Indian J Dairy Sci 77(1): 91-96
https://doi.org/10.33785/IJDS.2024.v77i01.013

SHORT COMMUNICATION

Prediction of first lactation milk yield on the basis of test day yield using artificial neural network versus multiple linear regression in Gir cows

K B Savalia[1], A R Ahlawat[2], T K Patbandha[3], Ankita D Verma[4], V V Gamit[3], P G Dodiya[1] And G M Chaudhary[5]

Received: 18 January 2023 / Accepted: 15 August2023 / Published online: 23 February 2024

Abstract: The test- day model is a method of choice for the study of milk yield traits and this method is very important in countries like India where herd size is generally smaller and lacking well-established milk recording system. The present study was aimed to predict first lactation milk yield on the basis of test day yield in Gir cows and comparison was made between the relative efficiency of Artificial Neural Network (ANN) and Multiple Linear Regression (MLR) models. First lactation records of 513 Gir cows sired by 75 bulls spread over a period of 34 years (1981 to 2014), maintained at Cattle Breeding Farm, Junagadh were used for the study. The data of monthly test-day milk yield (MTMY) was divided into seven sets. ANN was used with back propagation Bayesian regularization (BR) algorithm and MLR was used with backward elimination method. The accuracy of prediction of first lactation milk yield in MLR was lower than the accuracy of ANN in all the test data sets. The Root Mean Square Errors (RMSE) of prediction were lower in ANN as compared to MLR. The optimum equation had total four variables (test days) viz. TD2 to TD5 for prediction of First Lactation 305-Days Milk Yield (FL305DMY). This equation gave an accuracy of prediction of 76.02% by MLR and 87.69% by ANN model till 125th days of lactation i.e. 5th monthly test day.

Keywords: First lactation 305 day milk yield, MTDMY, ANN, MLR, Gir cow

As per the 20th livestock census India possess over 300 million bovines which includes 192.49 million cattle. A total 50 cattle breeds have been identified and registered by National Bureau of Animal Genetic Resources, Karnal. Gujarat has rich and bio diverse cattle genetic resources viz. Gir, Kankrej, Dangi and Dagri breeds. Gir cattle are well known milch cattle breed across the whole world. It is known for its integral heat tolerance and disease resistance capacity. The native tract of the breed is Gir hills and forests of Kathiawar including Junagadh, Bhavnagar, Rajkot and Amreli districts of Gujarat (Patbandha et al. 2020).

First lactation 305-days milk yield is considered as an important trait for selection of cows. Test-day milk yield (TDMY) is the measurement of the amount of milk produced by a cow for the period of 24 hours (Schaeffer and Jamrozik, 1996). The test- day model is method of choice for the study of milk yield traits in order to maximize the use of all available information. This method is very important in countries like India where herd size is generally smaller and lacking well-established milk recording. Test day yield model is a substitute of 305-day lactation model because early selection on the basis of test-days could reduce generation interval and using test day yield model, it is possible to economize the genetic evaluation with a better accuracy (Bilal and Khan, 2009).

Various regression models, which are used for future yield predictions are applied in milk yield analyses. However, artificial neural network (ANN) takes an entirely different approach. In practice, ANNs are primarily used in engineering, economic predictions, or in medical diagnoses. There has been relatively little research into the application of ANNs in the field of animal breeding. An artificial neural network, also called a neural network (NN) is a computational model based on biological neural network system. ANNs are used for the evaluation of different parameters (Chaturvedi et al. 2013). Artificial neural networks are based on

[1]Polytechnic in Animal Husbandry, Kamdhenu University, Junagadh, Gujarat, India. E-mail: drsavaliya87@kamdhenuuni.edu.in
[2]Dept. of Animal Genetics and Breeding, College of Veterinary Science & A.H., Kamdhenu University, Junagadh, Gujarat, India. E-mail: arahlawat@kamdhenuuni.edu.in
[3]Department of Livestock Production Management, College of Veterinary Science & A.H., Kamdhenu University, Junagadh, Gujarat, India. E-mail: patbandhavet@kamdhenuuni.edu.in
[4]Dept. of Animal Genetics and Breeding, College of Veterinary Science & A.H., Kamdhenu University, Junagadh, Gujarat, India. E-mail: ankita.agb@kamdhenuuni.edu.in
[5]Dept. of Veterinary Extension , College of Veterinary Science & A.H., Kamdhenu University, Junagadh, Gujarat, India. E-mail: girishchaudhary@kamdhenuuni.edu.in

K B Savalia (✉)
Email: drsavaliya87@kamdhenuuni.edu.in

the neural structure of the human brain, which processes information by means of interaction among many neurons. The basic components of an ANN are neurons, weights and learning rules (Stich et al. 2000). The main advantage of neural networks lies in the ability to represent both linear and non-linear relationships directly from the data being modeled. ANN approach needs specified algorithm to be transformed by a computer program (Grzesiaket et al. 2003). After applying successful training algorithms, the neural network will be capable to perform classification, estimation, prediction or simulation on new data from the same or similar sources. Therefore, the present study was carried out to predict the first lactation 305-days milk yield by using ANN & Multiple Linear Regression (MLR) model based on monthly test day milk yield records and to find the best early test-day milk yields combination in order to make an early selection of the animals for breeding program. The study also compared the effectiveness of MLR and ANN for prediction of first lactation 305-days milk yield in Gir cows.

First calving decreases the cost of raising the animals to productive life, increases the annual genetic gain and raises the average productive life of the animal. The estimate of additive genetic variability for traits of economic importance gives an idea about the scope of genetic improvement of the trait through selective breeding. The selection and evaluation of breeds to be used as the parental stock is an important step for the success of any animal breeding program.

Data pertaining to of 513 Gir cattle from pedigree cum lactation registers maintained at Cattle Breeding Farm, Junagadh Agricultural University, Junagadh for a period of 34 years 1981 to 2014 were used for the present study. Only those animals whose lactation was normal and has completed at least 100 lactation days were selected in the study. The outliers beyond three-standard deviation on both the tail ends of normal distribution were also excluded from the data. The first test day was considered 5th day (soon after colostrum), the next test day was calculated by adding 30 days to the preceding test day up to 305 days of lactation.

The prediction of First Lactation 305-Day Milk Yield (FL305DMY) was performed utilizing the MLR & ANN.

The MLR was used to develop prediction equations by estimating the regression coefficients for the test-day milk yield records in different combination. Stepwise backward multiple linear regression analysis was used to estimate 305- day milk yield (Singh et al. 2015).

$$\hat{Y}_i = a + \Sigma b_i X_i$$

Where,

$\hat{Y}_i$ = Estimated first lactation 300 day or less milk yield of the ith animal

X_i = Test day record of ith animal

a = Intercept

b_i = Regression coefficient of first lactation 305 day or less milk yield on text day records

ANN is a multilayer feed forward neural network with back propagation of error learning mechanism was developed using Neural Network Toolbox (NNT) of MATLAB 7.0 to predict the first lactation 305-day or less milk yield (FL305DMY). The network was trained and simulated using back propagation algorithms viz. Bayesian regularization (BR) (Singh et al. 2020) upto 4000 epochs or till the algorithms truly converged. Network parameters such as learning rate, momentum, and error goal were used as the default setting of the algorithms.

The entire data was divided as training (60%), testing (20%) and validation (20%). The network architecture was composed of monthly test day milk yield as the input layer, two hidden layers with five neurons each while the predicted first lactation milk yield was the output layer. To prevent the network from falling into local minima, the momentum coefficient was used of $\alpha = 0.5$. A general schematic diagram of multilayer feed forward network with input layer, hidden layers and output layer is shown in Figure 1.

The statistical analysis was carried out with help of Microsoft Office Excel-2016 and software MATLAB 7 was used for ANN. Effectiveness was compared in both ANN and MLR methods using percent coefficient of determination (R^2 - Value) and root mean square errors (RMSE).

The results showed that the least square mean of first lactation test day milk yield was consistently increasing up to peak yield 6.19kg (MTDMY3) and after that gradually decreasing in phase which showed the typical pattern of lactation curve in Gir cow (Table-1). Singh (1983) reported 8.38 kg peak yield in Gir cows. The overall mean of FL305DMY or less milk yield in the present study was 1448.19 ± 26.61kg in Gir cow. Savaliya et al. (2016) reported 1554.3 kg and Gadariya et al. (2017) reported 1427.49 kg FL305DMY in Gir cow.

The Multiple linear regression (MLR) analysis revealed that when all monthly test days (TD1 to TD10) were included in equation to predict the first lactation 305 Days milk yield (FL305DMY) the accuracy (R^2 value) was 93.20%. In consonance to the present study, Dongre et al. (2012) observed R^2 value of 92.6% in Sahiwal cows using MLR model. However, when nine variables were included in the equation (TD1 to TD9), 92.58% accuracy was observed. Further, when backward elimination method was

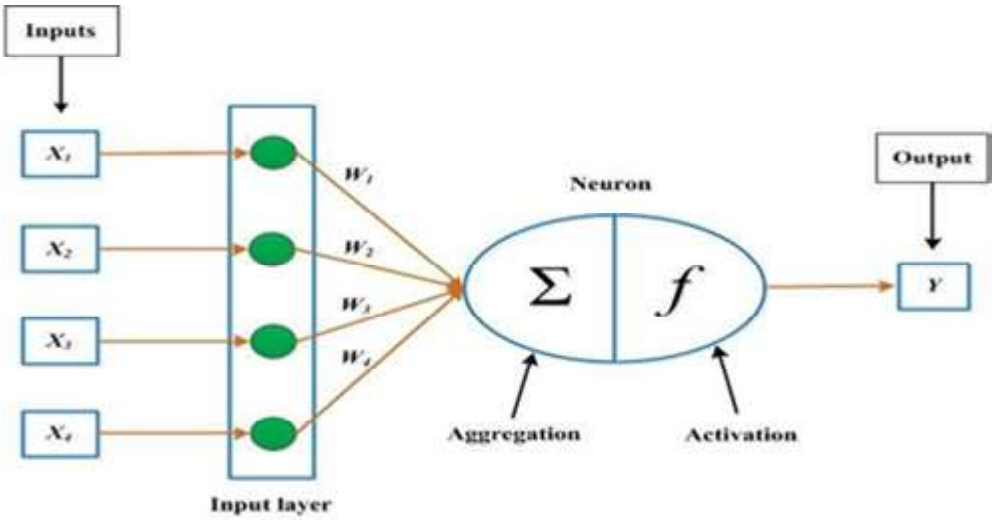

Fig. 1 Schematic diagram of ANN model

Table 1: First Lactation Monthly Test-Day Milk Yield (FLMTMY) of Gir Cows

Traits	Mean (kg)	Standard error	Standard deviation
MTDMY1	5.33	0.09	2.23
MTDMY2	6.12	0.09	2.25
MTDMY3	6.19	0.10	2.37
MTDMY4	5.84	0.10	2.34
MTDMY5	5.47	0.10	2.35
MTDMY6	5.17	0.09	2.25
MTDMY7	4.85	0.10	2.34
MTDMY8	4.68	0.10	2.26
MTDMY9	4.62	0.10	2.12
MTDMY10	4.37	0.10	2.04
FL305DMY	1448.19	26.61	602.71

Table 2: Test days involved in each input sets and R^2-value from MLR and ANN

Input sets	Test days included	R^2 (%)	
		ANN	MLR
Set-1	MTDMY1 to MTDMY10	97.00	93.20
Set-2	MTDMY1 to MTDMY9	96.59	92.58
Set-3	MTDMY1 to MTDMY8	95.32	90.83
Set-4	MTDMY1 to MTDMY7	94.36	88.03
Set-5	MTDMY1 to MTDMY6	91.81	82.25
Set-6	MTDMY1 to MTDMY5	88.03	76.42
Set-7	MTDMY2 to MTDMY5	87.69	76.02

applied to predict the first lactation 305 Day milk yield as early as possible and the degree of accuracy was observed decrease (Table 2). Subsequently, formulation of different prediction equations was done by step-wise backward elimination and total 7 sets were analysed. Table-2 revealed that set-7 included minimum no of TDMY (TD2 to TD5) and yielded >75% accuracy. In set-7 the 1st TDMY was not included because it is difficult to record under field condition and mostly missed. Hence, in this study 4 test day milk yields (TD2, TD3, TD4 and TD5) were found to be suitable for prediction of FL305DMY as early as 125th days and considered as a best formula (Table 3). Olori et al. (1999) stressed that $R^2 \geq 0.70$ indicated a very good fit of a model, while if $R^2 < 0.40$, such model should not be used for prediction. In this study (Table 2), the R^2 - values for all input sets were ≥0.76 in MLR model and ≥0.87 in ANN model. This indicates all models were very good fitted for Gir cows.

The optimum equation to predict FL305DMY as early as possible suggests that the accuracy obtained by ANN was 87.69% and was higher than the accuracy value obtained for MLR. Further, RMSE estimate was found lower in ANN model as compared to MLR model (Table 3). Similar findings were reported by Singh et al. (2022) in Murrah buffalo. Comparatively, higher estimate was observed in this study than the previous studies conducted

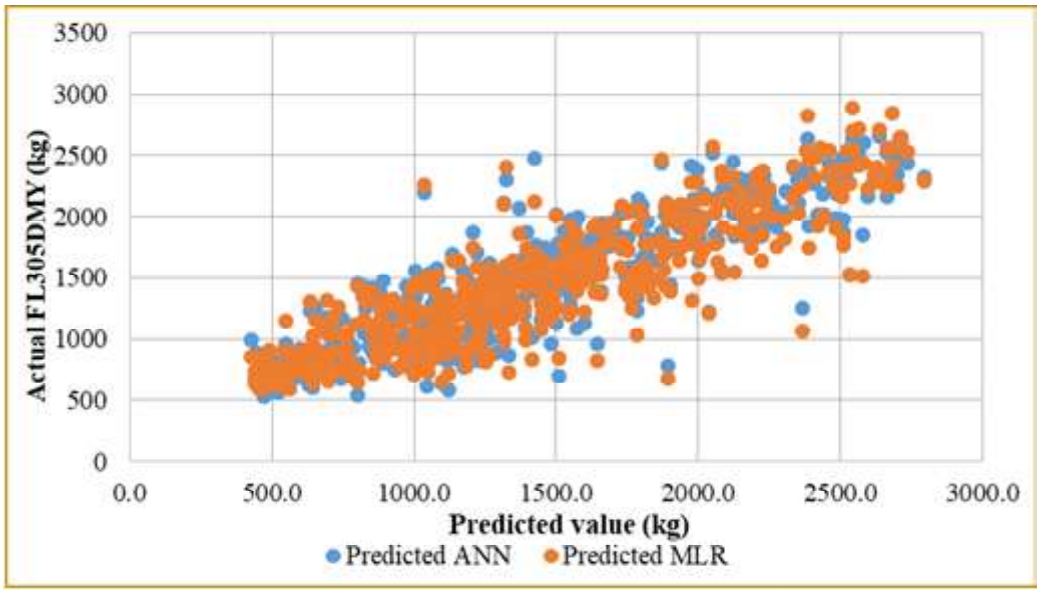

Fig. 2 Prediction of FL305DMY using ANN and MLR model

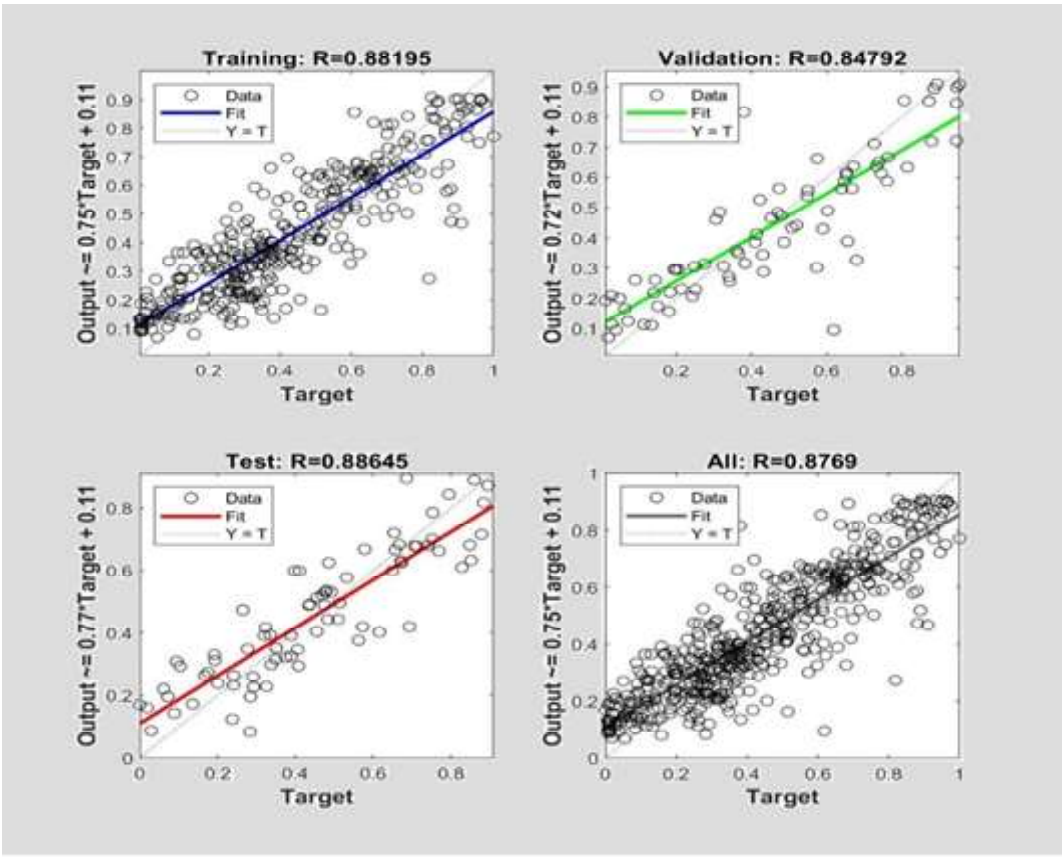

Fig. 3 ANN model for optimum equation

in other breeds of dairy cattle. Lower R^2 - value (72.34%) in MLR model was reported by Debbarma (2010) in Sahiwal cattle. However, Kokate (2009) predicted 61% (MTDY-6) accuracy in Karan-Fries cattle. Saini et al. (2005) reported an accuracy of 78.42% (1st, 2nd and 7th MTDMY) in Rathi cattle by MLR and Ramani (2016) reported 77.71% accuracy for prediction of FL305DMY from monthly test day milk yield using MLR analysis in Gir cattle. While, Dongre et al. (2012) predicted more than 80% accuracy in both methods ANN and MLR for prediction of lactation milk yield by using fortnightly test day yields in Sahiwal cattle, the value which was slightly higher than present estimate value. The higher accuracy obtained must be due to more data allotted for the training.

Table 3: The optimum equation along with their R^2-values developed using MLR and ANN

Input Set-7	Optimum Equation by MLR	R^2-value (%)		RMSE(kg)	
		MLR	ANN	MLR	ANN
MTMY2 to MTMY5	$\hat{Y}$ = 139.19 + (33.36)TD2 + (44.12)TD3 + (64.47)TD4 + (116.61)TD5	76.02	87.69	13.02	12.41

The difference between best ANN model and MLR model for prediction of FL305DMY are graphically presented in Figure 2. The R^2 value of optimum equation in ANN model is represented in different category like training, validation, test and based on all test days record (Figure 3). In the present investigation, difference was found between MLR and ANN model to predict FL305DMY in Gir cows. The best ANN network algorithm achieved 87.69% accuracy, whereas the MLR model achieved 76.02% of accuracy for prediction of FL305DMY in Gir cows. Similarly, Dongre et al. (2012) reported 86.08% prediction accuracy of FL305DMY in Sahiwal cattle using best ANN model. However, the accuracy was comparatively higher than the present study in best MLR model (85.16% vs. 76.02%). The variation of results might be attributed to the data sets used by different studies; in this study monthly test day milk yield records used whereas by Dongre et al. (2012) fortnightly test day milk yield records were used. Further, in crossbred cows like Karan Fries and exotic Holstein Friesian cows similar results were observed (Sharma et al. 2006; Njubi et al. 2010). On the other hand, Mundhe et al. (2015) reported higher accuracy for prediction of FL305DMY in Sahiwal cattle using MLR (88.80%) and ANN (89.29%) model. Moreover, the accuracy value shows minor difference between the two models. However, Rana et al. (2012) reported higher accuracy in MLR equation for an early prediction of FL305DMY than the ANN model in Murrah Buffaloes.

Conclusions

It was concluded that the optimum equation for prediction of first lactation milk yield in Gir cattle using Multiple linear regression was Y = 139.19 + (33.36)TD2 + (44.12)TD3 + (64.47)TD4 + (116.61) TD5 in this study. It provided an accuracy of 76.02% as early as 125th days of lactation. Artificial neural network proved a better approach with higher degree of accuracy (R^2-value 87.69%) for prediction of first lactation milk yield in Gir cattle. It was inferred that FL305DMY could be predicted as early as 125th days of lactation even without having 1st TDMY with high degree of accuracy. Hence, evaluation of Gir cattle at an early stage will help to make selection decisions and could lead to genetic improvement of the animals.

Acknowledgment

The authors are highly thankful to Principal & Dean, College of Veterinary Science & Animal Husbandry, K.U., Junagadh and Research Scientist (AGB), Cattle Breeding Farm, J.A.U., Junagadh for providing necessary facilities.

References

Bilal G, Khan MS (2009) Use of test-day milk yield for genetic evaluation in dairy cattle: a review. Pak Vet J 29: 35-41

Chaturvedi S, Gupta AK, Yadav RL, Sharma AK (2013) Life time milk amount prediction in dairy cows using artificial neural networks. Int J Curr Res 5:1-6

Debbarma M 2010 Genetic analysis of test day milk yield in Sahiwal cattle. M.V.Sc. Thesis, NDRI, Deemed University, Karnal, India

Dongre VB, Gandhi RS, Singh A, Ruhil AP (2012) Comparative efficiency of artificial neural networks and multiple linear regression analysis for prediction of first lactation 305-day milk yield in Sahiwal cattle. Livest Sci 147(1-3): 192-197

Gadariya MR, Vataliya PH, Murthy KS, Gajbhiye PU (2017) Breeding and lactation efficiencies, production profile, productive herd life and lifetime productivity of Gir cows in their home tract. Ind J Vet Sci, Biote 13(2): 26-30

Grzesiak WRL, Wojcik J, Blaszczyk P (2003) A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records. Canad J Anim Sci 83: 307-310

Kokate LS (2009) 'Genetic evaluation of Karan Fries sires based on test day milk yield records.' M.V.Sc. Thesis, NDRI, Deemed University, Karnal, Haryana, India.

Mundhe UT, Gandhi RS, Das D N, Dongre VB , Gupta A (2015) Prediction of FL305 DMY from monthly part lactation milk yield records using artificial intelligence in Sahiwal cattle. Ind J Anim Sci 85: 477-479.

Njubi DM, Wakhungu JW, Badamana MS (2010) Use of test-day records to predict first lactation 305-day milk yield using artificial neural network in Kenyan Holstein Friesian dairy cows. Trop Anim Health Prod 42(4): 639–644.

Olori VE, Brotherstone S, Hill WG, McGuirk BJ (1999) Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. Livest Prod Sci 58(1):55-63

Patbandha TK, Sabapara GP, Savaliya BD, Dash SK, Parikh SS, Ali M (2020) Physical Characteristics and Production Performance of Gir Cattle in India. Int J Livest Res 10(8): 1-11

Ramani AL (2016) 'Inheritance of test day milk yield in Gir cattle.' M.V.Sc. Thesis, JAU, Junagadh, Gujarat.

Rana E, Gupta AK, Singh A, Ruhil AP, Malhotra R, Yousuf S, Ete G (2021) Prediction of first lactation 305-day milk yield based on bimonthly test day milk yield records in Murrah buffaloes. Indian J Anim Res 55(4): 486-490.

Saini T, Gahlot GC, Kachwaha RN (2005) Prediction of 300 days lactation yield on the basis of test day milk yield in Rathi cows. Indian J Anim Sci 75(9): 1087-1089

Savaliya BD, Parikh SS, Gamit PM, Gajbhiye PU (2016) Environmental factors affecting economic traits in Gir cattle. Int J Sci Environ Technol 5(4): 2467-2475.

Schaeffer LR, Jamrozik J (1996) Multiple-trait prediction of lactation yields for dairy cows. J Dairy Sci 79: 2044-2055.

Sharma AK, Sharma RK, Kasana HS (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 Comput Appl 15: 359-365

Singh J (1983) 'Persistency of milk production in Gir cattle.' M.Sc. Thesis. Anand Agricultural University, Anand, Gujarat

Singh M, Singh A, Gupta AK, Dash SK, Gupta A, Sahoo SK, Das S and Shivahre PR (2015) Comparative evaluation of different lactation curve models in prediction of monthly test-day milk yields in Murrah buffaloes. Indian J Anim Res 5: 189-193

Singh NP, Dutt T, Sheikh MU, Baqir M, Tiwari R, Kumar A (2022) Prediction of first lactation 305 days milk yield using artificial neural network in Murrah buffalo. Indian J Anim Sci 92(9): 1116–1120

Singh NP, Usman SM, Maurya V, Dutt T, Bhatt N, Kumar A (2020) Comparative analysis of artificial neural network algorithms for prediction of FL305DMY in Murrah Buffalo. Int J Livest Res 10(9):205-209

Stich TJ, Spoerre JK, Velasco T (2000) The application of artificial neural networks to monitoring and control of an induction hardening process. J Ind Technol 16: 1-11