



## Comparative study of ANN and conventional methods in forecasting first lactation milk yield in Murrah buffalo

VIJAY KUMAR<sup>1</sup>, A K CHAKRAVARTY<sup>2</sup>, ANKIT MAGOTRA<sup>3</sup>, C S PATIL<sup>4</sup> and P R SHIVAHRE<sup>5</sup>

ICAR-National Dairy Research Institute, Karnal, Haryana 132 001 India

Received: 13 July 2018; Accepted: 4 April 2019

### ABSTRACT

Present investigation was undertaken to predict first lactation 305-day milk yield (FL305DMY) using monthly test day milk records. Under this study, multiple linear regression (MLR) and artificial neural network (ANN) approach were used. Effectiveness of both methods was also compared for prediction of FL305DMY in Murrah buffalo. The data on 3336 monthly test day milk yields records of first lactation pertaining to 556 Murrah buffaloes maintained at National Dairy Research Institute, Karnal; Central Institute for research on buffalo; Guru Angad Dev Veterinary and Animal Sciences University (GADVASU), Ludhiana and Choudhary Charan Singh Haryana Agricultural University (CCSHAU), Hisar were used in this study. In MLR study, it was observed that model 14 having four independent variable, i.e. FSP, TD2, TD4 and TD6 fulfilled most criteria such as highest R<sup>2</sup>, lowest MSE, lowest RMSE, lowest CP, lowest MAE, lowest MAPE, and lowest U value. In the present investigation, the accuracy of prediction obtained from ANN was almost similar to MLR for prediction of FL305DMY using monthly test day milk records in Murrah buffalo. The best ANN algorithm achieved 76.8% accuracy of prediction for optimum model, whereas the MLR explained 76.9% of accuracy of prediction of FL305DMY in Murrah buffalo. MLR method is simple as compared to ANN, hence MLR method could be preferred.

**Key words:** AIC, ANN, BIC, FL305DMY, MLR, Test-day milk yield

The aim of any animal breeder is to evaluate sire in younger age to minimise the cost of rearing. Mostly 305-day milk yield is used for evaluation of dairy animals in India. Dairy cattle evaluation using test-day milk yields (TDMY) has significant advantages over the 305-day milk yield (Mostert *et al.* 2006). The use of TDMY permits a more precise understanding of contemporary groups and associated environmental effects. In developing countries like India, there is inadequate milk recording facilities, and use of test day models would result in reduced cost of recording as we could have longer intervals between milk recording and less frequent collection of milk samples. The multiple linear regression (MLR) models are being extensively used in various disciplines including dairy science to predict milk production of dairy animals. In recent times, Artificial Neural Network (ANN) is also used in some

areas of animal genetics and husbandry, such as to predict swine daily gain in different ambient temperatures (Korthals *et al.* 1994), estimating meat quality (Brethour 1994), prediction of 305 day milk production from part lactation records (Lacroix *et al.* 1995), prediction and classification of dairy cows based on milk yield in one period (Salehi *et al.* 1998), detection of clinical disease (Yang *et al.* 1999), evaluation of physiological status of cows (Molenda *et al.* 2001), detection of mastitis in dairy cattle (López-Benavide *et al.* 2003) and prediction of slaughter value of bulls (Adamczyk *et al.* 2005).

Present investigation was undertaken to predict FL305DMY on the basis of first lactation traits by MLR and ANN approach and to compare their effectiveness for prediction in Murrah buffalo. These methods could be used as a tool for recognition of more producer buffaloes of high genetic merit as the parents of the next generation.

### MATERIALS AND METHODS

*Source of data:* The data on 3336 monthly test day milk yields records of first lactation pertaining to 556 Murrah buffaloes maintained at National Dairy Research Institute, Karnal; Central Institute for Research on Buffalo; Guru Angad Dev Veterinary and Animal Sciences University (GADVASU), Ludhiana and Choudhary Charan Singh Haryana Agricultural University (CCSHAU), Hisar were

Present address: <sup>1</sup>Assistant Professor (jaisvet@gmail.com), Department of Animal Genetics and Breeding, Pt. Deen Dayal Upadhyaya Veterinary University, Mathura, Uttar Pradesh. <sup>3,4</sup>Assistant Professor (dr.cspatil03@gmail.com, ankitoms@gmail.com), Lala Lajpat Rai University of Veterinary and Animal Sciences, Hisar, Haryana. <sup>5</sup>Assistant Professor (drpr06@gmail.com), Animal Husbandry and Dairying, Udai Pratap College, Varanasi, Uttar Pradesh. <sup>2</sup>Retired Principal Scientist (ak\_chakravarty@yahoo.co.in), Dairy Cattle Breeding Division.

used to predict first lactation 305-day milk yield (FL305DMY) using monthly test day milk records. The records of animals with lesser than 500 kg of milk production and lactation length lesser than 100 days were discarded due to their abnormal lactation. Out of all test days, test day records upto 6<sup>th</sup> TDMY were selected as 6<sup>th</sup> TDMY have highest genetic and phenotypic correlation with FL305DMY (Kumar *et al.* 2014). Other traits considered were age at first calving (AFC), and first service period (FSP) (Table1).

*Variance inflation factors (VIF):* When correlation exists among predictor's the standard error of predictors coefficients will increase and consequently the variance of predictor's coefficients are inflated. The VIF is a tool to measure and quantify how much the variance is inflated. To interpret the value of VIF the following rule is used in the Table 2. VIF can be calculated using the formula:

$$VIF = \frac{1}{1-R_i^2}$$

Where R<sub>i</sub><sup>2</sup> is the R<sup>2</sup>- value obtained by regressing the ith predictor on the remaining predictors. A variance inflation factor exists for each of the i predictors in a multiple regression model.

*Multiple regression analysis:* The multiple regression analysis was done as suggested by Draper and Smith (1987) with help of following model:

Model for prediction

$$Y_i = a + b_1X_1 + b_2X_2 + \dots + b_nX_n + e_i$$

where Y<sub>i</sub> is the variable to be predicted; a, b<sub>1</sub>, b<sub>2</sub>, ..., b<sub>n</sub> are unknown parameters to be estimated; X<sub>1</sub>, X<sub>2</sub>, ..., X<sub>n</sub> traits whose values are known; e<sub>i</sub>, random residual, NID (0, σ<sup>2</sup><sub>e</sub>)

*Criteria for predicting performance of model:* Fourteen models have been studied using four traits (Table 3). In order to judge the forecasting accuracy of a particular model or for evaluating and comparing different models, their

Table 1. Description of input and output variables used in the present study

Input Variable	Day of lactation	Output variable
TD1	5th day of lactation	305 days milk yield
TD2	35th day of lactation	
TD	3 65th day of lactation	
TD4	95th day of lactation	
TD5	125th day of lactation	
TD6	155th day of lactation	
AFC	-	
FSP	-	

Table 2. VIF rule followed in the study

VIF value	Conclusion
VIF = 1	Not correlated
1 < VIF ≤ 3	Moderately correlated
VIF > 3	Highly correlated

Table 3. Input sub-sets and traits involved in each set

Sub-set	Traits included
SET- 1	FSP
SET- 2	TD2
SET- 3	TD 4
SET- 4	TD6
SET- 5	FSP, TD2
SET- 6	FSP, TD4
SET- 7	FSP, TD6
SET- 8	TD2, TD4
SET- 9	TD2, TD6
SET- 10	FSP, TD2, TD4
SET- 11	FSP, TD2, TD6
SET- 12	FSP, TD4, TD6
SET- 13	TD2, TD4, TD6
SET- 14	FSP, TD2, TD4, TD6

relative performance, the following performance measures have been used.

*Akaike information criterion (AIC):* Akaike information criterion (AIC) is a tool for model selection (Akaike, 1974). It is a measure of the goodness of fit of an estimated statistical model.

$$AIC = n \times (\log \text{MSSe} / n) \times 2p$$

where,

n, no. of observations; MSSe, mean sum of squares due to error and p, no. of parameters in model.

Model with lowest AIC, considered as Optimal Model.

*Bayesian Information Criterion (BIC):* Bayesian information criterion (BIC) is a measure for model selection among a class of parametric models with dissimilar numbers of parameters developed by Schwarz (1978). It is strongly related to the Akaike information criterion and penalises additional parameters robustly than that of the Akaike information criterion. Optimum model is said to have low BIC value.

$$BIC = n \times \log (\text{MSSe} / n) + [k \times \log (n)]$$

where n, Sample size; k, no of free parameters to be estimated.

*Mallow's conceptual predictive value (MCP Value):* Mallow's conceptual predictive value is used for model selection in regression (Mallows 1973). The Cp statistic is defined as a criterion to assess fits when models with dissimilar numbers of parameters are being compared.

True Model (TM) i.e. model including all the independent traits.

Candidate Model (CM) i.e. model with different combination of independent traits.

Cp criterion compares Candidate Model with True Model.

$$C_p = p + \frac{(\text{MSSe of CM} - \text{MSSe of TM})}{\text{MSSe of TM}} \times (n-p)$$

where p, no. of parameters in CM; n, Sample size; MSSe, Mean sum of squares error.

If Cp > p, then the candidate model is too small i.e. missing some important independent traits. If Cp ≤ p, then candidate model is large enough i.e. all important

independent traits included in the model.

*Coefficient of determination ( $R^2$ ):* Coefficient of determination indicates that out of hundred per cent of variability of the dependent variable, how much variation was contributed by a set of independent variables and is expressed in terms of percentage.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

*Root mean square error (RMSE):* The root-mean-square error (RMSE) is a frequently used measure of the differences between values (sample and population values) predicted by a model or an estimator and the values actually observed.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

*Mean absolute error (MAE):* The mean absolute error measures average magnitude of the errors in a set of predictions, without considering their direction. It measures accuracy for continuous variables.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

*Mean absolute percentage error (MAPE):* The mean absolute percentage error (MAPE) is one of the most popular measures of the forecast accuracy due to its advantages of scale-independency and interpretability. MAPE is the average of absolute percentage errors (APE).

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100$$

*Theil's U-statistics (U):* Theil's U statistic is a relative accuracy measure that compares the forecasted results with the results of forecasting with minimal historical data. It also squares the deviations to give more weight to large errors and to exaggerate errors, which can help eliminate methods with large errors.

$$U = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i)^2} \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i)^2}}$$

Where,  $y_i$  is actual or observed value,  $\hat{y}_i$  is the predicted value,  $\bar{y}$  is the arithmetic mean of observed value and  $n$  is number of observation. For assessing forecast accuracy, it is desirable that the RMSE, MAE, MAPE and U-statistic should be close to zero and  $R^2$  should be close to unity.

*Selection of optimum model for prediction of FL305DMY in Murrah buffalo:* Optimum model(s) was developed for prediction of variables in Murrah buffalo using combination of AIC value, BIC value, MCp value,  $R^2$ , RMSE, MAE, MAPE and U-statistic for each model. Best model was also compared with models obtained using artificial neural network.

*Artificial neural network (ANN):* ANN model is principally an intelligent data processing system which learns the predictive capability mechanically from the information presented while training the network. A multilayer feed forward network with back propagation of error learning mechanism is mainly used neural network architecture and has shown outstanding effectiveness in dealing with functional approximation problems. Such neural network made up of input layer, hidden layer and output layer. Each layer has a definite role in execution of the ANN. In back propagation system, input data and the matching target are used to train a network until it can approximate a prediction function (Fausett, 1994). In the present study, a multilayer feed forward neural network with back propagation of error learning mechanism was constructed using R programme (R core team, 2013) to predict FL305DMY. The ANN was trained and simulated by means of resilient back propagation algorithms. The network was trained with training data set to get consistent results. The forecast capability was tested using a new dataset (test data). A properly trained ANN is likely to give reasonable output when presented with new inputs. The whole data set was divided at random into two subsets, viz. the training set consisting of 65% or 75% or 85% and testing subset comprising 35% or 25% or 15% of data. The training sets were used to train the neural network models and the testing sets were used to validate the models. The network was tested with one hidden layer (2, 3, 5, 7, 10 and 20 neurons) and two hidden layers (10:5, 5:3, 3:2 and 2:1 neurons). All four first lactation traits were used to predict FL305DMY by MLR as described by Draper and Smith (1987) and ANN and their effectiveness was compared from both the methods using  $R^2$ -value and root mean square errors.

## RESULTS AND DISCUSSION

Research workers often use models to approximate unknown relationship between a set of predictor variables and the response variable. They try different types of models, which explain the variability in the data in a better way. The main objective of the model building is to predict response variable using the predictor variables. Much of the researcher's effort is devoted to the estimation of model parameters, however it is desirable to have a model that is reasonably easy to construct and predicts well. Thus, the assessment of prediction performance of a model is critical and has practical importance. This is especially true for models with prediction as their primary objective.

*Development of optimum equation:* A total of 11 monthly test days records were collected with 30 days interval on 5<sup>th</sup> day, 35<sup>th</sup> day, 65<sup>th</sup> day, ... and 305<sup>th</sup> day of lactation. Along with test day records, AFC and FSP records were also collected. Out of all test days, a total of three test day records (2, 4 and 6 test day records) and FSP were selected using variance inflation factor (VIF) method to use as input variables for multiple linear regression method. The monthly test day milk yields were used to predict first

lactation 305-day milk yield (FL305DMY) by using multiple linear regression method. After removing TD3 and TD5 the VIF for all traits were below 3. AFC and TD1 were also removed to reduce number of input variables as on initial studies it was found that these variables had very poor influence on FL305DMY.

The models for prediction of FL305DMY using first lactation traits were developed and are presented in Table 4. The developed models were in different combinations of FSP, TD2, TD4 and TD6 traits. Fourteen models were developed for prediction of FL305DMY. It was observed that model 14 having four independent variables, i.e. FSP, TD2, TD4 and TD6 fulfilled most criteria such as highest R<sup>2</sup>, lowest MSE, lowest RMSE, lowest CP, lowest MAE, lowest MAPE, and lowest U value. AIC and BIC values were comparatively higher.

This may be due to increased number of parameters. Hence, model 14 was adjudged as the best optimum model for the prediction of FL305DMY. The data set was subjected to MLR and the linear regression equation developed is given below.

$$\text{FL305DMY} = 0.53X_1 + 55.314X_2 + 68.759X_3 + 126.765X_4 - 46.432$$

The accuracy of prediction from the above model was 76.9%. The R<sup>2</sup>-value of prediction was high suggesting that the relationship between the predictors and response variable is linear. Now-a-days, use of test day milk yields is receiving more importance for prediction of milk yield in dairy cattle. Prediction of FL305DMY using test day milk yields in an early stage of lactation with maximum accuracy is one of the criteria of selection for life time profit ability of dairy cows (Gandhi and Gurnani 1988, Kannan and Gandhi 2006, Dongre *et al.* 2012). Accuracy rate of judgment on high producing animals is essential, because feeding, breeding, maintenance costs etc can be minimized for best animals and also by wrong culling cows of high

Table 4. Estimation of intercept and regression coefficient (s) for FL305DMY in Murrah buffalo

Model	Variables	A	FSP	TD2	TD4	TD6
1	FSP	1668.484	0.759			
2	TD2	540.427		170.540		
3	TD4	379.546			196.469	
4	TD6	389.716				216.158
5	FSP, TD2	422.052	0.577	167.997		
6	FSP, TD4	279.683	0.583		193.926	
7	FSP, TD6	288.623	0.578			213.656
8	TD2, TD4	194.922		88.138	131.381	
9	TD2, TD6	145.112		83.210		158.675
10	TD4, TD6	127.913			99.468	145.433
11	FSP, TD2, TD4	104.687	0.547	86.439	130.247	
12	FSP, TD2, TD6	56.262	0.538	81.454		157.561
13	TD2, TD4, TD6	39.963		56.768	69.481	127.538
14	FSP, TD2, TD4, TD6	-46.432	0.530	55.314	68.759	126.765

genetic value, good sources of gene pool will be lost. In several countries, study of milk yield for 305 day lactation period is a basis for dairy animal genetic assessments. So, implementing statistical models for prediction of 305 day production in succeeding lactations from previous lactations or predicting total lactation yield from early records would be valuable.

Conventional models such as linear regression, multiple linear regression (MLR), stepwise multiple linear regression, partial least-squares regression, projection pursuit regression, logistic regression, etc. have been widely used as prediction tools for various real-life problems. Dongre *et al.* 2018 used multiple linear regression analysis for prediction of standard lactation milk yield from monthly milk yields for Marathwadi buffalo breed. Tyasi *et al.* 2018 used multiple linear regression analysis for prediction of carcass weight from body measurement traits of Chinese indigenous Dagu male chickens. Haile *et al.* 2008 also used MLR in their study. Few workers used ANN model for prediction of milk yield in various breeds of cattle (Grzesiak *et al.* 2003, Grzesiak *et al.* 2006, Sharma *et al.* 2006, Gandhi *et al.* 2009) and reported that the performance of artificial neural network model was somewhat superior to that of conventional regression model. ANN has found wide application in diverse areas of food science research, while its application animal science in general and animal breeding in particular is scanty (Grzesiak *et al.* 2003, Sharma *et al.* 2006, Hosseinia *et al.* 2007, Gandhi *et al.* 2010).

*Development of ANN model:* The ANN was trained on the training data set having FSP, TD2, TD4 and TD6 (variables) as that was incorporated in optimum equation for regression analysis. A total of 30 networks were evaluated for each multilayer perceptron. Several combinations of hidden layers (1–2 layers) with varying number of neurons (1–20 neurons) were experimented to train the network and the best results was obtained with the combination of 1 hidden layer and 2 neuron in that hidden layer. Various criteria of judging the effectiveness of MLR and ANN analyses are given in Table 5 and Table 6.

A decreasing trend by the root mean square error (RMSE) while decreasing the percentage of test data set was verified (Table 6). In SET-C (training data-test data: 85–15%), the artificial neural network explained 76.8% coefficient of determination with one hidden layer and 20 nodes. Further, mostly in all the SETs, the performance of artificial neural network was found to be better than multiple linear regression except SET-B (Training-test data: 75–25%) in which ANN shows slightly higher RMSE. However, in SET-A, the ANN shows slightly better results than that of MLR. The RMSE values for SET-A, SET-B and SET-C have been presented in Table 6. The R<sup>2</sup> value of prediction of FL305DMY was 76.8%. As the ANN predictions gave similar R<sup>2</sup> values with lower RMSE in comparison to MLR, it can be interpreted that ANN is comparatively more accurate to predict FL305DMY in Murrah buffalo. The regressions of predicted FL305DMY on actual yield



Table 5. Estimation of criterion values for judging the optimum model for FL305DMY in Murrah buffalo

Model	Variables	R <sup>2</sup>	RMSE	MSE	AIC	BIC	MCp	MAE	MAPE	U (10 <sup>8</sup> )	Parameters
1	FSP	0.041	484.8	234996.	35840.2	8.1	1736.5	380.8	25.3	14034.2	2
2	TD2	0.504	348.6	121484.9	5202.9	7.8	631.1	265.2	17.3	9924.0	2
3	TD4	0.568	325.2	105776.7	5069.2	7.8	478.1	246.9	15.9	9239.3	2
4	TD6	0.645	294.8	86918.8	4879.5	7.7	294.5	221.3	13.8	8352.6	2
5	FSP, TD2	0.528	340.5	115941.2	7736.7	10.6	577.1	257.8	16.5	9799.0	3
6	FSP, TD4	0.592	316.4	100096.9	7523.8	10.5	423.0	236.6	15.1	8971.7	3
7	FSP, TD6	0.669	285.1	81288.5	7222.3	10.4	240.2	211.5	13.0	8063.9	3
8	TD2, TD4	0.640	297.0	88229.6	7340.9	10.4	307.7	224.2	14.3	8409.2	3
9	TD2, TD6	0.720	262.3	68822.2	6981.0	10.3	119.0	196.5	12.3	7406.4	3
10	TD4, TD6	0.722	261.3	68283.9	6969.7	10.3	113.8	196.9	12.2	7376.8	3
11	FSP, TD2, TD4	0.662	288.5	83207.8	9674.8	13.2	259.4	213.1	13.5	8153.1	4
12	FSP, TD2, TD6	0.740	252.9	63940.9	9166.0	13.0	72.4	185.5	11.5	7127.3	4
13	TD2, TD4, TD6	0.749	248.2	61607.3	9094.2	13.0	49.8	185.3	11.5	6993.9	4
14	FSP, TD2, TD4, TD6	0.769	238.5	56887.4	11175.3	15.7	5.0	172.1	10.6	6707.8	5

Table 6. RMSE value for different ANN Model

No. of Hidden layer	Node	RMSE		
		Training: Testing:: 65:35 (SET A)	Training: Testing:: 75:25 (SET B)	Training: Testing:: 85:15 (SET C)
2	10:5	225.5	314.2	220.3
2	5:3	229.7	267.6	218.8
2	3:2	228.7	269.8	225.7
2	2:1	238.4	268.5	219.9
1	2	225.2	272.9	229.5
1	3	221.8	265.9	223.3
1	5	229.9	265.2	221.3
1	7	221.6	254.0	219.4
1	10	223.1	264.6	222.4
1	20	289.3	484.3	216.3

predicted by ANN and MLR were plotted and are presented in Figs 1 and 2.

A neural network model based on back-propagation learning has been found helpful for prediction of dairy yield (Salehi *et al.* 1988, Salehi *et al.* 1998, Grzesiak *et al.* 2003). Artificial Neural Networks have been used effectively in other investigations for dairy yield prediction and cow culling classification (Lacroix *et al.* 1997). Prediction of cow performance with connectionist model has shown better results than conventional methods (Lacroix *et al.* 1995). Milk production estimates have been effectively obtained in an investigation by using feed forward artificial neural networks (Sanzogni and Kerr, 2001). ANNs have been used to predict milk yield in dairy sheep (Salehi *et al.* 1988). ANNs have been applied for detecting influential variables in the prediction of incidence of clinical mastitis in dairy animals (Yang *et al.* 1999, Heald *et al.* 2000, Nielen *et al.* 1995a, Nielen *et al.* 1995b). A three-layer back-propagation connectionist model has been exploited for pattern recognition to develop Monterey jack cheese (Waisarayutt and Norback 2001), which permits study of

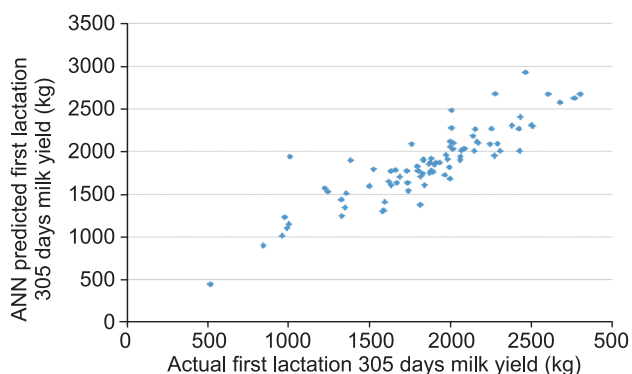


Fig. 1. Actual versus the best ANN model predicted first lactation 305-days milk yield (kg).

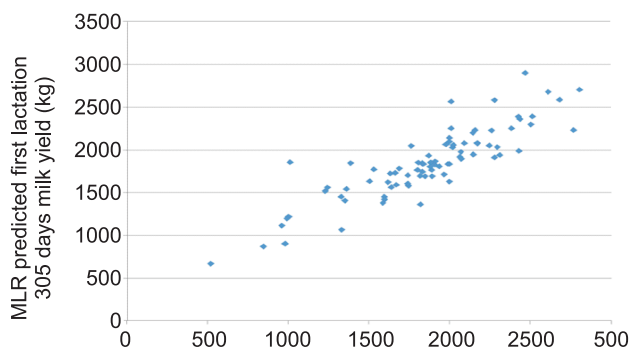


Fig. 2. Actual versus the best MLR model predicted first lactation 305-days milk yield.

real-time control process of cheese production. Also, ANN has been used in modeling of pH and acidity for cheese production (Paquet *et al.* 2000). ANNs have been effectively used to predict temperature, moisture and fat in slab shaped foods with edible coatings during deep-fat frying (Mittal and Zhang 2000). ANN has found broad applications in various areas of animal management, milk production of

dairy animals, viz. prediction of second parity milk yield and fat percentage of dairy cows based on first parity information using neural network system (Edriss *et al.* 2008), forecasts of 305-days milk yield using partial lactation records (Grzesiak *et al.* 2003), practical evaluations of feed forward connectionist and conventional multiple linear regression models for forecast of first lactation 305-days milk yield in Karan Fries dairy cows (Sharma *et al.* 2006), evolving prediction models for lifetime milk production by means of ANN technique in Sahiwal cattle (Gandhi *et al.* 2010).

*Comparison between multiple linear regression and artificial neural network:* In the present investigation the accuracy of prediction obtained from ANN was almost similar to MLR for prediction of FL305DMY using monthly test day milk records in Murrah buffalo. The best ANN algorithm achieved 76.8% accuracy of prediction for optimum model, whereas the MLR explained 76.9% of accuracy of prediction of FL305DMY in Murrah buffalo.

Correlation coefficient between the predicted values of both methods was found to be 0.98. High correlation value is confirming similar result by both methods. MLR method is simple as compare to ANN, hence MLR method should be preferred. The prediction accuracy from all the models increased with the addition of test day milk yields as input variables (Table 5).

In the present study, FL305DMY predictions made by the best ANN model and the MLR model developed here are graphically depicted in Fig.1 and Fig. 2, respectively. These graphs revealed that the accuracy of prediction obtained from ANN was almost similar to MLR for prediction of FL305DMY using monthly test day milk records in Murrah buffalo. Similar finding has been reported by Sharma *et al.* (2006), Njubi *et al.* (2010) and Dongre *et al.* (2012) in Karan Fries, Holstein Friesian and Sahiwal cattle, respectively. It is concluded that the ANN approach has definite application potential for prediction of the first lactation 305-day milk yield in Murrah buffalo but at the same time MLR method is comparatively simple and easy to understand. As in this study both methods are found to be similar in their performance to predict FL305DMY, hence MLR method should be preferred. These models can be improved with incorporation of other production and reproduction traits such as growth data, age at maturity and other traits for more accurate prediction of milk yield. Besides those functions which follow non-linear pattern of production, neural network can also be used for studying other aspects of animal breeding and management such as prediction of lameness, mastitis, prediction of body weight, etc. It can also be used for dam/sire evaluation using single/multiple traits and evaluation of other economic traits of importance in dairy enterprise.

#### REFERENCES

- Adamczyk K, Molenda K, Szarek J and Skrzynski G. 2005. Prediction of bulls slaughter from growth data using artificial

neural network. *Journal of Central European Agriculture* **6**: 133–42.

Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions Automatic Control* **19**: 716–23.

Brethour J R. 1994. Estimating marbling score in live cattle from ultrasound images using pattern recognition and neural network procedures. *Journal of Animal Science* **72**: 1425–32.

Dongre V B, Gandhi R S, Singh A and Ruhil A P. 2012. Comparative efficiency of artificial neural networks and multiple linear regression analysis for prediction of first lactation 305-day milk yield in Sahiwal cattle. *Livestock Science* **147**: 192–97.

Dongre V B, Kokate L S, Khandait V N, Karpe A G, Salunke V M and Patil P V. 2018. Development of standard lactation milk yield prediction models using monthly milk yield records in Marathwadi buffalo. *Indian Journal of Animal Sciences* **88**: 242–44.

Draper N R and Smith H. 1987. *Applied regression analysis*. John Wiley and Sons Inc., New York.

Edriss M A, Hosseinnia P, Edrisi M, Rahmani H R and Nilforooshan M A. 2008. Prediction of second parity milk performance of dairy cows from first parity information using artificial neural network and multiple linear regression methods. *Asian Journal of Animal and Veterinary Advances* **3**: 222–29.

Fausett L. 1994. *Fundamentals of Neural Network*. Prentice Hall, USA.

Gandhi R S and Gurmani M. 1988. Association amongst different productive and reproductive traits in Sahiwal cattle. *Asian Journal of Dairy Research* **7**: 171–74.

Gandhi R S, Raja T V, Ruhil A P and Kumar A. 2009. *Evolving prediction equations for lifetime milk production using artificial neural network methodology in Sahiwal cattle*. Proceedings of XXXVII Dairy Industry Conference. 7–9 February, Goa. pp 103.

Gandhi R S, Raja T V, Ruhil A P and Kumar A. 2010. Artificial neural network versus multiple regression analysis for prediction of lifetime milk production in Sahiwal cattle. *Journal of Applied Animal Research* **38**: 233–37.

Grzesiak W, Blaszczyk P and Lacroix R. 2006. Methods of predicting milk yield in dairy cows-predictive capabilities of wood's lactation curve and Artificial Neural Networks (ANN). *Computers and Electronics in Agriculture* **54**: 69–83.

Grzesiak W, Lacroix R, Wojcik J and Blaszczyk P. 2003. A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records. *Canadian Journal of Animal Science* **83**: 307–10.

Haile A, Joshi B K, Ayaleq W, Tegegne A, Singh A and Chakravarty A K. 2008. Prediction of first lactation milk yield of Boran cattle and their crosses with Holstein-Friesian in Central Ethiopia using multiple regression and principal components analysis. *Indian Journal of Animal Sciences* **78**: 66–69.

Heald C W, Kim T, Sicho W M, Cooper J B and Wolfgang D R. 2000. A computerized mastitis decision and using farm-based records: an artificial neural network approach. *Journal of Dairy Science* **83**: 711–20.

Hosseinnia P, Edrisi M, Edriss M A and Nilforooshan M A. 2007. Prediction of second parity milk yield and fat percentage of dairy cows based on first parity information using neural network system. *Journal of Applied Science* **7**: 3274–79.

Kannan D S and Gandhi R S. 2006. Prediction of lifetime production in Sahiwal cattle. *Indian Journal of Animal*

- Sciences* 9: 768–69.
- Korthals R L, Hahn G L and Nienaber J A. 1994. Evaluation of neural networks as a tool for management of swine environments. *Transactions of the American Society of Agricultural Engineers* 37: 1295–99.
- Kumar V, Chakravarty A K, Patil C S, Valsalan J, Sharma R K and Mahajan A, 2014. Genetic studies on first lactation monthly test day and 305 days milk yields in Murrah buffaloes. *Indian Veterinary Journal* 91: 31–33.
- Lacroix R, Salehi F, Yang X Z and Wade K M. 1997. Effects of data pre-processing on the performance of artificial neural network for dairy yield prediction and cow culling classification. *Transactions of the American Society of Agricultural Engineers* 40: 839–46.
- Lacroix R, Wade K M, Kok R and Hayes J F. 1995. Prediction of cow performance with a connectionist model. *Transactions of the American Society of Agricultural Engineers* 38: 1573–79.
- Lopez-Benavides M G, Samarasinghe S and Hickford J G H. 2003. *The use of artificial neural networks to diagnose mastitis in dairy cattle*. Proceedings of the International Joint Conference on Neural Networks, July 20–24, 2003, IEEE Computer Society, USA. pp: 582–86.
- Mallows C L. 1973. Some Comments on Cp. *Technometrics* 15: 661–75.
- Mittal G S and Zhang J. 2000. Use of artificial neural network to predict temperature, moisture, and fat in slab-shaped foods with edible coatings during deep-fat frying. *Journal of Food Science* 65: 978–83.
- Molenda K, Gil Z and Zychlinska J. 2001. Conception of computer system support in detection of pregnancy, oestrus and mastitis based on milk temperature. Proceedings of the 52nd Annual Meeting of EAAP, Budapest, Hungary, August 26–29, 2001, Wageningen Pers, pp: 132.
- Mostert B E, Theron H E, Kanfer F H J and vanMarleK Oster E. 2006. Comparison of breeding values and genetic trends for production traits estimated by a Lactation Model and a Fixed Regression Test-day Model. *South African Journal Animal Science* 36: 71–78.
- Nielen M, Schukken Y H, Brand A, Deluyker H A and Maatje K. 1995a. Detection of sub-clinical mastitis from on-line milking parlor data. *Journal Dairy Science* 78: 1039–49.
- Nielen M, Schukken Y H, Brand A, Haring S and Ferwerda-Van Zonneveld R T. 1995b. Comparison of analysis techniques for on-line detection of clinical mastitis. *Journal of Dairy Science* 78: 1050–61.
- Njubi D M, Wakhungu J W and Badamana M S. 2010. Use of test-day records to predict first lactation 305-day milk yield using artificial neural network in Kenyan Holstein Friesian dairy cows. *Tropical Animal Health and Production* 42: 639–44.
- Paquet J, Lacroix C and Thibault J. 2000. Modeling of pH and acidity for individual cheese production. *Journal of Dairy Science* 83: 2393–2409.
- R Core Team 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Salehi F, Lacroix R and Wade K M. 1988. Improving dairy yield predictions through combined record classifiers and specialized artificial neural networks. *Computers and Electronics in Agriculture* 20: 199–213.
- Salehi F, Lacroix R and Wade K M. 1998. Effects of learning parameters and data presentation on the performance of back-propagation networks for milk yield prediction. *Transactions of the American Society of Agricultural Engineers* 41: 253–59.
- Sanzogni L and Kerr D. 2001. Milk production estimates using feed forward artificial neural networks. *Computers and Electronics in Agriculture* 32: 21–30.
- Schwarz G E. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461–64.
- Sharma A K, Sharma R K and Kasana H S. 2006. Empirical comparisons of feed forward connectionist and conventional regression models for prediction of first lactation 305-day milk yield in Karan Fries dairy cows. *Neural Computing and Applications* 15: 359–65.
- Tyasi T L, Qin N, Niu X, Sun X, Chen X, Zhu H, Zhang F and Xu R. 2018. Prediction of carcass weight from body measurement traits of Chinese indigenous Dagu male chickens using path coefficient analysis. *Indian Journal of Animal Sciences* 88: 744–48
- Waisarayutt C and Norback J P. 2001. *Using a neural network for pattern recognition of Monterey jack cheese*. Proceedings of the Dairy Foods Session, IFT Annual Meeting, New Orleans, Louisiana.
- Yang X Z, Lacroix R and Wade K M. 1999. Neural detection of mastitis from dairy herd improvement records. *Transactions of the American Society of Agricultural Engineers* 42: 1063–71.