



ISSN: 2456-2912

NAAS Rating (2025): 4.61

VET 2026; 11(2): 01-04

© 2026 VET

www.veterinarypaper.com

Received: 06-11-2025

Accepted: 09-01-2026

All author's name and affiliations
are given below, after references

Prediction of first lactation total milk yield using part lactation performance trait in Haryana cattle

**Abhishek Saxena, Vijay Kumar, Satyendra Pal Singh, Deepak Sharma,
Avneesh Kumar, Jitendra Kumar Chaudhary, Manish Singh, Abhimanyu
Chouhan, Vidushi Aditya, Neelu Gupta and Satakshi Naithani**

DOI: <https://www.doi.org/10.22271/veterinary.2026.v11.i2a.3009>

Abstract

The present investigation was undertaken to develop prediction models for first lactation total milk yield (FLTMY) in Haryana cattle using lactation and reproductive traits. The study utilised first-lactation performance records of Haryana cows maintained at the Livestock Farm Complex, DUVASU, Mathura, during the period 1964 to 2025. Age at First Calving (AFC), first lactation length (FLL) and 30-days milk yield (30DMY) were considered as predictor variables, while FLTMY served as the dependent trait. Multiple linear regression models were developed using different combinations of predictor traits. Model performance was evaluated using coefficient of determination (R^2), root mean square error, mean absolute error, mean absolute percentage error and Theil's U statistic. Multicollinearity among predictors was examined through variance inflation factor values, which indicated absence of serious collinearity. Among seven models developed, the model including FLL and 30DMY proved most efficient, explaining 64.65 per cent variation in FLTMY with lower error statistics. Although the three-trait model showed marginally higher R^2 , the two-trait model was preferred for simplicity and practical utility. The final prediction equation developed was $FLTMY = -620.93 + 4.63(FLL) + 3.82(30DMY)$.

Keywords: FLTMY (First lactation total milk yield), AFC (Age at first calving), FLL (First lactation length), 30DMY (30-days milk yield)

Introduction

Haryana cattle, one of India's prominent dual-purpose zebu breeds, play a central role in the dairy economy of northern states owing to their adaptability, disease resistance, and suitability for smallholder production systems. The native breeding tract of Haryana cattle encompasses a large part of Rohtak, Hisar and Gurgaon districts of Haryana (Joshi *et al.*, 1995) [1]. Although the breed is valued for its resilience, its milk production levels are considerably influenced by early-life growth, reproductive performance, and management conditions. Among the various production traits, First Lactation Total Milk Yield (FLTMY) remains a critical benchmark because it provides the earliest and most dependable indication of an animal's future milk-yielding ability. Reliable prediction of FLTMY is therefore essential for informed selection, scientific heifer replacement decisions, and maximising genetic progress in Haryana herds.

Conventional selection tools for FLTMY often rely on pedigree-based estimates or single-trait assessments, which may overlook the complex interactions that shape milk performance in indigenous breeds. Haryana cattle display considerable variation in reproductive milestones, lactation dynamics, and environmental adaptability, making trait prediction more challenging but also more important. Daily milk production is a readily available and widely recorded parameter on farms adopting smart technologies, making it a primary focus of analysis (Song *et al.*, 2025) [5]. With the increasing availability of herd-level data, advanced statistical and computational approaches offer new possibilities to quantify trait relationships more accurately. Prediction models built on multivariate trait associations integrating reproductive parameters, early-lactation indicators, and biometric measures can provide a practical, field-oriented tool for forecasting FLTMY before calving.

Corresponding Author:

Amit Kumar

1) Division of Animal
Reproduction, ICAR-Indian
Veterinary Research Institute,
Izatnagar, Bareilly, Uttar Pradesh,
India

2) School of Genome Engineering,
ICAR-Indian Institute of
Agricultural Biotechnology,
Ranchi, Jharkhand, India

As per Wisnieski *et al.* (2019) ^[2] predictive modeling is a data-driven approach that efficiently integrates the most influential variables and their interactions to forecast current or future outcomes. Such models help identify promising heifers at an early stage, reduce uncertainty in breeding decisions, and improve economic efficiency for farmers and breeding organisations. The present study focuses on developing and evaluating robust prediction models specifically for Haryana cattle, with the aim of identifying potent predictor traits and generating reliable, easy-to-use equations for field applications. This work acknowledges the unique biological and management characteristics of the breed and contributes to data-driven improvement strategies for India's indigenous cattle resources.

Materials and Methods

Data source and traits studied

The present study was conducted using first-lactation performance records of Haryana cattle maintained at the Livestock Farm Complex (LFC), DUVASU, Mathura, over an extended period from 1964 to 2025. Only animals with complete and reliable first-lactation records were included in the analysis. The traits considered were age at first calving (AFC, days), first lactation length (FLL, days), 30-days milk yield (30DMY, litre) and first lactation total milk yield (FLTMY, litre). FLTMY was treated as the dependent variable, whereas AFC, FLL and 30DMY were considered as predictor variables for developing prediction models.

Data editing and preliminary analysis

Prior to statistical analysis, the data were screened to remove records with missing values or biologically inconsistent observations. Descriptive statistics including mean, standard deviation and range were computed for all traits. Outliers, when detected, were handled using the ± 3 standard deviation criterion. To ensure uniformity in scale and improve model stability, all variables were standardised before regression analysis.

Development of prediction models

Multiple linear regression models were developed to predict FLTMY using different combinations of predictor traits. The general linear regression model applied was:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Where Y represents FLTMY, X_1 , X_2 and X_3 denote AFC, FLL and 30 DMY in various combinations, β_0 is the intercept, β_i are the regression coefficients and ε is the random error term. Seven models were evaluated, including single-trait, two-trait and three-trait combinations of AFC, FLL and 30DMY. Regression coefficients were estimated using the least squares method.

Assessment of collinearity and model diagnostics

Multicollinearity among predictor variables was assessed

using Variance Inflation Factor (VIF) and tolerance values.

Model evaluation criteria

The predictive performance of different models was compared using coefficient of determination (R^2) as suggested by Saunders *et al.* (2012) ^[4], root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and Theil's U statistic as suggested by Maheta (2019) ^[3]. The best-fitting model was identified based on higher adjusted R^2 values coupled with lower error statistics, particularly RMSE and U values.

Software used

All statistical analyses and regression modelling were carried out using R, Microsoft Excel and SPSS, employing standard linear modelling procedures and diagnostic tools.

Results and Discussion

Multicollinearity among predictor variables

Multicollinearity among the independent variables considered for prediction of first lactation total milk yield (FLTMY) was assessed using the Variance Inflation Factor (VIF). The VIF values for first lactation length (1.236), age at first calving (1.008) and 30-days milk yield (1.235) were well below the critical threshold, indicating absence of serious multicollinearity among predictor traits. These low VIF values confirm the stability of regression coefficients and reliability of parameter estimates in the developed prediction models.

Table 1: VIF value for first lactation total milk yield using part lactation

Trait	VIF Value
First lactation length	1.236
Age at first calving	1.008
30 Days Milk Yield	1.235

Prediction models for first lactation total milk yield

Seven linear regression models were developed using different combinations of age at first calving (AFC), first lactation length (FLL) and 30-days milk yield (30DMY) to predict FLTMY in Haryana cattle. Estimates of intercepts, regression coefficients and model evaluation statistics including coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and Theil's U statistic are presented in Table 2.

Among the single-trait models, FLL was the most important predictor of FLTMY, explaining 56.94 per cent of the total variation. The model based on 30DMY accounted for 33.15 per cent variation, whereas AFC alone exhibited very poor predictive ability ($R^2 = 0.50\%$). These results indicate that lactation-related traits contribute more effectively to prediction of FLTMY than reproductive traits when used independently.

Table 2: Estimation of intercept and regression coefficient(s) for first lactation total milk yield (using part lactation) in Haryana cattle

Model	Variable	A	FLL	AFC	30DMY	RMSE	R2	MAE	MAPE	U
1.	FLL	-476.73	5.52	-	-	364.04	56.94	226.69	25.14	0.1558
2.	AFC	875.49	-	0.09	-	553.21	0.5	438.58	81.57	0.9098
3.	30 Days MY	273.29	-	-	7.15	453.63	33.15	362.90	61.35	0.7459
4.	FLL, AFC	-553.29	5.60	0.04	-	366.14	57.08	679.78	106.08	1.2295
5.	FLL, 30 Days MY	-620.93	4.63	-	3.82	332.28	64.65	214.64	23.83	0.5424
6.	AFC, 30 Days MY	42.76	-	0.13	7.21	453.36	34.20	357.86	60.88	0.7401
7.	FLL, AFC, 30 Days MY	-743.80	4.59	0.07	3.88	333.17	64.98	212.88	23.53	0.5399

Comparison of multi-trait prediction models

In two-trait combinations, the model incorporating FLL and 30DMY showed the highest predictive accuracy, explaining 64.65 per cent of the total variation in FLTMY with comparatively lower RMSE, MAE, MAPE and Theil's U values. The combination of FLL and AFC resulted in only marginal improvement over the single-trait FLL model, while

the AFC and 30DMY model showed inferior performance. The three-trait model including AFC, FLL and 30DMY yielded a slightly higher coefficient of determination (64.98%). However, the gain in R^2 was negligible compared to the two-trait model, while requiring inclusion of an additional predictor variable.

Table 3: Prediction equation for first lactation total milk yield

Sl. No.	Dependent variable	Independent variables	R ² (%)	Linear regression equation's
1.	FLTMY	FLL	56.94	FLTMY = -476.7 + 5.62FLL
2.		AFC	0.5	FLTMY = 875.49 + 0.09AFC
3.		30DMY	33.15	FLTMY = 273.29 + 7.15 30DMY
4.		FLL, AFC	57.08	FLTMY = -553.29 + 5.60FLL + 0.04AFC
5.		FLL, 30DMY	64.65	FLTMY = -620.93+4.63FLL+3.82 30DMY
6.		AFC, 30DMY	34.20	FLTMY = 42.76+ 0.13AFC + 7.21 30DMY
7.		FLL, AFC, 30 DMY	64.98	FLTMY = -743.80+4.59FLL+0.07AFC+3.88 30DMY

Best prediction equation

Considering both prediction accuracy and model simplicity, the two-trait model comprising FLL and 30DMY was identified as the most efficient and practical model for predicting FLTMY in Haryana cattle. The final linear regression equation developed was:

$$\text{FLTMY} = -620.93 + 4.63(\text{FLL}) + 3.82(30\text{DMY})$$

The superiority of this model highlights the biological significance of lactation length and early milk yield in determining total lactation performance. First lactation length reflects milk persistency, while 30-days milk yield serves as an early indicator of lactation potential. These findings suggest that early lactation traits can be effectively utilised for early selection and genetic improvement programmes in Haryana cattle.

Conflict of Interest: Not available

Financial Support: Not available

Reference

- Joshi BK, Tania MS, Kumar P, Gupta N, Vij PK, Nivsarkar AE, *et al.* Haryana cattle: A monograph on breed characteristics, National Bureau of Animal Genetic Resources; 1995.
- Wisniewski L, Norby B, Pierce SJ, Becker T, Sordillo LM. Prospects for predictive modeling of transition cow diseases. *Animal Health Research Reviews*. 2019 Jun;20(1):19-30.
- Maheta D. Forecasting nifty and sensex series using Theil's U statistic. *International Journal of Advanced Science and Technology*. 2019;28(16):246-251.
- Saunders LJ, Russell RA, Crabb DP. The coefficient of determination: what determines a useful R^2 statistic?. *Investigative Ophthalmology & Visual Science*. 2012 Oct 1;53(11):6830-6832.
- Song JW, Lee M, Cho H, Lee DH, Seo S, Lee WH. Development of individual models for predicting cow milk production for real-time monitoring. *Computers and Electronics in Agriculture*. 2025 Jan 1;228:109698.

All Author's Name and Details

Abhishek Saxena

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Vijay Kumar

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Satyendra Pal Singh

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Deepak Sharma

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Avneesh Kumar

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Jitendra Kumar Chaudhary

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Manish Singh

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Abhimanyu Chouhan

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Vidushi Aditya

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Neelu Gupta

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

Satakshi Naithani

Department of Animal Genetics and Breeding, College of Veterinary Sciences, DUVASU, Mathura, Uttar Pradesh, India

How to Cite This Article

Saxena A, Kumar V, Singh SP, Sharma D, Kumar A, Chaudhary JK, *et al.* Prediction of first lactation total milk yield using part lactation performance trait in Haryana cattle. International Journal of Veterinary Sciences and Animal Husbandry. 2026;11(2):01-04.

Creative Commons (CC) License

This is an open-access journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-Share Alike 4.0 International (CC BY-NC-SA 4.0) License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.