



## Neural Network and Regression Based Model for Cows' Milk Yield Prediction in Different Climatic Gradients

Bosedo Oyegbile<sup>1</sup>, Oludayo Michael Akinsola<sup>2\*</sup>, Okeke Rufina Obioma<sup>1</sup>,  
Adekola Omololu Atanda<sup>1</sup>, Balami Samuel Paul<sup>1</sup>, Mary Foluke Oladipo<sup>2</sup>  
and Zulfat Suleiman Abba<sup>3</sup>

<sup>1</sup>Department of Animal Science, Ahmadu Bello University, Nigeria.

<sup>2</sup>Department of Theriogenology and Production, University of Jos, Nigeria.

<sup>3</sup>Department of Animal Science, University of Maiduguri, Borno State, Nigeria.

### Authors' contributions

*This work was carried out in collaboration between all authors. Author BO designed the study. Author OMA performed the statistical analysis. Author AOA wrote the protocol. Author MFO wrote the first draft of the manuscript. Author BSP managed the literature searches. Author ORO managed the analyses of the study. Author ZSA managed the data collection. All authors read and approved the final manuscript.*

### Article Information

DOI: 10.9734/ARRB/2018/41947

#### Editor(s):

(1) Dr. J. David Puett, Professor, Department of Biochemistry and Molecular Biology, University of Georgia, Athens, USA.

(2) George Perry, Dean and Professor of Biology, University of Texas at San Antonio, USA.

#### Reviewers:

(1) Ozkan Gorgulu, Ahi Evran University, Turkey.

(2) Tamer Çağlayan, Selcuk University, Turkey.

Complete Peer review History: <http://www.sciedomain.org/review-history/25923>

Original Research Article

Received 7<sup>th</sup> May 2018  
Accepted 19<sup>th</sup> July 2018  
Published 17<sup>th</sup> August 2018

### ABSTRACT

The present study was designed to develop the prediction equations for 305 days fat corrected milk yield on the basis of part periods milk yield, milk component and conformation traits of multi-genotype cows. Artificial Neural Network model had the best prediction accuracy across varying environments, though Genetic Function Algorithm had the overall best adequacy for fat corrected milk yield predictions (FCM305d=1036.1-98.3RP+22FY+15.92UC-0.07RUH; Adj R<sup>2</sup>=0.997; RMSE=30.07; BIC=1997.28).

**Keywords:** Prediction; artificial neural network; genetic function algorithm; multi-genotype cows.

\*Corresponding author: E-mail: [dayoakinsola@gmail.com](mailto:dayoakinsola@gmail.com);

## 1. INTRODUCTION

The selection of dairy cattle at an early age to predict later yield on the basis of part yields is beneficial to the dairy farmer as it cuts down the over-head cost of rearing the animals. Part yields (daily or monthly milk records) have been shown to have a very high genetic and phenotypic relationship with 305 full records milk yield [1]. The ability to predict the complete lactation period of a cow from its part yields and conformation traits would determine the successes of dairy herd culling programmes [2]. Nigeria spends an estimated amount of \$1.3 billion on the importation of dairy products and government's target is to double milk production over the next three to four years in order to meet up with domestic consumption and export. An average cow in Nigeria is said to produce 1 kg of milk a day, showing there is a long way to go before reaching self-sufficiency [3]. Unfortunately, the domestic output of about 503,000 metric tonnes of milk from an estimated 14 million cattle can hardly satisfy the dairy demands of an ever increasing population of Nigerians [4]. The thinking is that with the on-going diversification agenda and the push for industrialization to mitigate the effects of the economic recession, an initiative in the mode of this research could not have come at a better time than this which will help prioritize a robust institutional framework and necessary policy formulations geared towards genetic improvement of milk yield to bridge the gap for protein sufficiency. These attempts are aided by increased computing power and software capability available today that have facilitated the use of more appropriate models (Multiple linear regression, Genetic function algorithm and neural network) and more sophisticated statistical procedures for prediction. The present study was, therefore, undertaken to develop the prediction equations for 305 days fat corrected milk yields on the basis of part period milk production, milk components and conformation traits for multi-genotype cows under different environmental sensitivity.

## 2. MATERIALS AND METHODS

The data for the present study were taken from the pedigree production records of multi-genotype cows (Holstein, Jersey, Simmental, Brown Swiss, Holstein XBunaji and Jersey XBunaji) maintained at private Dairy Farms (Sebore, Adamawa State; Shonga; Kwara State and Integrated Dairy Limited, Plateau State) in

Nigeria. For the present study, the monthly milk yield data of 12000 lactations were collected from 2000-2016. The complete lactation milk records of 1260 multi-genotype cows were extracted (1<sup>st</sup> lactation to 6<sup>th</sup> lactation) from the milk record database. Animal records with the history of abortion, stillbirth and records not complete were eliminated from the study. Data collected for 305 days milk yield, daily milk yield, milk components and conformation traits were subjected to analysis prediction equations for 305 days milk yield.

### 2.1 Milk Yield Correction for Fat

$$\text{Fat corrected milk (FCM)} = [(0.4 * \text{milk yield (kg)} + [15 * \text{fat yield (kg)}] \quad (1)$$

### 2.2 Milk Component Measurements

Butter fat and protein percentages were measured by infrared spectroscopy, using a Lactoscan analyser in the quality control laboratory of Shonga Dairy Holdings in Kwara State. Three (3) mls of raw milk from the bulk milk tank were injected into the automated lactoscan milk analyzers for determination of butterfat, protein and total solids in Shonga and Sebore farms while in Integrated Dairies Limited, analysis were done using the conventional method. For the determination of total solids content (milk solid with fat), two (2) mls of fresh cows raw milk sample were thoroughly mixed and 5 g was transferred to a pre-weighed and dried flat bottom crucible [6]. The milk samples were dried in a hot air oven (Serial No-96H203, Model-EDSC made in England) at 102 °C for 3 hours. Finally, the dried samples were taken out of the oven and placed in desiccators to cool to room temperature. Then samples were weighed again and total solids was calculated by the following formula [7].

$$\text{Total solids} = \text{Crucible weight} + \text{Oven dry sample weight} - \text{Crucible weight} \times 100 \text{ Sample weight}$$

### 2.3 Butter Fat Content of Milk

The fat content was determined by the Gerber method according to [7]. Ten ml of sulphuric acid (density 1.815 gm/ml at 20°C) was pipetted into a butyrometer. Then eleven ml of milk sample was added into the butyrometer and mixed with the sulphuric acid. This was followed by addition of one ml amyl alcohol into the butyrometer which was then closed with a lock stopper. Then the mixture was shaken and inverted several times until the milk was completely digested by

the acid. Finally, the butyrometer was kept in water bath for 5 minutes at 65°C and centrifuged in a Gerber centrifuge for 5 minutes. The butyrometer was placed in water bath again at 65°C for 5 minutes. At the end, the butyrometer reading was recorded.

## 2.4 Crude Protein Content of Milk

The crude protein content of milk samples was determined by the Kjeldahl method [6]. Five gramme (5 g) of milk sample was warmed in water bath at 38°C and poured into a Kjeldahl tube. A mixture of fifteen gramme (15 g) potassium sulphate, one ml of copper sulphate solution and 25 ml of concentrated sulphuric acid were added to the tube and mixed gently. The digestion was carried out for 120 minutes at 35°C using micro-Kjeldahl digester in the presence of catalyst (1 ml of copper sulphate and 15 g potassium sulphate) where sulphuric acid was used as an oxidizing agent. Then it was allowed to cool at room temperature over a period of 25 minutes. The digested solution was diluted with 250 ml of distilled water. The Kjeldahl tube was placed in the distillation equipment. Then, 75 ml of 40% sodium hydroxide solution was added into the tube. Then ammonia was distilled using 50 ml of 4% boric acid solution with bromocresol green/methyl red as indicators until blue color appears. Finally, the sample was

titrated with 0.1N hydrochloric acid solution until a faint pink color is formed and the burette reading was taken to the nearest 0.01 ml. Blank test was carried out using the above procedure except that water was used instead of the test sample. The percentage of nitrogen in the milk samples was calculated using the formula provided by [6].  $\% N = 1.4007x (v_s - v_b) \times N_{HCl} \times 100$  Weight of sample  $\% CP = \% N \times 6.38$  Where:  $\% N$  = percentage of nitrogen by weight;  $V_s$  = volume of HCl used for titration of sample;  $V_b$  = volume of HCl used for titration of the blank;  $\% CP$  = percent of crude protein Butter fat and protein yields were calculated by multiplying each percentage by the average between morning and evening milk yield.

## 2.5 Body Condition Scores

This was measured using a subjective visual score. Body condition scoring was measured on threshold scale which range from 1 through 5 according to the procedure of [5]

1. Implies severe under conditioning (Emaciated)
2. Implies frame obvious
3. Implies frame and covering well-balanced
4. Implies fat, no processes discernable
5. Implies severe over conditioning (grossly fat)

**Table 1. Definition of conformation traits of dairy cattle**

No	Measurements	Description	Instruments
1	Chest ligament	Measured as the depth of cleft at the base of the rear udder	Flexible tape
2	Chest width	Measured as the inside surface distance between the top of the front legs	Flexible tape
3	Body depth	Measured as the distance between the top of spine and the bottom floor of the abdomen at last rib	Flexible tape
4	Stature	Measured from the top of the spine in between hips to the ground	Measuring stick
5	Rump width	Measured between the inner walls of the two ischial tuberosities (i.e. pin bones)	Flexible tape
6	Heart girth	Measured behind the front legs and shoulder blades	Flexible tape
7	Udder clearance	Measured from the ground to the bottom of the udder	Measuring stick
8	Rear Udder height	Measured as the distance from the bottom of the vulva to the top of the rear udder	Flexible tape
9	Rear Udder width	Measured as the udder width at the point where the rear udder is attached to the body.	Flexible tape
10	Teat length	Measured as the distance from base to tip of the front teat.	Flexible tape

[8]

## 2.6 Statistical Model

The statistical model describing regression for 305day fat corrected milk yield was given as;  $Y = \beta_0 + \beta_1 100dFCM + \beta_2 FY + \beta_3 PY + \beta_4 BW + \beta_5 BD + \beta_6 BCS + \beta_7 UC + \beta_8 TL + \beta_9 RW + \beta_{10} RUW$ ; where  $Y = 305$  day fat corrected milk yield,  $\beta_0 = intercept$ ,  $\beta_{1-10} = coefficient$  of the independent variables (100 dFCM= 100 day fat corrected milk, FY= fat yield, PY= protein yield, BW= body weight, BD= body depth, BCS= body condition score, UC= udder clearance, TL= teat length, RW= rump width and RUW = rear udder width).

## 2.7 Model Selection for Regression Analysis

Selection of traits was done using backward elimination regression method and all regression were checked for multicollinearity and normal distribution of residuals. The Backward elimination method begins by placing all the predictor variables in the model and then removing the predictor variable which contributes the least to the model. The procedure is repeated with the left over predictor variables and for the predictor variable having least contribution is removed. The process is continued until a model in which all of the remaining predictor variables are statistically significant ( $P < 0.05$ ) is obtained.

## 2.8 Artificial Neural Network Computations

The whole data set for 305 day fat corrected milk yield were separated at random into two subsets (training and testing) using a supervised neural network. The training set consisted of 75 percent (experimental) and testing subset comprising of 25 percent (control). The training sets were used to train the neural network models and the testing sets were used to validate the models. The network were tested with 100 hidden layers with 3 to 25 neurons in each hidden layer to optimize the milk yield. Initial weights and bias matrix were randomly initialized between -1 to 1. A nonlinear transformation (or activation) function tangent sigmoid (eq. 1) were used to compute the output from summation of weighted inputs of neurons in each hidden layer. A pure linear transformation function were used as output layer for getting network response.

$$f^{(x)} = \frac{1}{1 + e^{-\alpha x}} \quad (2)$$

Where,  $x$  is weighted sum of inputs and  $\alpha = constant$

The modeled network were trained in supervisory mode with Bayesian regularization back propagation algorithm available in ANN tool box of python programming interface to evaluate the performance of ANN models.

The accuracy of the model was calculated using the coefficient of determination ( $R^2$ ). The  $R^2$  is a predictive strength of the variation in 305 day fat corrected milk yield was captured by the model.

$$R^2 = 1 - \left[ \sum_1^N \left( \frac{Q_{exp} - Q_{cal}}{Q_{exp}^2} \right)^2 \right] \quad (3)$$

$R^2$  value =

$$\frac{\text{Total sum of squares} - \text{Error sum of squares} \times 100}{\text{Total sum of squares}}$$

$Q_{exp}$  = Observed value

$Q_{cal}$  = Predicted value

$N$  = Number of observation

RMSE is defined as the square root of the mean squared error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{obs,i} - X_{model,i})^2}{n}} \quad (4)$$

where  $X_{obs}$  is observed values and  $X_{model}$  is modelled values at time  $i$ .

$n$  = sample size

BIC = Bayesian information criterion

$$BIC = 2 \log(L) + q \log(N) \quad (5)$$

## 2.9 Genetic Function Algorithm

The Genetic function algorithm method (GFA) begins with the formation of a populace of randomly produced parameter sets. The probability of a given parameter from the active set is 0.5 in any of the initial population sets. The parameters set used for the genetic algorithm incorporates the boundaries for mutation (0.1), hybrid (0.9), population (10000), number of model generation (1000),  $R^2$  floor limit (50%), and target capacity ( $R^2/\text{Number of parameters}$ ). The calculation keeps running until the wanted number of generations is achieved. Equations were generated between the experimental biological activity and the descriptors. The best mathematical statement was taken in light of statistical parameters such as adjusted

coefficient of multiple determination ( $\text{Adj } R^2$ ), root mean square error and Bayesian information criterion.

## 2.10 Data Analysis

The ordinary least squares method using the linear model procedures in R CRAN software was used to compute the regression analysis. Genetic function algorithm was modelled in Material studio software. Genetic function algorithm system as a selection tool was incorporated into Material studio 7 program (Acclerys Material Studio, 2014) and utilized in this study.

## 3. RESULTS AND DISCUSSION

Table 2 shows the MLR, NN and GFA models prediction of 305d FCM milk yield from milk production characteristics and conformation traits over a full cycle (305d) of lactation of genotype and environment interactions. All the models were good predictors of FCM305d yield. The NN model had the highest coefficient of determination and therefore ranked first ( $\text{Adj } R^2 = 0.99, 0.99, 0.99$  and  $0.96$ ); the second in rank was MLR ( $\text{Adj } R^2 = 0.96, 0.98, 0.98$  and  $0.94$ ) and third in rank was GFA ( $\text{Adj } R^2 = 0.87, 0.98, 0.95$  and  $0.90$ ) model for Holstein Friesian, FriesianxBunaji, Jersey and JerseyxBunaji in Kwara climatic gradients. The RMSE of the NN prediction (2, 4, 1 and 9%) was lower than that of the MLR (8, 6, 7 and 13%) and GFA (20, 7, 11 and 14%), indicating the presence of larger residual errors in the MLR and GFA forecasts. Neural Network had the best adequacy for model selection than MLR and GFA in prediction of FCM305d yield. The BD, FCM100, FY and PY were traits that featured in the prediction equation of Friesian and FriesianxBunaji cows. The RP and FCM100 were the major predictors of FCM305d yield using different models (MLR, NN and GFA) in Jersey purebred dairy cows while RP, FCM100 and FY were the markers that predicted FCM305d efficiently in JerseyxBunaji cows. It was observed that the accuracy of the prediction method was consistently higher in NN than MLR and GFA studied albeit with different precision across the four genotypes of cows in Kwara State.

### 3.1 Final Candidate Model for 305 Day Milk Prediction in Plateau State

Table 3 shows the model equation, adjusted coefficient of determination, root mean square error and Bayesian information criterion in

predicting FCM305d in Plateau State. Neural network model outclassed MLR and GFA in predicting FCM305d yield with lower RMSE and BIC for Plateau State in Holstein Friesian and FriesianxBunaji. The MLR was more accurate than the GFA with over 9% increment of the FCM305d yield prediction. The MLR and NN were more sensitive in Holstein Friesian dairy cows than FriesianxBunaji in modelling FCM305d yield. The FY and TL were the observed traits that predicted FCM305d in all the models accurately for FriesianxBunaji in Plateau State. It was observed that the accuracy of the prediction method was consistently higher in NN than MLR and GFA studied albeit with different precision across the two genotypes of dairy cows in Plateau State.

### 3.2 Final Candidate Model for 305 Day Milk Prediction in Sebore Farm in Adamawa State

Multiple linear and nonlinear regression equations derived from milk yield characteristics and conformation traits measurement were modelled for standardized 305 full lactation cycle (Table 4). The NN and MLR had the best accuracy of prediction with 99 % of all the FCM305d yield in a full lactation cycle with low noise (0.02 and 0.03) and minimum tears (961.06 and 983.41), respectively in Holstein Friesian breed. The NN was best in Jersey and Brown Swiss cows while in Simmental all the models were very efficient and similar in their efficiency of prediction ( $\text{Adj } R^2=99\%$ ) with varying goodness of fit (0.04 and 983.80, 0.03 and 984.12 and 0.05 and 983.68) in MLR, NN and GFA, respectively.

### 3.3 Final Candidate Model for 305 Day Milk Prediction for Pooled Analysis

Table 5 presents the 305dFCM yield equation pooled for all genotypes for the different models. Prediction equation of 305dFCM within models showed high adjusted  $R^2$  values ranging from 0.995 in the MLR to 0.997 in the NN and GFA, respectively. The BCS, HG, FCM100, FY, PY and RUW were traits that featured in the overall prediction equation for MLR. In the NN with  $R^2$  value of 0.997 consisted of RUW, BCS, HG, FCM100, FY and PY as predictors for 305dFCM while RP, FY, UC and RUH were observed in the GFA with  $R^2$  value of 0.997. It was observed that accuracy of the prediction method was best in GFA with three predictor variables as compared to NN and MLR with six predictors.

**Table 2. Prediction equation for FCM305d from milk and conformation traits using different models in Shonga dairy holdings in Kwara State**

Genotype and herd	Model	Equation	Adj R <sup>2</sup>	RMSE	BIC
Holstein Friesian	MLR	FCM305=-323.7+7.22BD-0.16FCM100+22.00FY+1.72PY	0.96	0.08	1024.18
	NN	FCM305=1125.36+0.26BD-0.01FCM100+0.44FY+0.48PY	0.99	0.02	988.34
	GFA	FCM305=1681.58+3.48BD-1.58FCM100+11.23FY+44.61PY	0.87	0.20	1360.09
FRxBJ	MLR	FCM305=26.83+1.18BD-2.38FCM100+79.78FY+1.43PY	0.98	0.06	1000.01
	NN	FCM305=868.43+0.01BD+0.08FCM100-2.08FY-0.05PY	0.99	0.04	990.06
	GFA	FCM305=36.08-12.77BD-0.22FCM100+21.99FY+15.92PY	0.98	0.07	1000.26
Jersey	MLR	FCM305=1722.5-13.70RP-0.33FCM100	0.98	0.07	1000.20
	NN	FCM305=1341.3+0.38RP+0.02FCM100	0.99	0.01	980.71
	GFA	FCM305=1246.5-62.8RP-66.2FCM100	0.95	0.11	1205.09
JxBJ	MLR	FCM305=1062.1-15.22RP-1.45FCM100-0.08FY	0.94	0.13	1232.43
	NN	FCM305=2628.6+10.18RP+0.17FCM100-8.24FY	0.96	0.09	1530.07
	GFA	FCM305=988.44-35.2RP-1.80FCM100-11.06FY	0.90	0.14	1298.22

FCM305-Fatcorrectedmilk for 305day; JXBJ-JerseyxBunaji; FRXBJ-FriesianxBunaji; JxBJ-JerseyxBunaji; NN-Artificial neural network; MLR-Multiple linear regression and GFA-Genetic function al; BD-Bodydepth; FY-Fat yield; RP-Rump; PY-Protein yield; Adj R<sup>2</sup>-Adjusted coefficient of determination, Bayesian information criterion; d=day; RMSE-Root mean square error

**Table 3. Prediction equation for FCM305d from milk components and conformation traits using different models in Jos**

Genotype and herd	Models	Equation	Adj R <sup>2</sup>	RMSE	BIC
Holstein Friesian	MLR	FCM305=1012.4+3.76ST-0.24FCM100+23.53FY+0.66PY-3.18UC-0.44BWT	0.98	0.06	1020.22
	NN	FCM305=3844.4-0.02ST+0.02BD +0.01FCM100+0.009FY+0.009PY -0.003UC-0.003BWT	0.99	0.01	960.42
	GFA	FCM305=348.7+6.82ST-8.93BD-0.13FCM100+20.3FY+1.28PY-9.39UC-1.18BWT	0.87	0.14	1290.50
FRxBJ	MLR	FCM305d=1512.8+21.1FY-50.9TL	0.96	0.05	1022.36
	NN	FCM305d=868.4 -2.08FY-0.17TL	0.98	0.03	1018.41
	GFA	FCM305d=1515.9-21.1FY-51.4TL	0.96	0.07	985.55

FCM305-Fatcorrectedmilk for 305day; FRXBJ-FriesianxBunaji; NN-Artificial neural network; MLR-Multiple linear regression; GFA-Genetic function algorithm approximation; BWT-Bodyweight; ST-stature; UC-Udder clearance; FCM100-Fat corrected milk at 100 day; BD-Bodydepth; FY-Fat yield; PY-Protein yield; Adj R<sup>2</sup>-Adjusted coefficient of determination; Bayesian information criterion; d=day, RMSE-Root mean square error

**Table 4. Prediction equation for FCM305d from milk and conformation traits using different model in Adamawa State**

Genotype and herd	Model	Equation	Adj R <sup>2</sup>	RMSE	BIC
Holstein Friesian	MLR	FCM305d=3048.5-26.1RP-3.23BD+24.1FY+3.1UC-21.5RUH	0.99	0.03	983.41
	NN	FCM305d=806.9-0.14RP+0.06FY+0.01UC-0.07RUH	0.99	0.02	961.06
	GFA	FCM305d=1036.1-98.3RP+22FY+15.92UC-0.07RUH	0.98	0.05	997.28
Jersey	MLR	FCM305d=-1059.93-26.78BCS-16.20HG+10.07UC	0.96	0.08	1050.82
	NN	FCM305d= 868.4-2.08FY+0.17TL	0.98	0.05	990.01
	GFA	FCM305d= 1515.9+21.1FY-51.4TL	0.96	0.10	1100.08
Brownswiss	MLR	FCM305d=1291.8-0.37FCM100+24.9FY	0.98	0.06	988.57
	NN	FCM305d=1208.8+0.0002RP+0.02FCM100+0.04FY	0.99	0.02	985.50
	GFA	FCM305d=-1008.92+0.05FCM100+22.05FY	0.97	0.08	993.84
Simmental	MLR	FCM305d=1241.3-0.29FCM100+23.2FY	0.99	0.04	983.80
	NN	FCM305d=1428.2-0.002FCM100+0.06FY	0.99	0.03	984.12
	GFA	FCM305d=1283.1-0.04FCM100-11.03FY	0.99	0.05	983.68

*FCM305d-Fatcorrectedmilk for 305day; NN-Artificial neural network; MLR-Multiple linear regression and GFA-Genetic algorithm function approximation; FCM100-Fatcorrectedmilk for 100day milk yield; FY-Fat yield; TL-Teat length; UC-Udder clearance; RP-Rump; BD-Body depth; BCS-Body condition score, RUH-Rear udder height; HG-Heart girth; Adj R<sup>2</sup>-Adjusted coefficient of determination; Bayesian information criterion; d=day; RMSE-Root mean square error*

**Table 5. Prediction equation for FCM305d from milk and conformation traits using different models pooled for all genotypes**

Model	Equation	Adj R <sup>2</sup>	RMSE	BIC
MLR	FCM305d=1425.49-7.83BCS-1.09HG-0.04FCM100+24.0FY+0.94PY-15.0RUW	0.995	36.77	7497.88
NN	FCM305d=-2848.99+0.03RUW+0.26BCS+0.01HG-0.002FCM100-0.008FY-0.0007PY	0.997	30.14	2424.69
GFA	FCM305d=1036.1-98.3RP+22FY+15.92UC-0.07RUH	0.997	30.07	1997.28

*FCM305d-Fatcorrectedmilk for 305day; NN-Artificial neural network; MLR-Multiple linear regression and GFA-Genetic algorithm function approximation; Adj R<sup>2</sup>-Adjusted coefficient of determination; Bayesian information criterion FCM100-Fatcorrectedmilk for 100day milk yield; FY-Fat yield; UC-Udder clearance; RP-Rump; HG-Heart girth; RUH-Rear udder height; BCS-Body condition score; RUW-Rear udder width; PY-Protein yield; RMSE-Root mean square error*

The observed high adequacy of NN as the best model for predicting fat corrected 305 day milk yield using part period milk production (FCM100), milk components and conformation traits within genotype x environment interactions was consistent with the report of several authors [9, 10,11]. [12] compared qualitative properties of MLR and two different models of NN. The MLR model, depending on the region, was characterised by  $R^2$  values ranging from 78 to 86%. A classical NN showed lower  $R^2$  values (0.74–0.82), while that with polynomial post-processing demonstrated higher values of  $R^2$  (0.80–0.90), which was still lower than the range of 0.96 – 0.99 (NN) reported in this study. The NN models applied in the studies were of better predictive properties, mainly due to larger and more comprehensive datasets and more independent variables used for NN design. MLR proposed by other authors for predicting the first lactation milk yield demonstrated moderate ( $R^2 = 0.64$ ; [13]), or low  $R^2$  values ( $R^2 = 0.36$ ; [14]). [15] analysed standard models for predicting some parameters of cow milk performance in a small herd. They obtained a very high  $R^2$  (0.94 to 0.99) for the peak yield, whereas for the overall lactation yield, the coefficients were much lower:  $R^2 = 0.66$ . When the data were pooled within genotype x environmental interactions in Nigeria, all the models were excellent predictor of 305-day fat corrected milk yield and Genetic function algorithm was the best with three predictor variables and lowest variance of estimation (RMSE and BIC). [16] suggested, however, that the lowest variance of estimation error, rather than coefficient  $R^2$ , should be taken into account as a criterion of the model performance. Linear regression models are simple in design and parameter interpretation. More simple models are less sensitive to fluctuations that appear in data; however, they do not fully reflect the real course of lactation. This speaks for a wider testing of the models in practice [16]. On the other hand, predictions based on GFA and NNs as demonstrated in this study, do deviate from a classical MLR model either by their quality parameters or their predictive properties. New, additional data require a new statistical model, whereas a genetic function algorithm and neural network can undergo additional training.

#### 4. CONCLUSION

The obtained results showed that our ANN-based model approach is very promising and may play a useful role in developing a better cost-effective strategy for milk production in the

future. The best model for predicting fat corrected 305 day milk yield was the Genetic Function Algorithm (FCM305d=1036.1-98.3RP+22FY+15.92UC-0.07RUH; Adj  $R^2=0.997$ ; RMSE=30.07; BIC=1997.28) in Nigeria dairy herd. The best neural network architecture for modeling 305 day fat corrected milk yield was 6-2-1 hidden nodes in the multilayer perceptron using back propagation algorithm with 88% learning rate and 2% bias.

#### COMPETING INTERESTS

Authors have declared that no competing interests exist.

#### REFERENCES

1. AOAC. Association of official and analytical chemists. 15<sup>th</sup> Edition William Press, Richard Virginia, U.S.A.; 1990.
2. Akinsola. Estimation of phenotypic and genetic parameters of milk yield, conformation and fertility traits in dairy cattle: A multi-genotype and multi-locational study. A PhD thesis submitted to the Department of Animal Science, Ahmadu Bello University, Zaria. 2017;280.
3. Akinsola. Estimation of genetic capacity of Simmental dual purpose cattle for milk and conformation traits in high arid climate. Asian Journal of Research in Animal and Veterinary Sciences. 2018;1(2):1-8.
4. Banos G, Van Ranst B, Coffey MP. (2010). Genetic association between body energy measured throughout lactation and fertility in dairy cattle. *Animal*, 4:189-199.
5. Gains WL. The energy basis of measuring milk yield in dairy cows. University of Lllinois. Agriculture Experiment Station. Bulletin No. 1928;308.
6. FAO (Food and Agriculture Organization of the United Nations). FAO Production year book. Rome; 2013.
7. Ferreira WJ, Teixeira NM, Torres RA, Silva MVGB. Utilizacao da produo de leite no dia do controlena avaliao genetica em gado de leite - Uma reviso. Archives Latinoam. de Production Animal. 2002;10:46-53.
8. ICAR. International Committee for Animal Recording Guidelines. Approved by the general assembly held in Koupio, Finland in June 2006; 2007.
9. Jensen J. Genetic evaluation of dairy cattle using test day models. Journal of Dairy Science. 2001;84:2803-2812.



10. Sanzogni L, Kerr D. Milk production estimates using feedforward artificial neural networks, computational, electronics and agriculture. 2001;32:21–30.
11. Schaeffer LR, Jamrozik J, Kistemaker GJ, Van Doornmaal BJ. Experience with a test-day model. Journal of Dairy Science. 2000; 83:1135-1144.
12. Olori VE, Brotherstone S, Hill WG, McGuirk BJ. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. Livestock Production Science. 1999;58:55-63.
13. Shive-kumar P, Arun K, Devendra S, kumar S, Kumar A, Singh D. Prediction of first lactation milk yield from body measurements in crossbred cows. Journal of Dairying Foods and Home Sciences. 1995;41-46.
14. Pan S, Koley N, Duttagupta R, Roy AM. Prediction of milk on the basis of peak yield and initial yield in Holstein-Friesian x Haryana cows (F1). Environment-and-Ecology. 1997;4:860-861.
15. Perz WP, Sobek Z. Evaluation of selected mathematical models of lactation curves. Biometrical Letters. 1999;36:169-178.
16. Yahuza ML. Smallholder dairy production and marketing constraints in Nigeria. In: Rangneker, D. and W. Thorpe Ed. Smallholder dairy production and marketing – Opportunities and Constraints. Proceedings of a South-South workshop held at National Dairy Development board (NDDB), Anand, India, 13-16 March 2001 and ILRI (International Livestock Research Institute) Nairobi, Kenya; 2001.

---

© 2018 Oyegbile et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Peer-review history:*  
*The peer review history for this paper can be accessed here:*  
<http://www.sciencedomain.org/review-history/25923>