HDRF Conference 2024:
Next Generation Data-Led Dairying
Hannah Dairy Research Foundation
Journal of Dairy Research
Edinburgh, November 18th and 19th 2024

Enhancing IgG Goat Colostrum Determination Using Color-Based Techniques and Data Science

Manuel Betancor-Sánchez, Marta González-Cabrera, Antonio Morales-delaNuez, Lorenzo E. Hernández-Castellano, Anastasio Argüello, Noemí Castro
IUSA-ONEHEALTH 4, Animal Production and Biotechnology, Institute of Animal Health and Food Safety

INTRODUCTION

Newborn goats rely on colostrum intake to acquire immunoglobulin G (IgG), essential for passive immunity due to minimal maternal antibody transfer during gestation. IgG protects the neonates against infectious diseases until their immune systems develop adequately.

This research builds upon a 2005 study by Anastasio Argüello (DOI:10.3168/jds.S0022-0302(05)72849-6), which attempted to predict IgG concentration using multiple regression techniques based on colostrum colour. The goal of the present study is to enhance the predictive accuracy of those initial findings by applying advanced data science methods.

Conventional methods to measure IgG, such as ELISA and RID, are reliable but often high-cost for small-scale farmers. The objective is to provide an accessible, farm-level alternative to laboratory testing, achieving results as close as possible to laboratory standards and facilitating timely, data-driven decisions in colostrum quality management.

MATERIALS AND METHODS

A dataset of 813 colostrum samples from Majorera goats was collected, measuring both IgG concentration (via ELISA) and colour attributes in the CIE LCh colour space. The data were split into training (80%) and testing (20%) sets, ensuring reproducibility through a fixed random seed.

For model selection, one Machine Learning and one Deep Learning method were applied to predict IgG concentration: decision trees and neural networks, respectively. Each model was chosen as a representative approach in its field, aiming to balance interpretability and predictive power with respect to the dataset characteristics. Decision trees offered a structured approach with interpretable splits on colour attributes, while neural networks provided the flexibility to capture non-linear relationships.

Model performance was evaluated using MAE, MSE, ROC-AUC, accuracy, sensitivity, specificity, R^2 and NPV, with a threshold of 20 mg/mL to classify high and low IgG concentrations.

RESULTS

The predictive models demonstrated high accuracy in estimating IgG concentration based on colostrum colour attributes. Both the decision tree model and neural network outperformed the multiple regression techniques used in the original 2005 study. These models were compared using multiple evaluation metrics to assess their predictive performance, highlighting significant improvements over previous methods.

The decision tree showed slightly higher performance in terms of accuracy and sensitivity, while both models achieved results closely aligned with IgG measurements from ELISA tests.

Metric	Decision Tree	Neural Network	Previous Study
MAE	0.3206	1.1076	---
MSE	3.6571	5.1804	---
ROC-AUC	AUC = 0.97	AUC = 0.96	---
r^2	0.9644	0.9541	0.695
Accuracy	0.9816	0.9632	0.8745
Sensitivity	1.0	0.94	0.9303
Specificity	0.9741	0.9735	0.7143
NPV	1.0	0.9735	0.7813

CONCLUSIONS

This study demonstrates the effectiveness of machine learning techniques, specifically decision trees and neural networks, in predicting IgG concentration in goat colostrum through colour analysis, improving upon the results obtained with multiple regression.

The decision tree model provided a reliable, accessible alternative for on-farm IgG measurement, achieving accuracy and sensitivity levels comparable to laboratory-based methods, allowing for cost-effective colostrum quality assessment that supports improved animal health and welfare.

This approach can be integrated directly into farm practices, enhancing decision-making and fostering a more efficient approach to animal care.