in the country and this can be broadly achieved by tackling supply side problems with increase number of cross bred populations, adequate nutrition while also creating an efficient market with clear demand signals.

## REFERENCES

- Deshmukh, S.S. and Paramasivam, R. (2016). Forecasting of milk production in India with ARIMA and VAR time series models. *Asian Journal of Dairy and Food Research*. 35(1): 17-22.
- FAO (2019). Food Outlook, Dairy Market Review Report. Rome, Italy.
- Government of India (2018). National Action Plan for Dairy Development Vision-2022. Department of Animal Husbandry and Dairying, Ministry of Fisheries, Animal Husbandry and Dairying.
- Government of India (2019). Basic Animal Husbandry Statistics-2019. Department of Animal Husbandry and Dairying, Ministry of Fisheries, Animal Husbandry and Dairying. Krishi Bhavan, New Delhi.
- Gujarati, D.N. and Porter, D.C. (2009). Basic Econometrics. Fifth Edition McGraw-Hill/Irwin, New York.
- Kaur, S. and Rakshit, M. (2018). Time series analysis of average retail price of milk with autoregressive integrated moving average in R. *International Journal of Research in Engineering Application Management*. 4(4): 649-655. DOI: 10.18231/2454-9150.2018.0557.
- Lohano, D.H. and Soomro, M.F. (2006). Unit root test and forecast of milk production in Pakistan. *International Journal of Dairy Science*. 1(1): 63-69.
- Mishra, P., Fatih, C., Niranjana, H.K., Tiwari, S., Devi, M. and Dubey, A. (2020). Modelling and Forecasting of Milk production in Chhattisgarh and India. *Indian Journal of Animal Research*. 1-6. DOI: 10.18805/ijar.B-3918.
- Mishra, P., Matuka, A., Abotaleb M.S.A., Weerasinghe, W.P.M.C.N., Karakaya, K. and Das, S.S. (2021). Modeling and forecasting of milk production in the SAARC countries and China. *Modeling Earth Systems and Environment*. DOI: 10.1007/s40808-021-01138-z.
- Paul, R.K. and Das, M.K. (2010). Statistical modelling of inland fish production in India. *Journal of Inland Fish Society of India*. 42(2): 1-7.
- Paul, R.K., Alam, W. and Paul, A.K. (2014). Prospects of livestock and dairy production in India under time series framework. *Indian Journal of Animal Science*. 84(4): 462-466.
- Sankar, T.J. and Vijayalakshmi, C. (2017). Implementation of ARIMA Model for Ghee Production in Tamil Nadu. *International Journal of Pure and Applied Mathematics*. 113(6): 56-64.
- Taye, B.A., Alene, A.A., Nega, K.A and Yirsaw, G.B. (2020). Time Series Analysis of Cow Milk Production at Andassa Dairy Farm. West Gojam Zone, Amhara Region, Ethiopia. *Modeling Earth Systems and Environment*. p 364.