Decoding lipolysis through individual free fatty FT-MIR model: Enhancing predictive accuracy through real-sample enrichment and advanced algorithms

Christophe Octave*[1], Veselko Didier[2], Leblois Julie[3], Dehareng Frédéric[1], Soyeurt Hélène[4]

[1] CRA-W, [2] Comité du lait, [3] Awé group, [4] TERRA Teaching and Research Center

Lipolysis is a longstanding and topic of interest within the dairy industry. Having been extensively researched throughout the twentieth century, lipolysis is often prematurely dismissed as a settled or outdated field of study. In recent years, the advent of robotic milking systems and their related increase of milking frequencies have led to a rise in lipolysis levels, sparking renewed interest in monitoring milk quality. For individual cows, this monitoring could help to select cows who could resist to lipolysis and/or highlight bad farming practices. Predicting lipolysis more effectively relies on the analysis of free fatty acid (FFA) concentrations.

Recent research has demonstrated the feasibility of quantifying individual free fatty acids in milk; specifically, the use of Fourier Transform mid-infrared (FT-MIR) spectroscopy on milk has shown promising results in predicting these traits. But this research has several limitations such as the use of unofficial reference analysis and the use of artificial lipolysis samples to gain in variability. To overcome these limitations, individual FFA models developed previously have been used to select, through the Walloon milk recording, 165 samples with a predicted high lipolysis level. Moreover, 41 samples were collected at the local farm of the CRA-W. All samples have been analyzed using the official reference analysis. FFA FT-MIR models built using Partial Least squares (PLS) regression and only the new dataset give the same performance than the literature. To improve the prediction performances as non-linear relationships can exist, advanced machine learning algorithms were tested using the full dataset (N=1004).

PLS regression, Kernel Ridge Regression (KRR), Gaussian process regression (GPR) and Support vector machine Regression (SVR) give better performance than other literature models with R^2CV of 0.72, 0.5 and 0.45 for Short Chain FFA, Mid Chain FFA and Long chain FFA respectively. Adding real lipolysis samples lead to a more robust models without changing the performances. This refinement of lipolysis has resulted in improved oversight of agricultural practices and enhanced valorization of milk for the purpose of its transformation.