Journal of Domestic Animal Practice & Research 1 (1): 33-39.

Journal homepage: http://www.medvet.co.in

Original research

Genetic evaluation of Surti buffalo to predict the 305 days milk yield using connectionists (Artificial Neural Network) methods.

Priyank Vyas[1], Pallavi Joshi[2], Urmilla Pannu[3], Ved Prakash[4]

1. Graphic Era Deemed to be University, Dehradun 248002, India
[2.] Graphic Era Hill University, Dehradun 248002, India
[3] National Research Centre on Camel, Bikaner 334001, India
4 Department of Animal Genetics & Breeding, Rajasthan University of Veterinary & Animal Sciences Bikaner 334001, India

***Correspondence**: priyankvyasbio@gehu.ac.in (P.V.); pallavijoshi@gehu.ac.in (P.V.)

Article Received: 18 May 2022
Accepted for publication: 19 June 2023

Abstract

ANN model is basically an intelligent data processing system that learns the predictive ability automatically from the data set presented while training the network. A properly trained network is likely to give a reasonable output when presented with new inputs. The literature available on the use of the artificial neural network for predicting milk yield on the basis of test-day records by different workers has been reviewed. In the present study, the ANN was trained on the training data sets having 2, 3, and 4 test days (variables) as that was incorporated in optimum equations from regression analysis. The network was trained with training data set (80%) a number of times to get consistent results. The prediction performance was tested using a new data set i.e. test data (20%). A total of five test day records (as incorporated in optimum equations from regression analysis) were used as input variables for the artificial neural network. In the present investigation, the efficiency of the artificial neural network algorithm was higher than the multiple linear regression models for the prediction of FL305DMY using fortnightly test day milk yield records in Surti buffaloes which confirms the efficiency of machine learning intelligence models over the basic mathematical formulas. It was concluded that the artificial neural network algorithm was better than the multiple linear regression model for the prediction of FL305DMY using fortnightly test day milk yield records in Surti buffaloes.

Key words: Artificial neural network, Surti buffalo, 305 days milk yield, genetic evaluation.

Introduction

The major thrust in buffaloes breeding programs in the country is to identify animals of high genetic merit to serve as parents of the next generation [1]. The selection of individuals is usually based on their first lactation performance evaluated on the basis of the first lactation 305-days or less milk production of their daughters. Though selection needs to be practiced in both sexes, the importance of sire selection far exceeds that of dam selection. Firstly, because the sire has a much greater impact on breed improvement in terms of progeny it can leave behind as compared to the dam, particularly through the use of artificial insemination. Secondly, because only a few sires are needed as replacements as compared to dams, therefore, a much more rigorous selection is possible in males. The performance of Surti buffaloes can be improved by selecting individuals based on early expressed traits that have a strong genetic correlation with milk production. This approach is essential for achieving a higher rate of genetic improvement. Therefore, it is crucial to evaluate both genetic and non-genetic factors of these traits in order to design an effective breeding program. Environmental factors such as the season and period of calving, as well as the parity of the animals, should also be taken into account during this assessment.

The concept of connectionist models (artificial neural networks) had been introduced recently in animal breeding for predicting milk yields [2]. ANN model is basically an intelligent data processing system that learns the predictive ability automatically from the data set presented while training the network. A properly trained network is likely to give a reasonable output when presented with new inputs. The literature available on the use of the artificial neural network for predicting milk yield on the basis of test-day records by different workers has been reviewed. Grzesiak et al. [3] predicted the standard lactation milk yield of dairy cows using the results of the first four days, mean herd production, number of days in milk, month of calving, lactation number, and percent of the HF genes using two artificial neural networks (ANN) models and reported that the predicted milk yields did not differ significantly from the actual mean. Grzesiak et al.[4] used artificial neural network (ANN) models to predict milk yield for both full and standardized lactations in dairy cows using 108,931 daily milk yield (dataset-A) and test-day milk yield records collected from three lactations of dairy cows managed in a production farm. Using the actual data on daily milk yields and the data recorded on official milk recording test-day, a number of neural networks were designed and parameters of Wood's model were estimated.

The quality of each network and regression model was measured using coefficients of determination, relative approximation errors (RAE) and root means square error (RMSE). In order to test the prognostic parameters of the models, a subset of cows was selected randomly from the studied population, which produced a dataset of 28,576 daily yields (dataset B). For those cows, daily and lactation yield forecasts were generated, which were next compared with their actual (observed) yield records and with the yields calculated by SYMLEK program. The results have shown that the quality parameters of the designed neural networks were better than those of the regression model, for both the daily yields and test day data (higher coefficients of determination and lower RMSE). The prognostic parameters estimated for the forecasts of the neural networks were characterized by lower errors of prediction for both the daily yields and test-day data and exhibited higher coefficients of correlation between the predicted and the actual data (or the yields produced by SYMLEK) [4]. The predictions by the neural networks were

more accurate than those by Woos models. Furthermore, the predictions by both analyzed models were closer to reality than the values estimated with the SYMLEK system. Sharma et al. [5] used two connectionist models based on different learning paradigms, viz., back propagation neural networks (BPNN) and radial basis function neural networks (RBFNN) for the prediction of 305-day milk yield.

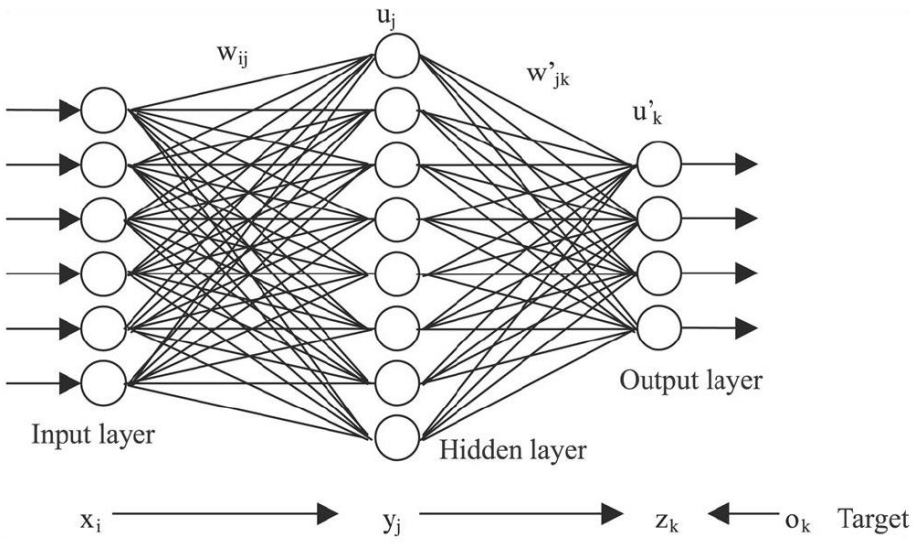
Several training algorithms, viz., a) gradient descent algorithm with adaptive learning rate; b) Fletcher-Reeves conjugate gradient algorithm; c) Polak-Ribiere conjugate gradient; d) Powell-Beale conjugate gradient algorithm; e) Quasi-Newton algorithm with Broyden, Fletcher, Goldfarb, and Shanno (BFGS) update; and f) Levenberg-Marquardt algorithm with Bayesian regularization; along with various network architectural parameters, i.e., data partitioning strategy, initial synaptic weights, number of hidden layers, number of neurons in each hidden layer, activation functions, regularization factor, etc., are experimentally investigated to arrive at the best model for predicting the first lactation 305-day milk yield. The performances of ANN and MLR models are compared to assess the relative prediction ability of two models. ANN model achieved 92.03 % prediction accuracy, whereas the MLR model attains 91.38 % accuracy. Singh [6] compared the artificial neural network (ANN) algorithm with multiple linear regression (MLR) models for the prediction of the first lactation 305- day milk yield in Murrah buffaloes and achieved an accuracy of 84.97% and 82.76% for ANN and MLR, respectively. In the present study we evaluated the ANN model in Surti buffalo to predict the life time performance.

Materials and methods

Source of Data

The data for the present investigation were collected from the history sheets and daily milk record registers of Surti buffaloes maintained at LRS (Livestock Research Station) Vallabhnagar, Navania, during the period 1987 to 2017. The information source for the present study was the Surti buffalo herd maintained under Network Project on Buffaloes, Livestock Research Station, Vallabhnagar, Rajasthan. This Farm is situated 582 meters above the mean sea level at 24º35N latitude and 73º 43E longitudes. The climate of the farm is tropical in nature. The yearly average minimum and maximum temperature range between 2.3ºC to 42.3ºC. The average rainfall and relative humidity are 660 mm and 31.5%, respectively. The performance of buffaloes of foundation stock and their progenies was recorded as per the technical program.

The technical program of the project involved testing a batch of (6-8) bulls in every 18 months. Each bull was mated to 35-40 buffaloes, with an aim to produce about 10-15 daughters. On the basis of the performance of their daughters, two bulls are to be selected from each set. In addition to their extensive use in the field, the selected bulls were mated to 70-80 elite buffaloes (giving more than 1200 kg of milk in the first lactation or more than 1500 kg in any other lactation of 305 days or less). To avoid inbreeding the Surti buffalo bulls of high genetic merit were also introduced in the herd from Central Cattle Breeding Farm, Dhamrod and Reproductive Biology Research Unit, Anand. In this way, the Surti germplasm at Vallabhnagar farm represented progeny of the Surti buffalo bulls of various centers situated in Gujarat state i.e., the home tract of Surti buffalo.



SD ratio is a quotient of two standard deviations of errors and data

$$\text{SD ratio} = \frac{\sum (e_i - e)^2}{\sum (y - y_i)^2}$$

Where,

e_i= the individual error of data set

e= mean error of data set

y = the mean value

y_i = the actual value

Root Mean Square and Standard Deviation (SD) ratio methods are used for evaluating the performance of model.

Root Mean Square (RMS) error

$$\text{RMS} = \frac{\sum (y - y_i)^2}{n}$$

Where,

y = value estimated by model

y_i = Actual value

n = Number of elements in the training and testing set

Fig 1 The artificial neural network model

Prediction by the connectionist model (Artificial neural network)

The data on fortnightly test-day milk yields were used to predict first lactation 305-day or less milk yield (FL305DMY) by MATLAB 2.9. ANN model is basically an intelligent data processing system that learns the predictive ability automatically from the data set presented while training the network. A neural network consists of the input layer, hidden layer(s), and output layer. Each layer has a specific role in the execution of the neural network. In the back propagation technique, the input vector and the corresponding target vectors are used to train a network until it can approximate a prediction function. Dataset were separated into training and testing dataset. The training dataset are used for training the connectionist network while the test dataset are used to validate the connectionist model (Fig. 1).

Results and Discussion

Artificial neural network (ANN)

The ANN was trained on the training data sets having 2, 3 and 4 test-days (variables) as that was incorporated in optimum equations from regression analysis. The network was trained with training data set (80%) for a number of times to get consistent results. The prediction performance was tested using a new data set i .e. test data (20%). A total of five test day records (as incorporated in optimum equations from regression analysis) were used as input variables for artificial neural network. Those five test day records were 110D (TD-8), 125D (TD-9), 170D (TD-12), 230D (TD- 16) and 290D (TD-20) day of lactation (Table 1). Further, a total of three input sub-sets have been prepared with a total of five test days which were used as input variables. The first input set included two test day milk yields records viz. TD-8 and TD-12; second input set included three test days namely TD-8, TD-12 and TD-16 and the third set included four test days namely TD-8, TD-9, TD-12 and TD-20 (Table 2).

The best artificial neural network algorithms achieved 8.72, 6.92 and 6.12% RMSE of prediction for optimum models in SET-1, SET-2 and SET-3, whereas the multiple linear regression model (MLR) explained 10.38, 9.17 and 8.73% RMSE in these sets, respectively for prediction of FL305DMY in Surti buffaloes. In the present investigation, the efficiency of the artificial neural network algorithm was higher than the multiple linear regression model for the prediction of FL305DMY using fortnightly test day milk yield records in Surti buffaloes. Sharma et al. [7,8], Gandhi et al. [9] and Singh [6] also reported higher accuracy of prediction of FL305DMY obtained from artificial neural network algorithm compared to multiple linear regression using weekly and monthly test-day milk records in Sahiwal cattle and Murrah buffaloes, respectively.

Conclusion

It was concluded that the artificial neural network algorithm was better than the multiple linear regression model for the prediction of FL305DMY using fortnightly test day milk yield records in Surti buffaloes.

Table: 1 Description of input & output variables used in ANN models.

Input Variables	Day of lactation	Output Variables
TD-8	110th day of lactation	First lactation 305 day or less milk yield (FL305DMY)
TD-9	125th day of lactation	
TD-12	170th day of lactation	
TD-16	230th day of lactation	
TD-20	290th day of lactation	

Table: 2 Comparison of MLR and ANN for prediction of FL305DMY using Fortnightly test-day milk yield records

Input Sets	Test days included	Training Data-Test Data (%)	RMSE (%)	
			MLR	ANN
SET-1	TD-8 and TD-12	80-20	10.38	8.72
SET-2	TD-8, TD-12 and TD-16	80-20	9.17	6.92
SET-3	TD-8, TD-9, TD-12 and TD-20	80-20	8.73	6.12

Acknowledgement:

The authors want to acknowledge Rajasthan University of Veterinary & Animal Sciences & Livestock Research Centre for providing all the Laboratories & Financial facilities for the successful designing of experiments and concluding the research.

References:

[1] Vaas KK, Gupta VK. Breeding Policy for Cattle and Buffalo in India. National Acad. Agric. Sci.2016 http://naas.org.in/Policy%20Papers/policy%2082.pdf
[2] Gorgulu O. Prediction of 305-day milk yield in Brown Swiss cattle using artificial neural networks. South Afr. J. Anim. Sci. 2012; 42(3):280-287.

[3] Grzesiak W, Lacroix R, Wojcik J, Blaszczyk P. A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records. Can. J. Anim. Sci.2003; 83(2): 307-310.
[4] Grzesiak W, Błaszczyk P, Lacroix R. Methods of predicting milk yield in dairy cows—Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs), Comput. Electr. Agric. 2006; 54(2):69-83,
[5] Sharma A K, Sharma RK, Kasana HS. Prediction of first lactation 305- day milk yield in Karan Fries dairy cattle using ANN modeling. Applied Soft Computing 2007; 7(3):1112-1120.
[6] Singh M, Singh A, Gupta AK, Chakravarty AK, Prakash V, Singh A, Sahoo SK. Relative effectiveness of best linear unbiased prediction-animal model vis-à-vis other sire evaluation methods for genetic evaluation of Murrah sires. Indian J. Dairy Sci. 2014; 67(6):510-514.
[7] Sharma AK, Sharma RK, Kasana HS. 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. 2006; 15(3-4):359-365.
[8] Sharma AK, Jain DK, Chakravarty AK, Malhotra R Ruhil AP. Predicting economic traits in Murrah buffaloes with connectionist models. J. Indian Society Agri. Statis. 2013; 67(1): 1-11.
[9] Gandhi RS, Raja TV, Ruhil AP, Kumar A. Prediction of lifetime milk production using artificial neural network in Sahiwal cattle. Indian J. Anim. Sci. 2009; 79(10): 1038-1040.