B-4409
[1-6]

RESEARCH ARTICLE Indian Journal of Animal Research, Volume Issue : ()

Future Milk Production Prospects in India for Various Animal Species using Time Series Models

Monika Devi[1], Umme Habibah Rahman[2], W.P.M.C.N. Weerasinghe[3], Pradeep Mishra[4], Shiwani Tiwari[4], Kadir Karakaya[5]

10.18805/IJAR.B-4409

ABSTRACT

Background: The Indian dairy industry is contributing significantly to the country's economic growth. Since the variations in milk production will be a huge matter for dairy products as well as for farmers, investors and policymakers in the country, an accurate forecast of milk production is extremely very important.

Methods: This study represents an ARIMA modelling approach for forecasting the milk production in India and milk production by five major milk producing animal species namely, Cow, Buffalo, Goat, Sheep and Camel by using annual data from 1961 to 2018. ARIMA (0,2,1) model was selected as the best model in forecasting milk production in India.

Result: There will be an increment in the overall milk production in India according to the study. Further, there will be an increase in buffalo, cow and goat milk production while a decrease in milk production by camels and sheep.

Key words: ARIMA, Milk availability, Milk Production.

INTRODUCTION

Milk is a very essential part of our daily diet and has a number of health benefits. We consume milk in liquid form directly or in processed form as Butter, cheese, Skim Milk Powder (SMP), Whole Milk Powder (WMP), casein, cream, skim milk, condensed and evaporated skim milk, whey and yoghurt. With accounting for almost 22 percent of total milk production in the world, India is the leading milk producing country followed by the European Union, USA, Pakistan, Brazil and China (Anonymous, 2019). Global milk production has reached up to 852 MT in the year 2019 from 530 MT production of milk in 1988, almost 60 percent increment has been noted in last three decades and main contributors in this growth are South Asian countries; India, China and Pakistan. Holding a share of about 90 percent, India and Pakistan were the top two milk producers in Asia's total milk production of 360 MT in the year 2019. In the year 2019, the whole fresh milk market of Asia alone was above 300 Billion dollars and India led the whole fresh milk market with a value of around 145 Billion dollars alone. With an annual increment of about 1.5 percent since the last ten years, the number of milk producing animals has been reached upto 427 Million heads in Asia. Whereas Global milk production growth was 1.4 percent in the year 2019 over the previous year's production, India has recorded a positive growth of about 4.5 percent. India produced 196.17 MT of milk in the year 2019and was the leading contributor among all milk producing countries which are responsible for this milk production expansion in Asia as well as in the world. India is not among the main milk exporter but has fourth place in Butter export according to data of the year 2019. Due to the fast growing urbanization, increased demand for processed food products is one of the main reasons for India's milk production growth. In many states of India, milk production is a tradition as milk and milk products have a very important place in their diet. Like other developing countries in India also milk is produced by smallholders (around 70 million rural households) to meet their household livelihood as it provides quick return and works as a source of cash income. Every year Government of India release funds to strengthen the infrastructure for quality and clean milk production and assistance to cooperative under National Program of Dairy Development (NPDD) and in the year 2018-19 Government released an amount of Rs. 26986 corers for the same. Uttar Pradesh, Rajasthan, Madhya Pradesh, Andhra Pradesh and Gujarat were the top five milk producing states in the year 2018-19.Per capita availability of milk in India has almost doubled in last fifteen years with 394 (gms/

[1]Department of Mathematics and Statistics, Chaudhary Charan Singh Haryana Agricultural University, Hisar-125 004, Haryana, India.
[2]Departments of Statistics, Assam University, Silchar-788 011, Assam, India.
[3]Department of Statistics and Computer Science, University of Kelaniya, Kelaniya, Sri Lanka.
[4]College of Agriculture, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Powarkheda-461 110, Madhya Pradesh, India.
[5]Department of Statistics, Faculty of Science, Selçuk University, Konya, Turkey.

Corresponding Author: Monika Devi, Department of Mathematics and Statistics, Chaudhary Charan Singh Haryana Agricultural University, Hisar-125 004, Haryana, India. Email: mscagstats@gmail.com

How to cite this article: Devi, M., Rahman, U.H., Weerasinghe, W.P.M.C.N., Mishra, P., Tiwari, S. and Karakaya, K. (2021). Future Milk Production Prospects in India for Various Animal Species using Time Series Models. Indian Journal of Animal Research. DOI: 10.18805/IJAR.B-4409.

Submitted: 21-01-2021 **Accepted:** 24-04-2021 **Online:** 19-06-2021

day) in 2018-19 with the top place of Punjab with 1181 gms/ day followed by Haryana (1084 gms/day). Buffalo, cow, goat and sheep are the main milk producing animals and count of buffaloes, cows (crossbreds and indigenous) and goats was 44767, 52840 and 36834 (in thousands) respectively with an average milk yield of 5.62, 5.48 and 0.45 (Kg/day) respectively in the year 2018-19 (Anonymous 2018).

To know the availability and need of milk, forecasting of milk production is required so that necessary policy formations can be done (Mishra *et al,* 2020).When policy matters are discussed it is important to have estimates of future production that is likely to take place in the region wise (Mishra *et al*, 2020). In this direction, Sharma *et al.* (2018) investigated the monthly arrival of Rohu fish using ARIMA in the Jammu Region of J&K State. Deshmukh and Paramasivam (2016) evaluated milk production forecasting using ARIMA and VAR time series model. Chaudhari and Tingre (2014) considered egg production in India using ARIMA modelling. Mishra *et al* (2020) investigate time series investigation of milk production in major states of India using ARIMA Model. Mishra *et al* (2020) also studied modelling and forecasting of milk production in Chhattisgarh and India. Li *et al.* (2020) also studied the genome-wide association study of milk production traits in a crossbred dairy sheep population using three statistical models. Mishra *et al.* (2021) also studied modelling and forecasting of milk production in SAARC countries and China. The present study is devoted to meet the future demand of various animal species in India.

MATERIALS AND METHODS

Indian dairy industry provides livelihood to about 70 million households. A key feature of India's dairy sector is the predominance of small producers. The livestock sector of this industry will approve the growth of both the socio-economic as well as the national economy. This investigation brings out the important features of the results obtained by employing various statistical modelling procedures to milk production of India collected for during 1961-62 to 2018-19 from www.fao.org.

Auto-regressive integrated moving (ARIMA) approach

Time series is a branch of Statistics; the object is to study variables over time. Among its main objectives is the determination of trends within these series as well as the stability of values (and their variation) over time. Unlike traditional econometrics, the purpose of time series analysis is not to relate variables to one another, but to focus on the "dynamics" of a variable. In particular, linear models (mainly AR and MA, for Auto-Regressive and Moving Average), (Box and Jenkins, 1976), conditional heteroscedasticity models, notably ARCH (Auto-Regressive Conditional Heteroscedasticity), (Engel, 1982) are used in modeling time series. In this study, we deal with Auto-Regressive Integrated Moving (ARIMA) process, (called Box-Jenkins Approach) to estimate and forecast the milk production in India and five major milk producing animal species namely, cow, buffalo, goat ,sheep and camel over the period 1961 to 2018. The data for analysis was collected from the website of www.fao.org.in.

In practice, it is impossible to know the probability distribution of a time series; y_t, t ≥0; therefore, when primary interest is in the modeling of the conditional distribution (a priori constant in time) of y_t *via* its density:

$$f\,(y_t \mid y_{t-p}),\ t,\ p \geq 0; \ \ldots\ldots(1)$$

Conditioned on the history of the process: $y_t = y_t, y_{t-1}, \ldots, y_0$. It is therefore a necessity to model y_t on its past values.

Auto-regressive model, AR (p)

The conditional approach in Equation (1) provides a decomposition prediction error, according to which:

$$y_t = E\,(y_t \mid y_{t-p}) + \epsilon_t \Leftrightarrow y_t = \sum_{t=1}^{p} b_i\, y_{t-p} + \epsilon_t \ \ldots\ldots(2)$$

Where

$E\,(y_t \mid y_{t-p})$, is the component of y_t, that can give rise to a forecast, when the history of the process, $y_{t-1}, y_{t-2}, \ldots y_0$ are known and ϵ_t represents unpredictable information. We suppose, $\epsilon_t \sim WN\,(0, \sigma^2)$, is a white noise process. The equation (2) represents an autoregressive model (AR) of order *p*. As an example an autoregressive process of order 1, AR (1) is defined:

$$y_t = c + \alpha y_{t-1} + \epsilon_t \ \ldots\ldots(3)$$

The value y_t depends only on its predecessor. Its properties are functions of α which is a factor of inertia. Autoregressive processes AR(p) assume that each observation y_t can be predicted by the weighted sum of a set of previous observations $y_{t-1}, y_{t-2}, \ldots y_{t-p}$, plus a random error term. The other type of process of the box-Jenkins approach is Moving Average, MA(q).

Moving-average process MA (q)

The moving average processes assume that each observation y_t is a function of the errors in the preceding observations, $\epsilon_{t-1}, \epsilon_{t-2}, \ldots \epsilon_{t-p}$, plus its own error. A moving average process is given as Mishra *et al*. (2021):

$$y_t = \sum_{i=1}^{q} \theta_i \epsilon_{t-1} + c \ \ldots\ldots(4)$$

The combination of the two models, AR (p) in equation (3) and MA(q) in equation (4) is an ARMA (p, q) process; which is the most popular models of the Box Jenkins for its flexibility and suitability for various data types. The model is designed as follow:

$$\text{ARMA}\,(p,q): \sum_{i=1}^{p} \beta_i y_{t-i} = \sum_{i=1}^{q} \theta_i\, \epsilon_{t-1} \ \ldots\ldots(5)$$

With: $\beta_i\,(i=1, \ldots, p)$, $\theta_i\,(i=1, \ldots, q)$, $\in \mathbb{R}$, $\epsilon_t \sim WN\,(0, \delta_\varepsilon)$

The time series y_t must be stationary to be fitted by an ARMA models. We take the case of weak stationary and we put its definition:

Definition: A time process y_t with real values and discrete time $y_1, y_2, \ldots y_t$ It is stationary in the weak sense (or "second order", or "in covariance") if:

- $E(y_i) = \mu \forall\ i=1, ..., t.$
- Var $(y_i) = \sigma^2 \neq \infty\ \forall\ i=1, ..., t.$
- $Cov(y_i, y_{i-k}) = f(k) = p_k\ \forall i=1, ..., t.$

When one or more stationary conditions are not met, the series is said to be non-stationary. This term, however, covers many types of non-stationary, (non-stationary in trend, stochastically non-stationary), we focused on the later. Thus, if y_t is stochastically non-stationary, a difference stationary technique should be applied. Consequently, a series is a stationary in difference if the series obtained by differentiating the values of the original series is stationary. Generally, we used the KPSS test, (Kwiatkowski *et al.,* 1992; Leybourne and McCabe, 1994).

The difference operator is given by $\Delta(y_t) = y_t - y_{t-1}$: if the series is differentiated **d** times, we say that it is integrated of order I (d). The process will be noted as ARIMA (p,d,q), defined by the equation:

$$\beta(L)(1-L)^d y_t = \theta(L)\ \varepsilon_t \quad ......(6)$$

With, L: is the lag operator (L) or backshift operator (B); If the time series $X_t = (1-L)^d y_t$ is stationary, then, estimating an ARIMA (p,d,q), process on y_t is equivalent to estimating an ARMA (p, q) process on X_t.

Box and Jenkins (1976) proposed a prediction technique for a univariate series that is based on the notion of the ARIMA process. This technique has three stages: identification, estimation and verification. The *first step* is to identify the ARIMA model (p, d, q) that could spawn the series. It consists, first of all, in transforming the series in order to make it stationary (the number of differentiations determines the order of integration: d) and then to identify the ARMA model (p, q) of the series transformed with the correlogram and partial correlogram. The graph of autocorrelation (correlogram) and partial autocorrelation coefficients (partial correlogram) give information on the order of the ARMA model. Thus, if we observe that the first two autocorrelation coefficients are significant, we will identify the following model: MA (2).The *second step* is to estimate the ARIMA model using a non-linear method (nonlinear least squares or *maximum likelihood*). These methods are applied using the degrees p, d and q found in the identification step.

Generally, we use the *Maximum Likelihood method*; by consider that the errors ε_t follow a normal distribution, $N(0,\sigma^2_\varepsilon)$. The log-likelihood function of ARMA (p,q)process is defined as Lama *et al*. (2021) :

$$\text{Log}Lt = -\frac{T}{2}\log 2\pi - = -\frac{T}{2}\log\sigma^2_\varepsilon - \frac{T}{2}\log(det\,[\psi'\psi]) - \frac{\varpi(\beta,\theta)}{2\sigma^2_\varepsilon} \quad ......(7)$$

With:

- T: number of observations**,**
- ψ a matrix of (p+q+T, p+q) dimensions, dependent of β_i = (i = 1,...,p) and θ_i = (i = 1,...,q)**,**
- $\varpi(\beta,\theta) = \sum_{t=-\infty}^{T} (\mathbb{E}\ [\varepsilon_t / Xt, \beta i, \theta i, \sigma^2_\varepsilon])^2$,

with i =1,...,p; j =1,...,q.

The *third step* is to check whether the estimated model reproduces the model that generated the data. For this purpose, the residuals obtained from the estimated model are used to check whether they behave like white noise errors using a “portmanteau” test (a global test that makes it possible to test the hypothesis of independence of residues). The common tests are based on residuals analysis for normality and autocorrelation is Durbin and Watson (1950), test for Homoscedasticity: Breusch (1978); Breusch and Pegan (1979), ARCH Test, Engel (1982).The last point under this step is the prediction of future values of y_t by the selected model.

RESULTS AND DISCUSSION

From Table 1 any one can see, that milk production in India has been increased from 10929 to 86262 thousand tones, 8 to 17 thousand tones, 6900 to 83634 thousand tones, 535 to 6166 thousand tones and 173 to 220 thousand tones

Table 1: Per se performance of Milk production in India (Thousand Tonnes).

	Mean	Minimum	Maximum	Standard deviation	Skewness	Kurtosis	SGR (%)
Buffalo milk	33723.47	10929	86262	21561.252	.781	-.481	386.46
Camel milk	12.61	8	17	3.175	.231	-1.412	-14.42
Cow milk	27378.73	6900	83634	20217.373	1.138	.477	487.62
Goat milk	2346.33	535	6166	1702.655	.564	-1.015	599.88
Sheep milk	199.04	173	220	13.001	-.389	-.667	4.26

Table 2: Selecting best model for forecasting.

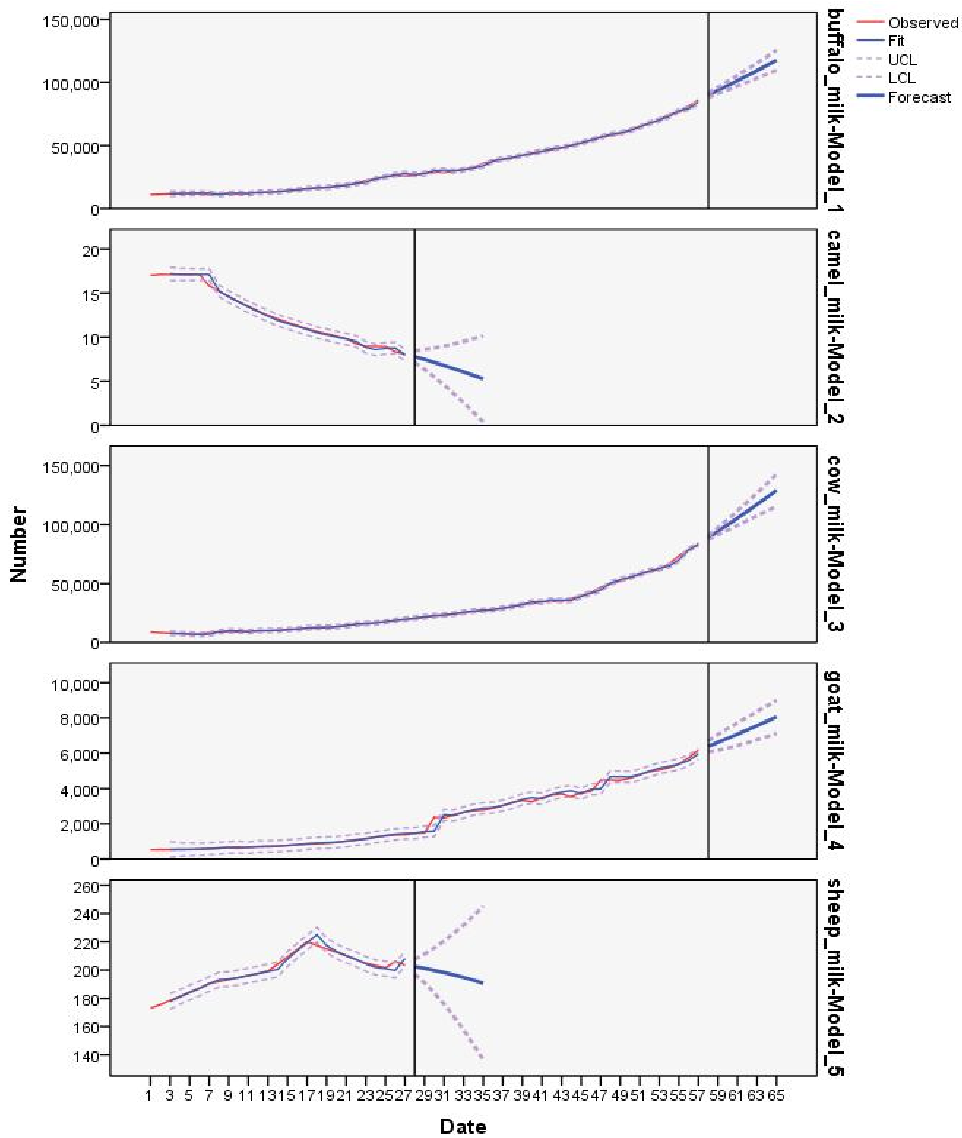
		Model	Stationary R-squared	R-squared	RMSE	MAPE	MaxAPE	MAE	MaxAE	Normalized BIC
Milk Production '000t (India)	Buffalo milk	ARIMA(0,2,1)	0.349	.999	803.606	2.128	11.776	570.791	1985.847	13.524
	Camel milk	ARIMA(0,2,1)	0.236	.990	.315	1.464	8.151	.164	1.288	-2.051
	Cow milk	ARIMA(0,2,1)	.182	.998	876.118	2.407	17.340	554.212	3756.591	13.697
	Goat milk	ARIMA(0,2,1)	.566	.992	157.746	3.209	33.466	79.698	796.837	10.269
	Sheep milk	ARIMA(0,2,1)	.089	.952	2.522	.673	3.478	1.390	7.564	2.108

for buffalo, camel, cow, goat and sheep respectively till 2018-2019.According to figures overall milk production and availability are growing at a good pace but the milk production of camel and sheep have been decreased during the study period. Except the sheep milk, the rest of having positive skewness,which indicates that during the study investigation production of milk has been increased for other breeds.

For the study, have 8 time series, for the stage of identification of the integration orders of the time series of milk production: buffalo milk, camel milk, cow milk, goat milk and sheep milk respectively, by using the tests of ADF. All the series are of deterministic non-stationarity (DS), after the first differentiation $\Delta (Z_t) = Z_t - Z_{t-1}$ t= 1,2, ...57 and the application of (ADF and KPSS) unit root tests has indeed shown that all these series are stationary; to model them using ARMA-type processes, followed by the steps of the Box and Jenkins approach cited above. All the criteria (Stationary R-squared, R-squared, RMSE, MAPE…etc) lead us to select the models (column 2, Table 1) to represent the dynamics of the 8 time series, the results are detailed in Table 2. As indicated in the theoretical section, the last step of the Box and Jenkins methodology is to forecast the series studied based on the selected (validated) processes in the second column of Table 2. The best model selected is an ARIMA (0,2,1) for milk production in India time series. The model equation is given by

$$Z_t = 2 * Z_{t-1} - Z_{t-2} + \varepsilon_t,\ E(\varepsilon_t) = 0$$

According to the forecasts of our study see Table 2, milk production continues its upward trend in India; it is expected to record 201376 thousand tones and 1507.2



Plot 1: Forecasting plot.

Table 3: Forecasting of various breeds milk production in India (Thousand Tonnes)

Years	PF	L80	U80	L95	U95	PF	L80	P80	L95	P95
			Buffalo_milk					**Camel_milk**		
2017	84561	83520	85602	8295	86170	8	8	8	7	9
2018	89955	88914	90996	88346	91564	8	7	8	7	8
2019	93716	92094	95338	91209	96222	7	7	8	6	9
2020	97544	95372	99717	94187	100902	7	6	8	5	9
2021	101441	98714	104168	97226	105656	7	5	8	5	9
2022	105405	102109	108700	100311	110498	6	5	8	4	9
2023	109436	105553	113319	103435	115437	6	4	8	3	10
2024	113535	109045	118026	106595	120475	6	3	8	2	10
2025	117702	112583	122821	109790	125614	5	2	8	0	10
			Cow_milk					**Goat_milk**		
2017	83094	81957	84230	81336	84851	5967	5764	6171	5653	6282
2018	88973	87836	90110	87216	90730	6388	6185	6592	6074	6703
2019	94417	92420	96414	91331	97503	6615	6325	6905	6166	7064
2020	99966	97034	102898	95435	104497	6846	6488	7205	6292	7401
2021	105620	101667	109573	99510	111729	7082	6664	7499	6437	7727
2022	111378	106320	116437	103560	119196	7322	6851	7792	6595	8048
2023	117242	110998	123486	107591	126893	7566	7047	8085	6764	8368
2024	123210	115704	130716	111608	134812	7814	7250	8379	6942	8687
2025	129283	120442	138124	115619	142947	8067	7459	8674	7127	9006
			Sheep_milk							
2017	208	205	211	203	213					
2018	202	199	206	197	208					
2019	201	165	201	191	211					
2020	200	190	210	184	216					
2021	198	184	213	176	221					
2022	197	178	215	167	226					
2023	195	171	219	158	232					
2024	193	164	222	147	238					
2025	191	156	225	136	245					

thousand tones in 2020-21, 219730 thousand tones and 1600 thousand tones in 2022- 23 and in 2025-26. Also for validation of these forecasted values are very close to actual values for the year 2018-19 and 2019-20 (Ministry of Agriculture and Farmers Welfare, GoI, 2020). This is well explained in part, also by the forecasts of augmentation of the population of buffaloes, camel, cow, goat and sheep in Table 3.The production of buffalo milk may has increased from 84561 thousand tones (2017) to 117702 thousand tones (2025); in the same way, 83094 thousand tones to 129283 thousand tones for cow milk and 5967 to 8067 thousand tones for goat milk have been showing an increasing result. On the other way, there are decline results for camel and sheep milk production. The prediction values are decreases for these two categories. As Table 3 shows in the case of camel and sheep, the milk production decreases from 8 to 5 thousand tones and 208 to 191 thousand tones respectively.

CONCLUSION

Forecasting the milk production in India for five major milk producing animal species namely, Cow, Buffalo, Goat, Sheep and Camel with a time series modelling approach was carried out through this study. The study was carried out by using the data which was obtained from the website of www.fao.org.in over the period 1961 to 2018. ARIMA process followed by the Box and Jenkins methodology was used in developing the model in order to identify the future movements of the milk production in India for the five breeds. For all five breeds, ARIMA (0,2,1) model was evidently selected as the best model for forecasting the milk production in India. According to the findings, the overall milk production in India shows an upward trend which leads to production of 201376 thousand tones, 1507.2 thousand tones, 219730 thousand tones for the period of 2020-22 respectively. And when considering the milk production in detail according to the breed, there can be seen a vast increase in milk production in Buffaloes, Cows and Goats in future years. But at the same time, there is a decreasing trend for milk production by Camels and Sheep. The existing data to shows a decrease in milk production in these two categories. From the above, it is evident that the ARIMA time-series modelling approach is the best one for the data sets under consideration. Accordingly, this approach is used to forecast the milk production of different breeds' population of India. The highest milk production would be 129283 thousand tones for the local cows in India in 2025-26. Also forecasting values give direction that the milk production of camel and sheep has been decreased during the study period. Overall Milk availability would be increase for total India by the next 5 years.Thedemand for milk grows when there is an increasing consumer's preference for high quality milk as well as for dairy products. So the forecasts from the model also depict an important piece of information for potential investors in the dairy products market.

REFERENCES

Anonymous. (2018). National Action plan of Diary Development, Ministry of Agriculture and Farmers Welfare.

Anonymous. (2019). Department of Animal Husbandry, Dairying and Fisheries, Ministry of Agriculture and Farmers Welfare, GoI.

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

Breusch, T. (1978). Testing for Autocorrelation in Dynamic Linear Models.Australian Economic Papers. 17: 334-355. doi: 10.1111/j.1467-8454.1978.tb00635.

Breusch, T.S. and Pagan,A.R. (1979). A simple test for heterosce--dasticity and random coefficient variation, Econometrica. 47: 1287-1294.

Chaudhari, D.J. and Tingre, A.S. (2015). Forecasting eggs production in India. Indian Journal of Animal. Research. 49(3): 367-372.

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.

Durbin, J. and Watson, G.S. (1950). Testing for serial correlation in least squares regression: I. Biometrika. 37 (3/4): 409-428.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica. 50: 987-1008.

Intriligator, M.D, Bodkin, R.G and Hsio, C. (1996). Econometric Models, Techniques and Applications, Prentice Hall edition.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. Journal of Econometrics. 54: 159-178.

Landes, M., Cessna, J., Kubarka, L. and Jones, K. (2017). India's Dairy Sector: Structure, Performance and Prospects, Economic Research Service/USDA. LDPM-272-01.

Leybourne, S.J. and McCabe, B.P.M. (1994). A Consistent Test for a Unit Root, Journal of Business and Economic Statistics, 12: 157-166.

Lama, A., Singh, K. N., Singh, H., Shekhawat, R., Mishra, P. and Gurung, B. (2021). Forecasting monthly rainfall of Sub-Himalayan region of India using parametric and non-parametric modelling approaches. Modeling Earth Systems and Environment. 1-9.

Ministry of Agriculture and Farmers Welfare. (2018). National Action Plan of Diary Development, Ministry of Agriculture and Farmers Welfare.

Mishra, P., Fatih, C., Niranjan, 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. 54: 912-917

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, 1-13.

Mishra, P., Fatih, C., Vani, G.K., Tiwari, S., Ramesh, D., and Dubey, A. (2020). Time Series Investigation of Milk Production in Major States of India Using ARIMA Modeling. Journal of Animal Research. 10(1): 77-84.