



Modeling and forecasting of milk production in different breeds in Turkey

HARUN YONAR¹, AYNUR YONAR², PRADEEP MISHRA^{3✉}, MOSTAFA ABOTALEB⁴,
ABDULLAH MOHAMMAD GHAZI AL KHATIB⁵, TATIANA MAKAROVSKIKH⁴ and MUSTAFA CAM⁶

Selçuk University, Konya, Turkey

Received: 27 July 2021; Accepted: 12 October 2021

ABSTRACT

Identification of milk production is one of the key activities for the Turkish economy in terms of providing economic income for dairy smallholders and meeting consumer demands. This study aims to predict milk production in Turkey using various time series models which are BATS, TBATS, Holt's Linear Trend, ARIMA models, and NNAR. Yearly data from 1991 to 2019 on Milk Production is used in this study. The best time series model is selected from the testing data set (2015 to 2019) based on the lowest Mean Absolute Percentage Error (MAPE). The results of this study showed that the best predicts are obtained for Culture purebred milk production by an ARIMA (1,2,1) model, for Crossbreed milk production by a TBATS Model, for Indigenous milk production by ARIMA (0,2,0) model for total milking cows' production by Holt's Linear Trend model. Furthermore, these models forecasted an increasing trend in the production of milk from 2020 to 2025. The percentage increases for culture purebred milk, Crossbreed milk production, and Indigenous milk production from 2020 to 2025 are projected to be 40, 20.9 and 10.9%, respectively. Overall, the total milking cows' production is projected to increase by 25% in 2025.

Keywords: ARIMA, Forecasting, Holt's Linear Trend, Milk production, TBATS

Milk is a source of many essential components which play an indispensable role in body development and health. One of the most important subsectors of livestock production in Turkey, milk is consumed directly or processed into different products such as butter, cheese, powder, kefir, buttermilk, ice cream, or yogurt. According to FAO, about 81% of global milk production is obtained from cow milk. Annually, current milk production in Turkey is about 23 million tonnes and nearly 21 million tonnes of those are obtained from milking cows. Turkey has been a global dairy leader, ranking third in Europe and tenth in the world in terms of milk production.

The dairy industry was reported to have comprised at least 18% of all Turkish agriculture sectors, and small-scale dairy holders have comprised the majority of labour forces. Most of the owners hold fewer than 10 cattle due to inadequate knowledge of intensive farming (Krdar 2017).

Present address: ¹Faculty of Veterinary Medicine, Department of Biostatistics, Selçuk University, Konya, Turkey. ²Faculty of Science, Department of Statistics, Selçuk University, Konya, Turkey. ³College of Agriculture, Powarkheda, Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur, Madhya Pradesh. ⁴Department of System Programming, South Ural State University, Chelyabinsk, Russia. ⁵Department of Banking and Insurance, Faculty of Economics, Damascus University, the Syrian Arab Republic. ⁶Veterinary Faculty, Department of Animal Science, Selçuk University, Konya, Turkey. ✉Corresponding author email: pradeepjnkvv@gmail.com

To compare worldwide cow milk production, which has reached 524 MT in 2019 with the 2018–19 years, significant improvement in annual cow milk production has been noted over a 3% growth rate from the previous year in Turkey. Identification of milk production is one of the key activities for the Turkish economy to provide economic income for dairy smallholders and meet consumer demands. With a growing population, an outbreak of pandemic disease, and refugee migrations from undeveloped countries, milk consumption is expected to increase further in the coming years. Turkey should take critical political steps to determine short- and long-term future milk production in order to meet rising consumer demand. Therefore, forecasting of milk production, particularly cow milk, plays an essential role for policymakers in the regulation of agricultural policies. There are various studies in the literature on modeling and forecasting of milk production for different countries. Using the ARIMA Model, Mishra *et al.* (2020) investigated milk production in major states of India. Mishra *et al.* (2020) conducted a study interested in modelling and forecasting of milk production in Chhattisgarh and India with different Autoregressive Integrated Moving Average (ARIMA) models. Mishra *et al.* (2021) also studied modelling and forecasting of milk production in SAARC countries and China. Devi *et al.* (2021) investigated the future milk production prospects in India for various animal species using time series models. This study aims to predict milk production in Turkey with various time series models which

are BATS, TBATS, Holt's Linear Trend, ARIMA models, and NNAR.

MATERIALS AND METHODS

In the present investigation, data were collected for different breeds of milk production for Turkey from 1991 to 2019 (www.fao.org). In present study different time series models were used for projection purpose.

ARIMA Model

ARIMA models are the most widely used statistical models for time series forecasting, this is done by describing the autocorrelation in the data (Box *et al.* 2015). These models are divided into three parts, according to their nomenclature (Auto Regressive – Integrated – Moving Average) (p, d, q).

Autoregressive (p) refers to predicting a variable using a linear set of its preceding values, the model of order p can written as:

$$y_t = c + \beta_p y_{t-p} + \varepsilon_t \quad \dots (1)$$

where, β_q parameters of model; q, lag order of the moving average; ε_t , error term. Integrated (d) refers to the degree of stationarity of a variable that is determined using ADF test.

Moving average (q) uses past forecast errors in in regression. The equation will be in the form:

$$y_t = c + \varepsilon_t + \beta_q \varepsilon_{t-q} \quad \dots (2)$$

where β_q , parameters of model; q, lag order of the moving average; ε_t , error term.

Whereas (d) is determined by ADF test, (p) and (q) are determined by the autocorrelation function $\rho(p)$, and the partial autocorrelation function $\rho(p)$, which are given according to the following:

$$\rho(p) = \frac{\text{Cov}(y_t, y_{t-p})}{\sigma^2} \quad \dots (3)$$

$$(\rho(p-1) \rho(p-2) \dots \rho(0)) (\beta_q) = R(p) \quad \dots (4)$$

There is a difference between the TBATS model and dynamic harmonic regression in that the seasonality is allowed to change slowly over time in a TBATS model. (Both the BATS and TBATS models allow for a varied seasonality. TBATS is a (real improvement) modification of BATS that allows multiple seasonal incorrect cycles. TBATS has the following Equation (Livera *et al.* 2011, Hyndman and Athanasopoulos 2018)).

Equation (1) is a Box-Cox transformation

$$Y_t^{(\omega)} = \begin{cases} \frac{y_t^{(\omega)} - 1}{\omega} & \omega \neq 0 \\ \log y_t & \omega = 0 \end{cases} \quad \dots (5)$$

Equation (6) represents the seasonal M pattern

$$Y_t^{(\omega)} = l_{t-1} + \phi b_{t-1} + \sum_{i=1}^T S_{t-m_i}^{(i)} + d_t \quad \dots (6)$$

Equations (7), (8) and (9) are global trends and local trends

$$l_{t-1} + \phi b_{t-1} + \alpha d_t \quad \dots (7)$$

$$b_t = \phi b_{t-1} + \beta d_t \quad \dots (8)$$

$$S_t^{(i)} = S_{t-m_i}^{(i)} + \gamma_i d_t \quad \dots (9)$$

Equation (10) is the error modeled by ARMA

$$d_t = \sum_{i=1}^T \phi_i d_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad \dots (10)$$

where $m_1, \dots, m_T$ denote that seasonal period; l_t and b_t represent to the level and trend of components of the time series at time t; $s_t^{(i)}$ represents the i th seasonal component at time t; d_t denotes an ARMA (p, q) process; ε_t Gaussian white noise process when the mean equal to zero and constant variance σ^2 ;

Trigonometric exponential smoothing models for seasonal data (Taylor 2003).

$$s_t^{(i)} = \sum_{j=1}^{k_i} a_{j,t}^{(i)} \cos(\lambda_j^{(i)} t) \quad \dots (11)$$

$$a_{j,t}^{(i)} = a_{j,t-1}^{(i)} + k_1^{(i)} d_t \quad \dots (12)$$

$$\beta_{j,t}^{(i)} = \beta_{j,t-1}^{(i)} + k_2^{(i)} d_t \quad \dots (13)$$

where $k_1^{(i)}$ and $k_2^{(i)}$ are the smoothing parameters $\lambda_j^{(i)} = 2\pi j/m_i$. This is an extended, modified single sources of error version of single seasonal multiple source of error representation suggested by (Hannan *et al.* 1970) and is equivalent to index seasonal approaches when $k_i = m_i/2$ for even values of m_i and when $k_i = (m_i-1)/2$ for odd values of m_i but most seasonal terms will require much smaller values of k_i , thus reducing the number of parameters to be estimated.

ETS Model

Whereas ARIMA model describes autocorrelation in the data, exponential smoothing model (ETS) are based on describing the trend in the data, which was suggested by (Holt 1957). ETS models are a systematic development in which exponential smoothing models (ETS) are combined into a nonlinear dynamic model. Analysis of these models using state-space based likelihood calculations, with support for model selection and calculation of forecast standard errors (Hyndman *et al.* 2002). Interested in the model in three main components of time series: trend (T), seasonal (S), error (E). Reflects the trend term of the long-term movement of time series, and the error term is the unpredictable component of the time series. In our case, do not care about the seasonal term because the data annual. The components we need are combined in our model, in various additive and multiplicative combinations to produce y_t . We have additive model $y_t = T + E$ or multiplicative model like $y_t = T.E$ where the individual components of the model are given as follows:

E [A, M]
T [N, A, M, AD, MD]
S [N, A, M]

where N, none; A, additive; M, multiplicative; AD, additive dampened and MD, multiplicative dampened (damping uses an additional parameter to reduce the impacts of the trend over time) the models that we are interested in estimating State space equations for each of the models in the Holt's nonlinear

Trend	Additive error models	Trend	Multiplicative error models
N	$y_t = l_{t-1} + \varepsilon_t$ $l_t = l_{t-1} + \alpha \varepsilon_t$	N	$y_t = l_{t-1} (1 + \varepsilon_t)$ $l_t = l_{t-1} (1 + \alpha \varepsilon_t)$
A	$y_t = l_{t-1} + b_{t-1} + \varepsilon_t$ $l_t = l_{t-1} + b_{t-1} + \alpha \varepsilon_t$ $b_t = b_{t-1} + \beta \varepsilon_t$	M	$y_t = (l_{t-1} + b_{t-1}) (1 + \varepsilon_t)$ $l_t = (l_{t-1} + b_{t-1}) (1 + \alpha \varepsilon_t)$ $b_t = b_{t-1} + \beta (l_{t-1} + b_{t-1}) \varepsilon_t$
AD	$y_t = l_{t-1} + \phi b_{t-q} + \beta \varepsilon_t$ $l_t = l_{t-1} + \phi b_{t-q} + \alpha \varepsilon_t$ $b_t = \phi b_{t-q} + \beta \varepsilon_t$	MD	$y_t = (l_{t-1} + \phi b_{t-1}) (1 + \varepsilon_t)$ $l_t = (l_{t-1} + \phi b_{t-1}) (1 + \alpha \varepsilon_t)$ $b_t = \phi b_{t-1} + \beta (l_{t-1} + \phi b_{t-1}) \varepsilon_t$

Mishra *et al.* 2021.

can be written (after selecting S [N]) in the following Table.

where parameters: α smoothing factor for the level, β smoothing factor for the trend, ϕ : damping coefficient. And initial states l initial level components, b initial growth components, which is estimated as part of the optimization problem.

Performance indicators

To compare the prediction performance of the three models used, we first test the validity of the model by calculating mean absolute percentages error (MAPE) between the estimated data and the actual data during the period (2015–2019):

$$\frac{1}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t - y_t}{y_t} \right| \times 100 \quad \dots (14)$$

Then we evaluate the performance of the model by calculating root mean square error (RMSE) and (MAPE) between the estimated data and actual data during the period (1961–2015):

$$\sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}} \quad \dots (15)$$

where $\hat{y}_t$: the forecast value, y_t : the actual value, n : number of fitted observed. The last stage is to predict the pulses production for the countries of the study sample until 2027, the model that has the least values of (RMSE–MAPE) is the best, and the uncertainty is included in the expectations 95% prediction interval is given by:

$$\widehat{y_{T+h}} \pm Z_{\alpha/2} \sqrt{\hat{\sigma}_h} \quad \dots (16)$$

NNAR Model

Neural network autoregressive are statistical models that allow complex nonlinear relationships to predict a variable using its lagged values. Where lagged values of the time series can be used as inputs to a neural network (Hyndman *et al.* 2012) previously suggested this method. These models are distinguished from ARIMA models by the presence of a hidden layer, in which the linear weighted input is modified by a nonlinear function before it is output:

$$N_i = \beta_i + \sum_{j=1}^i \omega_j y_i \quad \dots (17)$$

In the hidden layer, this is modified using a nonlinear function:

$$f(y) = \frac{1}{1 + e^{-y}} \quad \dots (18)$$

where β_i and ω_i parameter of model are learned from the data. This model can be written as NNR(p, k) where p lagged input and k nodes in the hidden layer. Model is neural network with observations (y_{t-p}) used as inputs for forecasting the output y_t , and with k neurons in the hidden layer, with neglecting the effect of seasonality because the data is annual. The optimal number of lag p as (p, q) in ARIMA model is chosen using akaike information criterion (AIC), which is given as follow:

$$-2 \log L(\hat{\theta}) + 2k \quad \dots (19)$$

where $\hat{\theta}$ maximum value of the likelihood function. We remind that this model does not assume restriction about the stationary and therefore the random part is included in the predictions.

BATS, TBATS, Holt's Linear Trend, and ARIMA Models: BATS, TBATS, Holt's Linear Trend, ARIMA Models, and NNAR were fitted by using R software. Dataset is yearly data from 1991 to 2019 about Milk Production, we used training data from 1991 to 2014 and testing last 5 years from 2015 to 2019.

RESULTS AND DISCUSSION

The descriptive statistics for cases of milk production are tabulated in Table 1, which show that in Turkey from 1991 to 2019, purebred milk production increased during the period from 1,913,438 tonnes to 155,986,883 tonnes. The mean value of cultured purebred milk production was 54,150,108 tonnes. The Kurtosis value is 2.5, which indicates the data follows a platykurtic distribution, which means the number of outliers will be small followed by a positive value of skewness (0.80), which indicated there is some possibility of growth in the culture of purebred milk

Table 1. Descriptive statistics of milk production from 1991 to 2019

Case of milk production	Mean	Minimum	Maximum	Standard Deviation	Skewness	Kurtosis
Culture purebred	54,150,108	1,913,438	155,986,883	45,011,874	0.8071815	2.504330
Crossbreed	70,564,755	4,188,398	149,176,001	42,749,861	0.1958773	1.932994
Indigenous	26,319,341	2,514,576	43,654,989	12,572,076	-0.2966318	1.866264
Total milking cows	151,034,203	8,616,412	348,817,873	99,296,619	0.4103277	2.069020

Table 2. BATS model fitted for milk production for training data from 1991 to 2014

Case of milk production	Model	*Box-Cox transformation (Lambda)	Smoothing parameter		Damping parameter for trend	ARMA coefficients		Prediction error	
			Alpha	Beta		AR coefficients	MA coefficients	Sigma	AIC
Culture purebred	BATS(1, {0,0}, 1, -)	1	0.6656	1.6706	1	-	-	472044.3	711.3850
Crossbreed	BATS(1, {0,0}, 1, -)	1	0.8384	1.3843	1	-	-	303847.1	690.2387
Indigenous	BATS(1, {0,0}, 0.98, -)	1	0.9785	-0.0295	0.980449	-	-	406208.1	706.1751
Total milking cows	BATS(1, {0,0}, 1, -)	1	0.8091	1.3164	1	-	-	875131.1	741.0155

Table 3. TBATS model fitted for milk production for training data from 1991 to 2014

Case of milk production	Model	*Box-Cox transformation (Lambda)	Smoothing parameter			Damping parameter for trend	Prediction error		
			Alpha	Beta	Gamma-1 values		Gamma-2 values	Sigma	AIC
Culture purebred	TBATS(1, {0,0}, 1, {<6,2>})	1	-0.02	0.39	0.00	1.00	1,856,179.00	789.11	
	TBATS(1, {0,0}, 1, {<6,2>})	1	1.07	0.78	0.10	1.00	500,364.40	726.18	
	TBATS(1, {0,0}, 1, {<6,2>})	1	1.03	0.23	-0.01	1.00	422,319.30	718.04	
	TBATS(1, {0,0}, 1, {<6,2>})	1	1.59	0.47	1.53782e	1.00	2,127,571.00	795.66	

production. Crossbreed milk production increased during the period from 4,188,398 tonnes to 149,176,001 tonnes. The mean value of Crossbreed milk production is 70,564,755 tonnes and Kurtosis value 1.9, indicating that the data follows a platykurtic distribution, meaning that the number of outliers will not be large followed by a positive value of skewness (0.19) which is between -0.5 and 0.5, so the distribution is approximately symmetric. Indigenous milk production has increased during the period from 2,514,576 tonnes to 43,654,989 tonnes indicating the data follows a platykurtic distribution. The total milking cows' production increased during the period from 8,616,412 tonnes to 348,817,873 tonnes. The average total milking cows' production is 151,034,203 tonnes and Kurtosis value 2.06, indicating the data follows a platykurtic distribution. Followed by a positive value of skewness (0.41) which is between -0.5 and 0.5, which means the distribution is approximately symmetric.

From Table 2, the culture purebred milk production: BATS is the best-suited model (1, {0,0}, 1, -), in this model, Box-Cox transformation =1, the order of ARMA error is (0, 0), the damping parameter = 1 (essentially doing nothing). Crossbreed milk production: BATS is the best-suited model (1, {0,0}, 1, -), in this model, Box-Cox transformation =1, the order of ARMA error is (0, 0), the damping parameter = 1 (essentially doing nothing). Indigenous milk production: BATS is the best-suited model (1, {0,0}, 0.98, -), in this model, Box-Cox transformation =1, the order of ARMA error is (0, 0), the damping parameter =0.98. Total milking cows' production: BATS is the best-suited model (1, {0,0}, 1, -), in this model, Box-Cox transformation =1, the order of ARMA error is (0, 0), the damping parameter = 1 (essentially doing nothing). From Table 3, the Culture purebred milk production: TBATS is the best-suited model (1, {0,0}, 1, {<6,2>}) in this model, Box-Cox transformation =1, (doing nothing), the order of ARMA error is (0, 0), the damping parameter = 1 (essentially doing nothing). Crossbreed milk production: TBATS is the best-suited model (1, {0,0}, 1, {<6,2>}) in this model, Box-Cox transformation =1 (doing nothing), the order of ARMA error is (0, 0), the damping parameter = 1 (essentially doing nothing). Indigenous milk production TBATS is the best-suited model (1, {0,0}, 1, {<6,2>}) in this model, Box-Cox transformation =1 (doing nothing), the order of ARMA error is (0, 0), the damping parameter = 1 (essentially doing nothing). The total milking cows' production: TBATS is the best-suited model (1, {0,0}, 1, {<6,2>}) in this model, Box-Cox transformation =1, (doing nothing), the order of ARMA error is (0, 0), the damping parameter = 1 (essentially doing nothing). By comparing Table 2 with Table 3, we found that forecasting accuracy by the BATS model outperformed the forecasting accuracy of TBATS Model, because the values of the sigma, AIC for BATS Model were lower than values of the sigma, AIC for TBATS Model at all series.

The level smoothing parameter was denoted by Alpha, and the trend smoothing parameter is denoted by Beta, a

Table 4. Holt's Linear Trend model fitted for milk production for training data from 1991 to 2014

Case of milk production	Box-Cox transformation (Lambda)	Smoothing parameters		Initial states		Sigma	AIC
		Alpha	Beta	L	B		
Culture purebred	0.5568	0.6486	0.6425	4,367.4945	2,979.9086	459.2392	376.1170
Crossbreed	0.8934	0.9119	0.9119	299,275.8654	762,443.8882	68,234.2100	616.1712
Indigenous	1.3122	0.9999	0.8494	123,626,109.0125	304,405,299.6126	36,233,855.0000	917.3618
Total milking cows	0.7957	0.9006	0.5897	267,057.2479	249,331.8722	41,401.5400	592.1891

and β are constrained to 0–1 with higher values meaning faster learning and lower values meaning slower learning. From Table 4, it can be seen that the values of α and β for all series were close to 1 which means a fast learning in the year-to-year milk production in all cases. In Table 5, ARIMA (1, 2, 1) is seen as best fitted model for Culture purebred milk production, ARIMA (0,2,0) is seen as best fitted model for Crossbreed milk production, ARIMA (0, 2, 0) is seen as best fitted model for Indigenous milk production, ARIMA (1, 2, 0) is seen as best fitted model for Total milking cows' production. In Table 6, the best-fitted models on training data set (1991 to 2014), based on, lowest values of ME, RMSE, MAE, MPE, MAPE, MASE and ACF1, NNAR (1,5) model is the best model for all the series. In other words, the forecasting accuracy by the NNAR (1, 5) model is very high and outperforms the forecasting accuracy of the other models, because the most values of the accuracy criteria (RMSE, MAE, and ME, MASE) were lower than

Table 5. ARIMA Model fitted for Milk production for training data from 1991 to 2014

Case of milk production	Model	AR (1)	AR (2)	AR (3)	MA (1)
Culture purebred	ARIMA (1,2,1)	0.9536	–	–	–0.6052
Crossbreed	ARIMA (0,2,0)	–	–	–	–
Indigenous	ARIMA (0,2,0)	–	–	–	–
Total milking cows	ARIMA (1,2,0)	0.3494	–	–	–

the values of the accuracy criteria of other Models. In Table 7, the best-fitted models on testing data set (2015 to 2019), based on, lowest values of MAPE, were: ARIMA (1, 2, 1) Model is the best model for the series (Culture purebred

Table 6. BATS, TBATS, Holt's Linear Trend, ARIMA, and NNAR model fitted for milk production for training data from 1991 to 2014

Model	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
<i>Milk production in case of culture purebred (tonnes)</i>							
BATS	208,134.1000	472,044.3	335,703.3	2.4470	4.1778	0.0785	–0.1374
TBATS	867,638.0000	1,856,179.0	1,583,520.0	8.7220	12.9600	0.3702	0.5189
Holt's linear Trend	140,084.8000	575,005.8	464,152.0	–2.1530	4.2718	0.1085	0.5134
ARIMA(1,2,1)	83,732.9500	314,200.1	221,067.9	0.2290	0.6256	0.0517	–0.0388
NNAR(1,5)	–433.2704	200,296.2	162,468.5	–0.1470	0.7901	0.0380	0.2680
<i>Milk production in case of crossbreed (tonnes)</i>							
BATS	68,176.3300	303,847.1	231,739.0	–0.0411	0.8124	0.0478	–0.0501
TBATS	114,585.3000	500,364.4	422,334.4	0.1808	1.0930	0.0871	–0.0361
Holt's linear Trend	56,113.8200	387,193.7	288,267.9	–0.2794	1.5117	0.0595	0.0916
ARIMA(0,2,0)	99,908.1000	308,532.7	221,620.8	0.1939	0.4348	0.0457	0.2026
NNAR(1,5)	–660.3522	217,117.2	173,952.1	–0.0264	0.3529	0.0359	0.2820
<i>Milk production in case of indigenous (tonnes)</i>							
BATS	–11,150.2600	406,208.1	240,003.4	–1.9360	4.0470	0.1490	0.0060
TBATS	–126,118.3000	422,319.3	285,842.6	–2.0970	4.2440	0.1780	–0.0470
Holt's linear Trend	–13,833.5000	175,051.5	125,386.0	–0.0220	1.0180	0.0780	–0.0900
ARIMA(0,2,0)	–59,304.8400	171,109.3	114,846.5	–0.3660	0.5970	0.0710	–0.1010
NNAR(1,5)	246.6742	123,126.7	92,906.4	–0.0230	0.4540	0.0580	0.2710
<i>Milk production in case of total milking cows (tonnes)</i>							
BATS	303,016.2000	875,131.1	583,745.0	0.2530	1.3058	0.0544	–0.1249
TBATS	683,110.0000	2,127,571.0	1,800,999.0	0.9105	2.5656	0.1678	–0.3283
Holt's linear Trend	294,232.7000	1,331,912.0	1,050,027.0	–0.4398	2.9423	0.0978	0.2045
ARIMA(1,2,0)	229,552.2000	730,908.9	465,884.0	0.1595	0.4052	0.0434	–0.1804
NNAR(1,5)	2,593.3430	471,350.5	358,133.0	–0.0329	0.3745	0.0334	0.2315

Table 7. MAPE (%) for testing data last 5 years from 2015 to 2019

Case of milk production	BATS (%)	TBATS (%)	Holt's Linear Trend (%)	ARIMA (%)	NNAR (%)	Best model
Culture purebred	2.38	4.70	0.54	0.48	7.29	ARIMA (1,2,1)
Crossbreed	0.43	0.26	0.60	0.43	1.71	TBATS
Indigenous	2.29	3.11	0.99	0.91	1.10	ARIMA (0,2,0)
Total milking cows	0.86	2.70	0.59	0.79	1.97	Holt's Linear Trend

Table 8. Forecasting from 2020 to 2025 for milk production by using best models

Year	Culture purebred	Crossbreed	Indigenous	Total milking cows
2020	166,942,108.09	154,691,560.14	45,442,518	361,328,093.98
2021	179,571,425.99	161,012,969.75	46,429,219	379,670,939.15
2022	192,582,466.02	167,052,302.98	47,415,920	398,196,699.40
2023	205,957,527.64	173,342,117.13	48,402,621	416,900,291.11
2024	219,679,731.09	180,219,893.08	49,389,322	435,776,950.20
2025	233,732,979.33	187,065,637.50	50,376,023	454,822,201.57

milk production); TBATS Model is the best model for the series (Crossbreed milk production); ARIMA (0, 2, 0) Model is the best model for the series (Indigenous milk production); Holt's Linear Trend model is the best model for the series (Total milking cows' production).

In Table 8, using the best models of our study, we found that according to ARIMA (1, 2, 1), purebred milk production continues its upward trend in Turkey. The production of cultured milk will increase from 166,942,108.09 tonnes in 2020 to 233,732,979.33 tonnes in 2025. According to the TBATS Model, crossbreed milk production will rise from 154,691,560.14 tonnes in 2020 to 187,065,637.50 tonnes in 2025. According to the ARIMA (0, 2, 0) Model, indigenous milk production will rise from 45,442,518 tonnes in 2020 to 50,376,023 tonnes in 2025. According to Holt's Linear Trend Model, the total milking cows' production will rise from 361,328,093.98 tonnes in 2020 to 454,822,201.57 tonnes in 2025.

Forecasting

The data is divided into two sets of the year, which are: 1991 to 2014 as training data, and the data from 2015 to 2019 as testing data, and the data from 2020 to 2025 is used as out-of-sample forecast. The residuals of the chosen models were found stationary and white noise in all time series. The 95% confidence limit of predicted values from 2020 to 2025, using best-fitted models, is shown in the Table 8. The figures showed the predicted values lie within the 95% confidence intervals. All forecasted lines in the figures are close to the actual values, which emphasize the quality of the selected models. From the forecasted figures, it can be seen that milk production in Turkey would increase continually in all cases (culture purebred, crossbred, indigenous and total milk cows).

Milk production is a vital part of the international food system and it plays a main role in the sustainability of rural areas especially. In the forecast models, each model has a

different error rate. From testing data, the best models with the least MAPE error were: ARIMA (1,2,1) for the series of Culture purebred milk production, TBATS Model for the series of Crossbreed milk production, ARIMA (0,2,0) Model for the series of Indigenous milk production, Holt's Linear Trend model for the series of total milking cows' production. According to the best forecasting models, the Culture purebred milk production continues its upward trend in Turkey. This study can be useful for policymakers in understanding and developing plans for next years. In future studies, it can help make a production policy forecast which includes milk prices, supply, and demand of milk production, marketing of milk production, and export of milk production for all cases.

ACKNOWLEDGEMENTS

The work was supported by Act 211 of the Russian government, contract no. 02.A03.21.0011 and by the Russian Ministry of Science and Higher Education (Government order no: FENU-2020-0022).

REFERENCES

- Box G E P, Jenkins G M, Reinsel G C and Ljung G M. 2015. Time series analysis: Forecasting and control (5th ed). John Wiley & Sons, Hoboken, New Jersey.
- 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* 1–6.
- Dickey D and Fuller W. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* **49**: 1057–72.
- 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.
- De Livera A M, Hyndman R J and Snyder R D. 2011. Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*

- 106**(496): 1513–27.
- Hyndman R J, Koehler A B, Snyder R D and Grose S. 2002. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting* **18**(3): 439–54.
- Holt C E. 1957. Forecasting seasonal and trends by exponentially weighted averages (ONR Memorandum No. 52). Carnegie Institute of Technology, Pittsburgh, USA.
- Winters P R. 1960. Forecasting sales by exponentially weighted moving averages. *Management Science* **6**(3): 324–42.
- Hyndman R J and Athanasopoulos G. 2018. Forecasting: Principles and practice. OTexts.
- Hannan E J, Terrell R D and Tuckwell N E. 1970. The seasonal adjustment of economic time series. *International Economic Review* **11**(1): 24–52.
- Kirdar S S. 2017. An overview of the Turkish dairy sector. *Indian Journal of Dairy Science* **70**: 249–55.
- 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.
- 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* **54**: 912–17.
- 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, M G Al Khatib, Sardar I *et al.* 2021. Modeling and forecasting of sugarcane production in India. Sugar Tech 1–8. <https://doi.org/10.1007/s12355-021-01004-3>.