Machine Learning-Based Prediction of Milk Yield from Early-Life Data

C. Ferrari[1], A. M. Vergani[2], C. Punturiero[1], A. Delledonne[1], M. G. Strillacci[1], A. Bagnato[1]

(1) Department of Veterinary and Animal Science – Università degli Studi di Milano, Lodi, Italy
(2) Politecnico of Milan, Milan

carlotta.ferrari@unimi.it

Artificial Intelligence 4 Animal Science

EAAP European Federation of Animal Science

1. Introduction

Early-life prediction of milk yield is a key innovation in dairy herd management. Machine learning (ML), especially when leveraging genomic breeding values (gEBV) and phenotypic traits, offers new tools to forecast performance before first lactation. This study validates and expands a previously developed neural network model to predict 305-day milk yield using only early-life data.

2. Materials and Methods

- **Model**: Feed-forward neural network with one hidden layer trained on 502 cows, using gEBV, parity, days in milk, age at calving, and month of calving.
- **Validation Set**: 252 lactation curves from genetically related cows not seen during training.
- **Simulations:** alteration of calving month and age (±2 and ±4 months) to evaluate impact on total milk yield.

Table 1 - MSE and MAE of FFNN 1 HL model evaluated on simulated data vs. validation set.

Metric	FFNN 1 HL (Simulation)	Neural network (Validation)
MSE [kg²]	41.97 (σ = 4.61)	35.92 (σ = 52.96)
MAE [kg]	5.08 (σ = 0.31)	4.80 (σ = 3.58)

3. Results

Figure 1 - Predicted lactation curves for baseline and simulated calving month shifts (±2, ±4 months), compared to observed production.

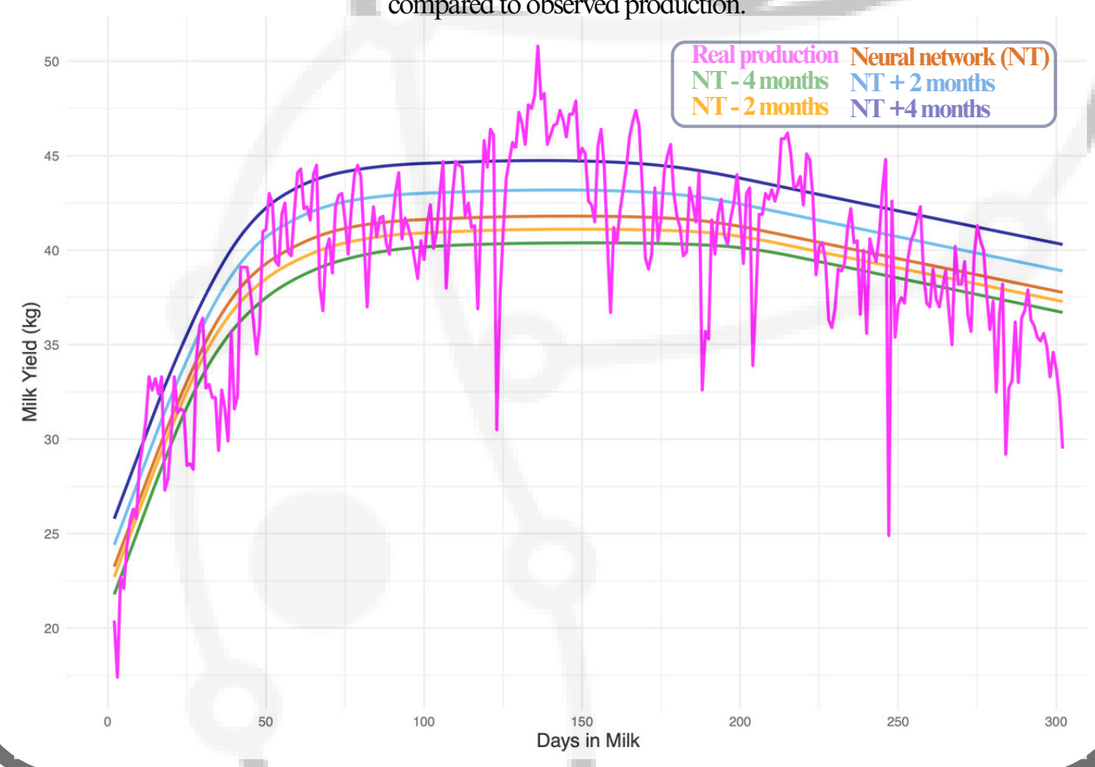

Table 2 - Cumulative MAE values for simulated calving shifts (−4, −2, +2, +4 months) compared to baseline

Simulation	MAE
Neural_network simulation	1012.96
Neural_network simulation -2	1185.10
Neural_network simulation -4	1298.58
Neural_network simulation 2	1044.48
Neural_network simulation 4	1063.75

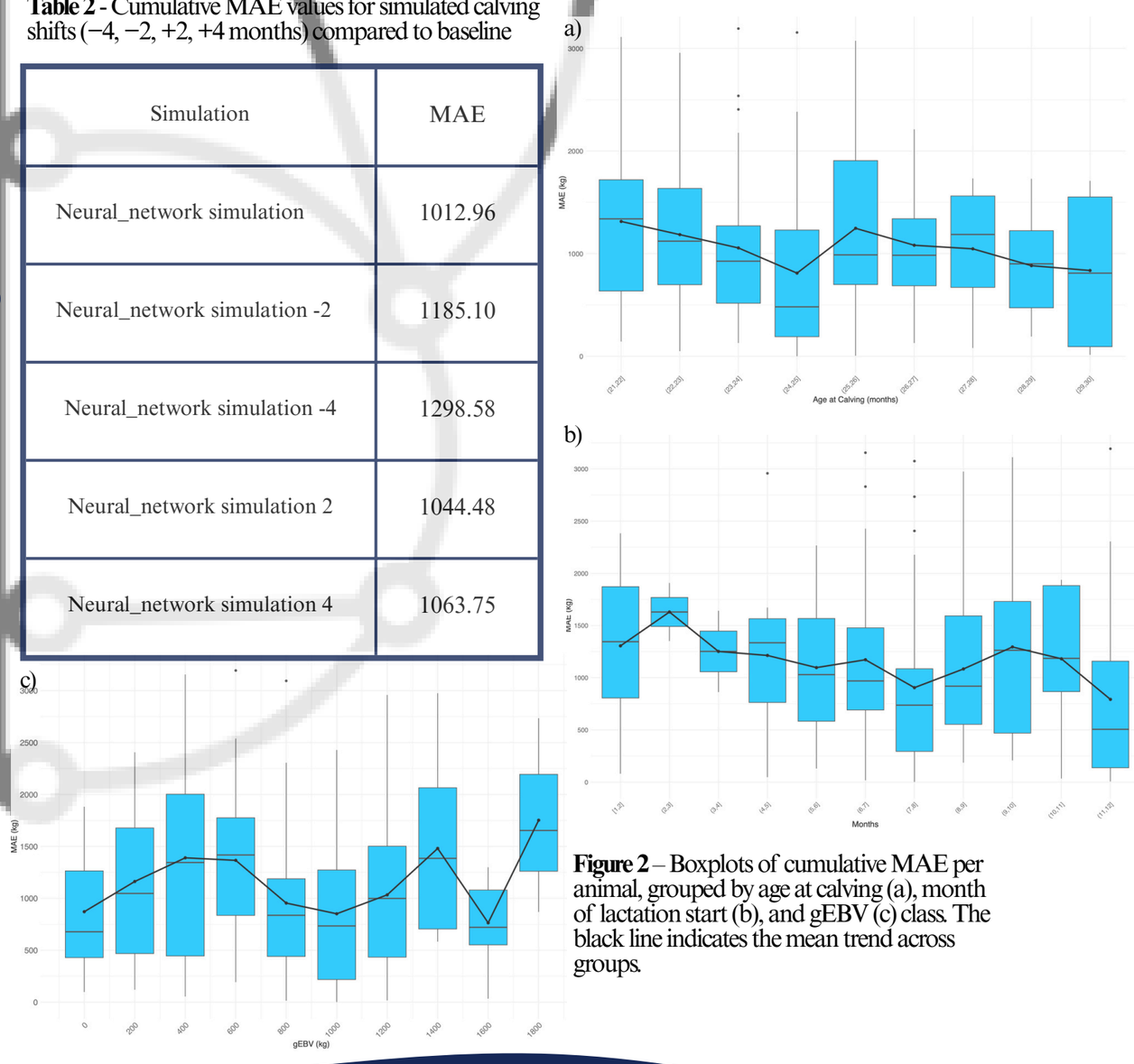

Figure 2 – Boxplots of cumulative MAE per animal, grouped by age at calving (a), month of lactation start (b), and gEBV (c) class. The black line indicates the mean trend across groups.

4. Discussion and conclusions

- This study demonstrates that a feed-forward neural network trained on early-life data can accurately predict 305-day milk yield in dairy cow
- Validation results (MAE = 4.80 kg) slightly outperform simulated data (MAE = 5.08 kg), showing strong generalization.
- These results highlight the value of machine learning for early selection and improved herd management decisions

Funded by the European Union

This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022.
This work was supported by COST (European Cooperation in Science and Technology) under COST Action CA22112 – EU-LI-PHE