Indian J Dairy Sci 77(3): 267-274
https://doi.org/10.33785/IJDS.2024.v77i03.011

RESEARCH ARTICLE

Time series analysis and ARIMA models for milk yield data

Navneet Kaur and Vinod Kumar(✉)

Received: 24 September 2023 / Accepted: 16 December 2023 / Published online: 23 June 2024

Abstract: The present study is concerned with the time series analysis of milk yield data collected from Instructional Dairy Farm (IDF), Pantnagar, India. ARIMA models have been constructed for milk yields of Crossbred and Sahiwal cows and Murrah buffaloes. It is found that ARIMA models (0,1,1) $(1,1,1)_{12}$ for Crossbred cows, (1,1,0) $(0,1,1)_{12}$ for Sahiwal cows and (0,1,1) $(2,1,0)_{12}$ for buffaloes are more appropriate for predicting milk production.

Keywords: Trend analysis, stationarity, forecasting, white noise, autocorrelation

Introduction

Dairy is a growing industry in India and a large part of the country's population is engaged in dairy production, making it a source of financial benefits to the agricultural population and an important contributor to the country's economy. India ranks first in terms of milk production and milk is the second most important commodity after rice in our country, accounting for 20% of overall production and contributing about 5.3% of India's agricultural GDP (Belhekar and Dash, 2016). The National Dairy Development Board (NDDB) estimates that demand for milk is expected to reach 180 million tons by 2022 and to meet this demand, production must be increased by an average of 5 million tons per year over the next 15 years, otherwise, India will have to depend on milk imports from the world market.

Department of Mathematics, Statistics and Computer Science, G.B. Pant University of Agriculture & Technology, Pantnagar-263 145, Uttarakhand, India

Vinod Kumar(✉)
E-mail: vinod_kumarbcb@yahoo.com

Suresh and Priya (2011) forecasted the sugarcane area, production and productivity of Tamilnadu through fitting of univariate Auto Regressive Integrated Moving Average (ARIMA) models. Using the Augmented Dickey Fuller unit root test, Etuk & Mohamed (2014) found that seasonal series of precipitation data are stationary for one year. To predict future sustainable milk production and its policy implications, Deshmukh and Paramasivam (2016) used vector autoregression (VAR) and autoregressive integrated moving average (ARIMA) models. Model reliability was assessed using the Akaike Information Criterion (AIC), Schwarz-Bays Criterion (SBC), The Mean Absolute Percentage Error (MAPE), R-square and RMSE. Thakur and Gupta (2020) used ARIMA time series (p,d,q) to estimate monthly milk production over time. The study showed that climatic factors and weather conditions were responsible for the year-to-year variation in monthly milk production for all crossbred cows. Kamble et al. (2020) also published their paper on growth and performance of dairy sector in India and emphasized the need for increasing milk production to meet out the needs of the people. Taye et al. (2021) used inference statistics to test hypotheses, estimate trends, fit models and make predictions. They found that milk production was time-dependent and that the average amount of milk produced was decreasing.

Milk production is an integral part of Instructional Dairy Farm (IDF) Nagla, Pantnagar of the College of Veterinary and Animal Science, G.B. Pant University of Agriculture and Technology. Initially, the Tarai State Dairy Farm was established in 1948 with a herd of Murrah buffaloes and Haryana cows. Later, in 1956, Sahiwal cows were also included in the farm. As the population and consequently the demand for milk increased, more cows and buffaloes were introduced to the farm. Therefore, the future milk production trend needs to be monitored so that preventive measures can be taken to meet the future demand and hence the study was undertaken.

Materials and Methods

Data collection

The daily milk yield data of dairy cattle namely crossbred and Sahiwal and Murrah buffaloes for a period of 10 years (2011-2012 to 2019-2020) were collected from the Bureau of records maintained at Instructional Dairy Farm (IDF) Nagla, Pantnagar. Daily data were converted to monthly data and the averaged monthly milk yield data for a period of 2011-2018 were used for the development of the model and data of the period 2019 and 2020 were used for prediction and testing the efficiency of the model.

Tools

Statistical techniques commonly used by researchers, statisticians and meteorologists to assess the past trend and predict the future accordingly, include multiple regression analysis, analysis of variance (ANOVA), auto-regression, moving average, regression analysis and several other advanced techniques available nowadays, such as single-equation regression models, exponential smoothing method, simultaneous equation regression, autoregressive integrated moving average (ARIMA) models, vector autoregression, artificial neural network (ANN), fuzzy logic etc. which are frequently used in data analysis in various fields of science and technology using the latest software like SAS, SPSS, STATISTICA, JMP, EViews, GRETL, Rstudio etc. The present study focuses on ARIMA models for the analysis of milk productivity and forecasts future supply based on past trends to meet the forthcoming demands so that necessary actions can be taken.

Trend analysis

The Mann-Kendall test is a non-parametric test used to observe a trend in a data series, even if the data series has a seasonal component.

The Mann-Kendall test statistic is calculated as follows;

$$\tau = \frac{S}{n(n-1)/2}$$

It ranges from -1 to +1 and corresponds to the correlation coefficient in regression analyses. The null hypothesis of no trend is then rejected if S and τ are significantly non-zero.

Or Compute the M-K test statistic, Z_{MK} as: $Z_{MK} = \frac{S-1}{\sqrt{VAR(S)}}$

The statistical value of the standard Z-test is compared to the normal variance (within ±1.96) in the table at the 5% significance level for the null hypothesis of no trend and the alternative hypothesis of trend.

Stationarity detection by unit root test

KPSS test considers null hypothesis (H_0) that time series is stationary, and tests such as ADF and P-P, for which null hypothesis is on the contrary, is that the series possesses a unit root and data is not stationary. In the interpretation of the KPSS test, if the p-value of its test statistic is greater than the critical value of say 0.463, then we reject the null hypothesis of having a level stationary series and we accept the alternate hypothesis that it has a unit root. The P-P and ADF tests, on the other hand, test for the null hypothesis of unit root against an alternative hypothesis of stationary by rejecting the null hypothesis, if its *p*-value is less than the critical value.

Model-identification

Model-identification can be done by examining the significant spikes of the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. Software packages like GRETL and SPSS are used for the easy calculations of autocorrelation functions and partial autocorrelation functions.

Estimation of model parameters

By comparison, the residuals of the models with the lowest AIC and BIC values resemble white noise.

AIC= 2m-2 log L

Where, m= p+q (Estimated parameters)

L= the maximized value of the likelihood function of the model.

BIC is calculated as -2 ln(L)+ln(n)k, where L is the likelihood, n is the number of residuals and k is the number of free parameters.

Diagnostic check

The Ljung-Box Q statistic (LBQ) tests the null hypothesis that autocorrelation is zero up to lag k, which means that the data values are random and independent up to a certain number of lags. If the LBQ statistic is greater than a certain critical value, the autocorrelation of one or more lags may deviate significantly from zero, indicating that the values are not random and independent over time. Because the samples are finite, the autocorrelations for each sample may not be exactly zero. White noise ACFs are said to have a sampling distribution that can be approximated by a normal curve with zero mean and standard error. In the white noise process, 95% of all autocorrelation values in the sample must lie within the range defined by the mean ±1.96 standard error. In this case, since the process has a mean of zero and a standard error of about 95% of all autocorrelation values in the sample should be expected to fall within this range.

Ljung-Box test statistic

$$Q = T(T+2)\sum_{K=1}^{m}\frac{\hat{\tau}_K^{\ 2}}{T-K} \sim \chi_m^{\ 2}$$

T= Sample size

τ_K= Autocorrelation at lag K

K= lags, 1st lag, 2nd lag…………

Critical region

If $\chi_m^{\ 2}$ cal ≥±1.96 the null hypothesis is accepted, otherwise rejected.

If this condition is not met, the fitted model indeed follows a process with white noise, or the residuals are not white noise.

The quantitative performance of developed models may also be evaluated by applying various statistical indices viz. Root Mean Square Error (RMSE), Coefficient of determination (R^2), Coefficient of efficiency (CE) and Mean Absolute Error (MAE)

Statistical analysis

The time series analysis of milk yield data was carried out by using statistical tools described above and ARIMA models were developed for the milk yield data for different breeds of cows and buffaloes.

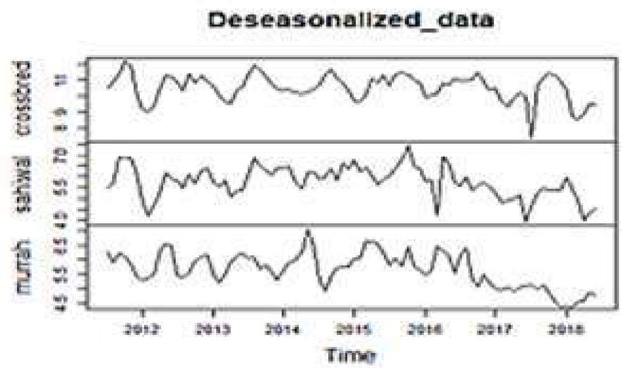

Fig 1: Seasonally differenced time series plot of monthly milk yield

Results and Discussion

The trend component of this time series was confirmed by Mann-Kendall's test. The values of the Mann-Kendall statistical test (τ) were -0.264986,-3.182073e-01 and -5.159664e-01 and Z values were -4.2892, -5.1512 and -8.3539 for Sahiwal and Murrah crosses, respectively. The calculated value was found to be not within the range of +1.00 and -1.00 and +1.96 and -1.96 for the calculated τ and Z value respectively, so the null hypothesis of the existence of a trend in the series was rejected at a 5% significance level. So, the data exhibits a trend component in the correlation function.

When developing any time series model, the first step is to determine stationarity. Since there is a trend in the data series, it is necessary to differentiate them (i.e. d=1). However, the milk production data are not stationary because the data are seasonal, with a periodicity of 12 months, and a seasonal differential is needed to achieve stationarity (d=1). Figure 1 shows a plot of the seasonally differentiated time series of monthly milk production.

Therefore, the long-term value of the trend component was taken for further analysis. Data without trends were used to test for stationarity, so the data still show non-stationarity. To achieve stationarity, seasonal differentiation is performed. After

Table 1: AIC values of the selected models for Crossbred, Sahiwal and Murrah

Crossbred		Sahiwal		Murrah	
Model	AIC	Model	AIC	Model	AIC
ARIMA (0,1,1) $(1,1,1)_{12}$	238.85	ARIMA (0,1,0) $(3,1,2)_{12}$	201.162	ARIMA (0,1,2) $(2,1,2)_{12}$	103.304
ARIMA (0,1,0) $(1,1,1)_{12}$	255.349	ARIMA (3,1,2) $(0,1,0)_{12}$	234.934	ARIMA (0,1,1) $(2,1,0)_{12}$	98.503
ARIMA (1,1,1) $(0,1,0)_{12}$	263.093	ARIMA (0,1,2) $(3,1,0)_{12}$	196.857	ARIMA (0,1,0) $(2,1,4)_{12}$	108.553
ARIMA (0,1,3) $(1,1,0)_{12}$	251.445	ARIMA (1,1,0) $(0,1,1)_{12}$	189.982	ARIMA (2,1,4) $(0,1,0)_{12}$	129.589
ARIMA (1,1,3) $(0,1,0)_{12}$	267.026	ARIMA (1,1,1) $(3,1,0)_{12}$	192.154	ARIMA (2,1,0) $(0,1,4)_{12}$	102.812

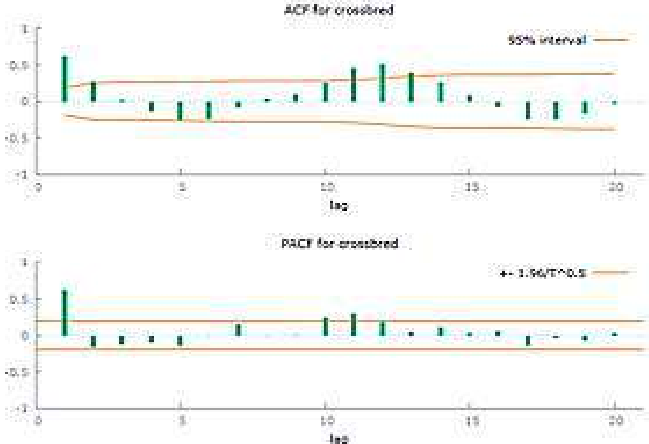

Fig 2: ACF & PACF graphs of de-trended & de-seasonalized data for Crossbred cows

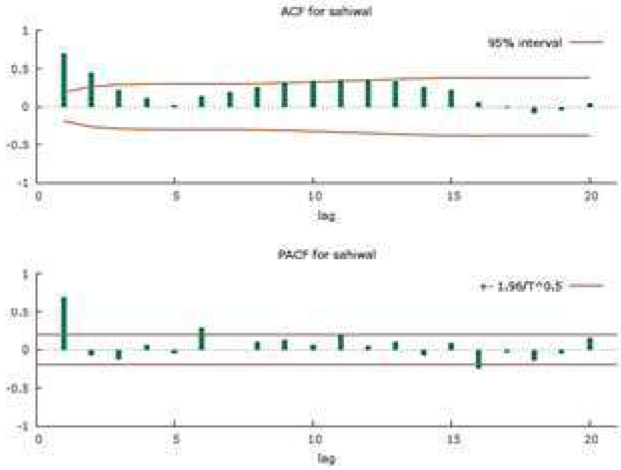

Fig 3: ACF & PACF graphs of de-trended & de-seasonalized data for Sahiwal cows

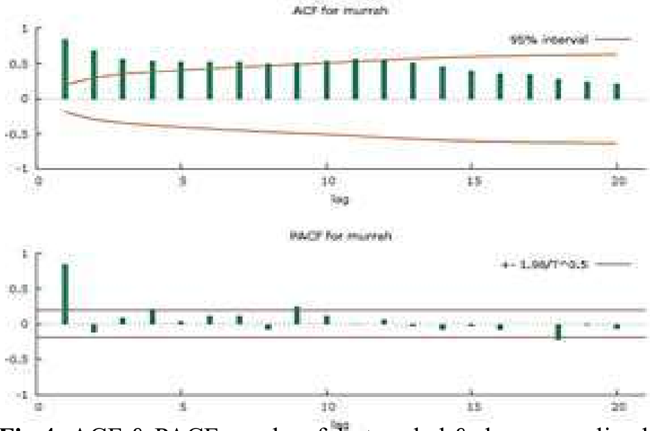

Fig 4: ACF & PACF graphs of de-trended & de-seasonalized data for Murrah buffaloes

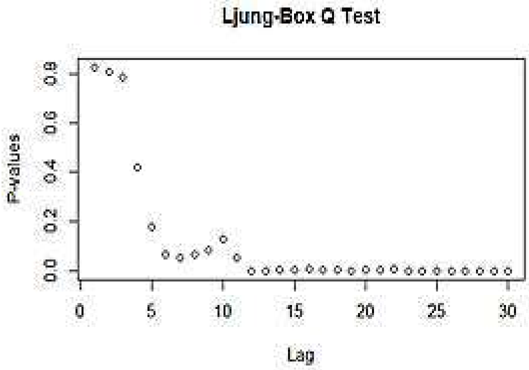

Fig 5: Ljung-Box Q plot for Crossbred

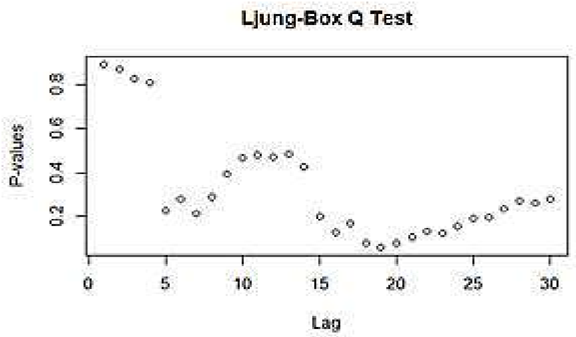

Fig 6: Ljung-Box Q plot for Sahiwal

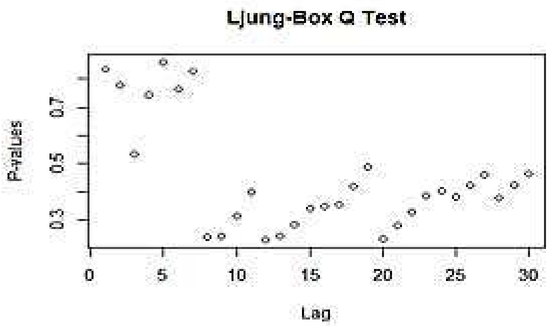

Fig 7: Ljung-Box Q plot for Murrah

Table 2: MAE, RMSE, R^2, MAPE and CE values of Crossbred, Sahiwal and Murrah

Performance indices	Crossbred		Sahiwal		Murrah	
	ARIMA (0,1,1) $(1,1,1)_{12}$		ARIMA (1,1,0) $(0,1,1)_{12}$		ARIMA (0,1,1) $(2,1,0)_{12}$	
	Training	Testing	Training	Testing	Training	Testing
RMSE	0.585	0.664	0.469	0.687	0.330	0.546
MAE	0.416	0.393	0.370	0.414	0.250	0.301
R^2	0.580	0.612	0.649	0.683	0.819	0.725
MAPE	4.04	4.18	6.68	6.01	4.69	4.90
CE	0.654	0.557	0.766	0.620	0.755	0.722

deseasonalization, the data are used again to test for stationarity using unit root tests.

The calculated values of the Augmented Dickey-Fuller (ADF) test for crossbred, Sahiwal and Murrah are 6.84e-010, 4.14e-010 and 2.76e-016, respectively, which are very small than the critical value of 0.05. Also, the critical value of the Kwiatkowski-Phillips-Schmidt-Shini test (KPSS) is 0.462 at the 5% significance level. The calculated p-values are 0.0162417, 0.0583065, and 0.0359997 for crosses, Sahiwal, and Murrah, respectively, which are smaller than the critical range, therefore the null hypothesis is accepted and the critical value of the P-P test is -2.886074 at the 5% significance level. The calculated values of Philips-Perron (PP) test statistics are -17.5211, -15.5328 and -16.4597 for crosses, Sahiwal and Murrah animals respectively, which are not in the critical range. Consequently, the null hypothesis is rejected, which means that the data are stationary.

Tentative models based on the seasonally differenced and non-trended series can be found. Figures 2, 3, and 4 show the ACF and PACF plots of the seasonally differenced and trendless series for crossbred, Sahiwal, and Murrah animals, respectively.

After carefully examining the ACF and PACF plots, the following preliminary models with their Akaike's information criterion (AIC) values were determined. The model with the lowest AIC value is selected and considered the best model.

It is evident from the Table 1 that ARIMA (0,1,1) $(1,1,1)_{12}$ for Crossbred cows, ARIMA (1,1,0) $(0,1,1)_{12}$ for Sahiwal cows and ARIMA (0,1,1) $(2,1,0)_{12}$ for Murrah buffaloes have the minimum value of AIC. Therefore, the ARIMA (0,1,1) $(1,1,1)_{12}$ for Crossbred cows, ARIMA (1,1,0) $(0,1,1)_{12}$ for Sahiwal cows and ARIMA (0,1,1) $(2,1,0)_{12}$ for Murrah buffaloes are found to be the best models for the study.

For the diagnostic check, the Box-Ljung test is applied to the residuals of a fitted ARIMA model, not the original series, and in such applications, the hypothesis being tested is that the residuals from the ARIMA model have no autocorrelation. The p-values for the Ljung-Box Q test all are well above 0.05, indicating "non-significance" (Figures 5, 6 and 7).

The regeneration performance of the ARIMA model was evaluated by the statistical parameters such as mean and standard deviation. The (mean, standard deviation) of regenerated series of Crossbred, Sahiwal and Murrah were found to be (10.37, 0.66), (5.56, 0.63) and (5.40, 0.72) respectively. These values are close to the (10.38, 0.90), (5.59, 0.79) and (5.39, 0.77) of Crossbred, Sahiwal and Murrah respectively of the historical series.

Qualitative evaluation

The observed and predicted values of Milk Yield of Crossbred, Sahiwal and Murrah of IDF Nagla, Pantnagar using the selected ARIMA model for the training period of eight years (2011-2018) are shown in Figures 8, 9 and 10 for Crossbred, Sahiwal and Murrah respectively. Similarly, the observed and predicted values for the testing period of two years (2019-2020) are also presented in the form of bar graphs in Figures 11, 12 and 13 for Crossbred, Sahiwal and Murrah respectively. Based on these figures, it is clear that there is good agreement between the predicted and

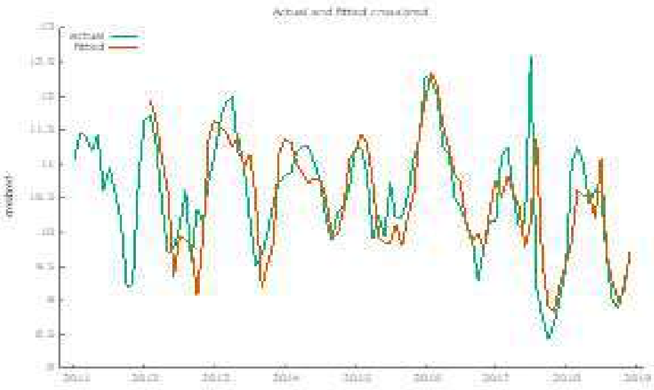

Fig 8: Actual vs Fitted Milk Yield data of Crossbred Cows

Fig 9: Actual vs Fitted Milk Yield data of Sahiwal Cows

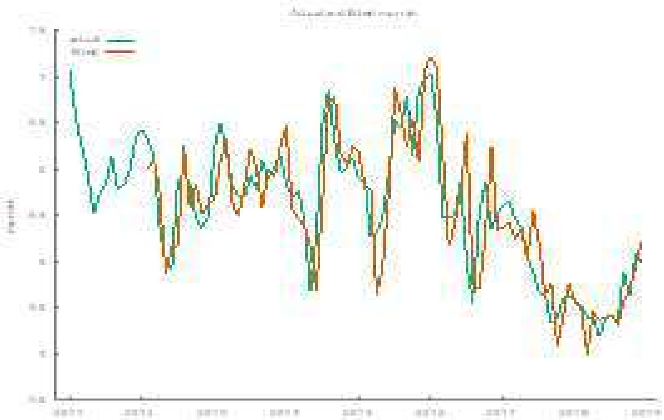

Fig 10: Actual vs Fitted Milk Yield data of Murrah Buffaloes

observed Milk Yield data, and the overall shape of the predicted Milk Yield data of Crossbred is similar to that of the observed Milk Yield.

Quantitative evaluation

For SARIMA based milk yield prediction model, the values of Root Mean Square Error (RMSE), Coefficient of determination (R^2), Coefficient of efficiency (CE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) for Pantnagar during calibration (2011-2018) and validation (2019-2020) periods for the selected models are close to each other.

The developed models were used for the forecasting of monthly Milk Yield data of two years ahead i.e., for the years 2019 and 2020 (Figures 14, 15 and 16).

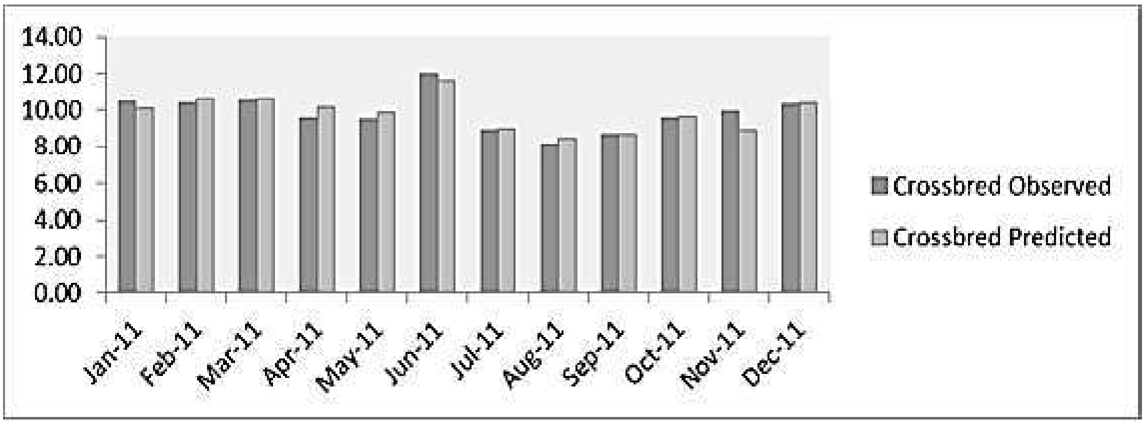

Fig 11: Mean monthly values of historical and regenerated data of Crossbred cow

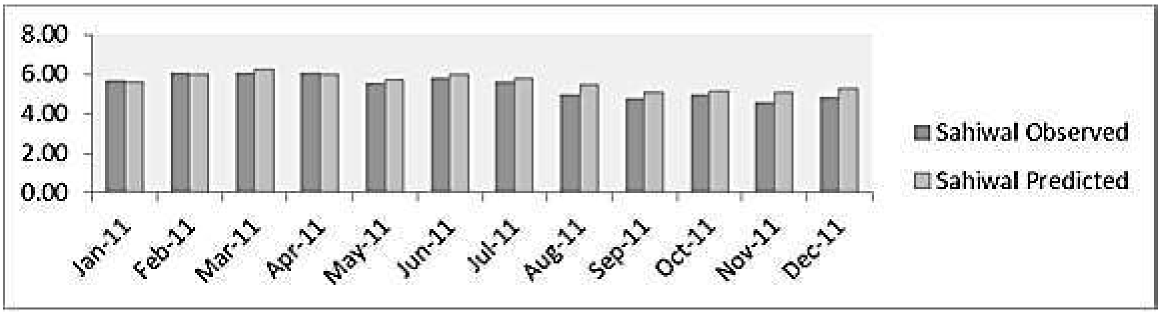

Fig 12: Mean monthly values of historical and regenerated data of Sahiwal cow

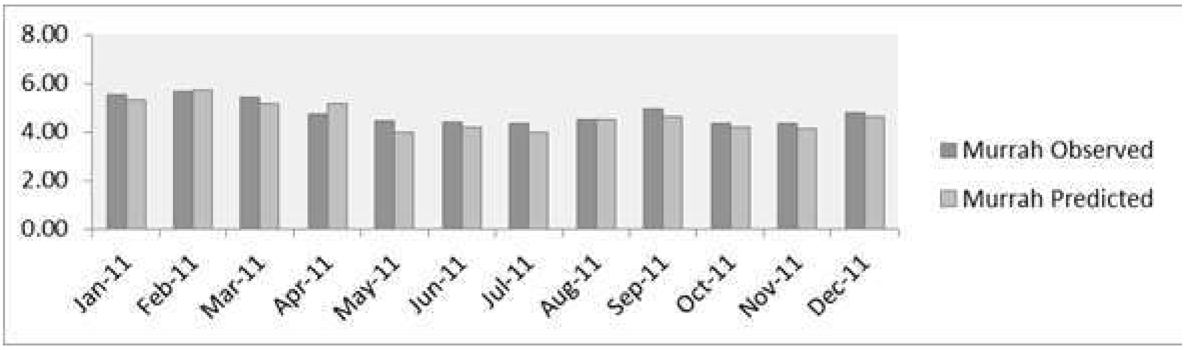

Fig 13: Mean monthly values of historical and regenerated data of Murrah buffalo

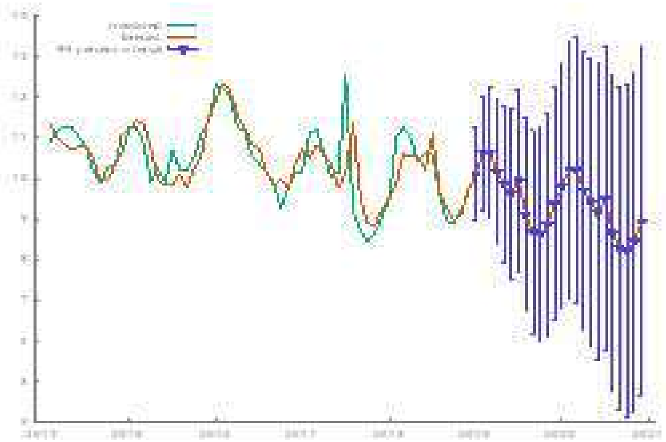

Fig 14: forecasts monthly Milk Yield for the years 2019 and 2020 for Crossbred Cows

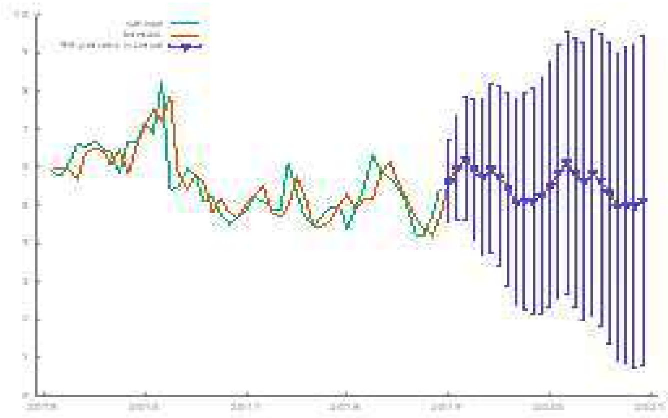

Fig 15: forecasts monthly Milk Yield for the years 2019 and 2020 for Sahiwal Cows

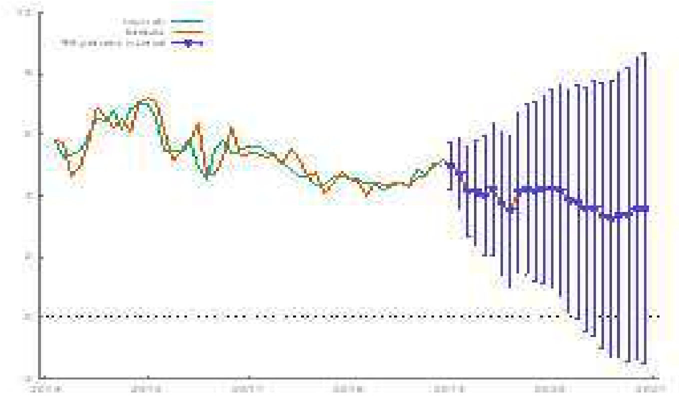

Fig 16: forecasts monthly Milk Yield for the years 2019 and 2020 for Murrah Buffaloes

Conclusions

The Auto-Regressive Integrated Moving Average (ARIMA) modelling was carried out by using GRETL, SPSS and Eviews software. The trend component of this time series was confirmed by the Mann-Kendall test. Therefore, the long-term value of the trend component was taken as the unit for further analysis. However, the milk yield data was still non-stationary due to seasonality, as the data had a 12-month periodicity and seasonal differencing (D=1) was required to achieve stationarity.

The seasonally adjusted data were used to check stationarity using unit root tests such as Augmented Dickey Fuller (ADF), Kwiatkowski-Phillips-Schmidt-Shin (KPSS) and Philips-Perron (PP) tests, which showed that the data were stationary. Then, after a thorough examination of the ACF and PACF plots, some tentative models with their Akaike Information Criterion (AIC) values were identified. The models with the lowest AIC value were selected and considered the most appropriate. Thakur and Gupta (2020) used ARIMA time series (p,d,q) to estimate monthly milk production over time and found that ARIMA model (6,0,2) was the best fitted model for prediction of monthly milk yield. However, for the present study we have found different models for different breeds and animals. ARIMA (0,1,1) $(1,1,1)_{12}$ for Crossbred cows, ARIMA (1,1,0) $(0,1,1)_{12}$ for Sahiwal cows and ARIMA (0,1,1) $(2,1,0)_{12}$ for Murrah buffaloes had the lowest AIC value and were considered the most suitable for the study. To verify the applicability of the models, monthly milk yield data for the period 2011-2018 were regenerated. The model also used for short term forecasting or prediction of monthly milk yield for the years 2019-2020.

As discussed above, Taye et al. (2021) estimated trends, fitted models and made predictions, hence, found that milk production was time-dependent and that the average amount of milk produced was decreasing. Similarly for the present study, the milk production for the predicted years (2019-2020) is found to be decreasing with the time as the milk production was time-dependent. This agrees with the findings of Taye et al. (2021).

The performance of developed models was evaluated qualitatively and quantitatively. The quantitative performance of developed models was assessed using statistical indices such as Root Mean Square Error (RMSE), Coefficient of Determination (R^2), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

Acknowledgement

The authors are thankful to the employees and In-charge, IDF, Pantnagar for their help in providing the necessary data for the purpose. The authors are also indebted to the learned referees and Editorial Board of the Indian Journal of Dairy Science for their valuable comments and suggestions which has helped us in improving the quality of this manuscript.

References

Australian Bureau of Statistics. *Time Series Analysis: The Basics.* April 06, 2021

https://www.abs.gov.au/websitedbs/d3310114.nsf/home/time+series+analysis:+the+basics

Belhekar S, Dash S (2016) Role of dairy industry in rural development in India. Indian J Res 5(11): 509-510

Berryman D, Bobée B, Cluis D, Haemmerli J (1988) Nonparametric tests for trend detection in water quality Time Series. J American Water Res Assoc 24(3): 545-556

Box–Jenkins method April 08 (2021) https://en.wikipedia.org/wiki/Box%E2%80%93 Jenkins_method

Department of Statistics *A brief history of time series analysis.* March 21 (2021) https://www.statistics.su.se/english/research/time-series-analysis/a-brief-history-of-time- series-analysis-1.259451

Deshmukh SS, Paramasivam R (2016) Forecasting of milk production in India with ARIMA and VAR time series models. Asian J Dairy and Food Res 35(1): 17-22

Etuk EH, Mohamed TM (2014) Time series analysis of monthly rainfall data for the Gadaref rainfall station, Sudan, by SARIMA methods. Int J Scientific Res Knowledge 2(7): 320

Kamble et al. (2020) "Growth and performance of dairy sector in India". FnBnews.com. April 16 (2020). http://www.fnbnews.com/Top-News/growth-and-performance-of-dairy-sector-in- india-54646

Mann-Kendall Test for Monotonic Trend. April 04, (2021) https://vsp.pnnl.gov/help/vsample/design_trend_mann_kendall.htm

Prabhakaran S (2014) Augmented Dickey Fuller Test (ADF Test)—Must Read Guide. Machine Learning Plus.

Pranovich M (2021) "Modeling non stationary variables". Joint Vienna Institute/IMF ICD. April 07 (2021). https://en.ppt-online.org/48550

Sahu PK, Das AK (2009) Agriculture and Applied Statistics-II. New Delhi, Kalyani Publishers. 50 p.

Suresh KK, Priya SK (2011) Forecasting sugarcane yield of Tamilnadu using ARIMA models. Sugar Tech, 13(1): 23-26.

Taye BA, Alene AA, Nega AK, Yirsaw BG (2021) Time series analysis of cow milk production at Andassa dairy farm, west Gojam zone, Amhara region, Ethiopia. Modeling Earth Sys Env 7(1): 181-189.

Thakur R, Gupta RK (2020) Prediction of milk yield using ARIMA. J Entomol Zool Stud 8(3): 1119-1123

Unit Root Tests. April 06, (2021) http://sfb649.wiwi.hu-berlin.de/fedc_homepage/xplore/tutorials/sfehtmlnode53.html#dfstatistik