Robustness of derivative-based AI pipelines for inferring herbage consumption from milk FT-MIR indicator

Killian Dichou[1]*, Didier Veselko[2], Antonino Marvuglia[3,] and Hélène Soyeurt[1]

[1] TERRA Research and Teaching Centre, Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium

[2] Comité du Lait. Route de Herve, 104. B-4651 Battice (Belgique)

[3] Luxembourg Institute of Science and Technology (LIST), 4362 Esch-sur-Alzette, Luxembourg

At present, the dairy sector relies on grazing calendars and farm inspections to assess grazing practices for milk production. These methods are labour-intensive, costly, and hard to standardise at scale. Indicators derived from Fourier-transform mid-infrared (FT-MIR) milk spectra provide a promising alternative for automated, reproducible monitoring of feeding practices. This approach could enable to count grazing days, supporting verification of pasture-related labels. In a previous work, we developed a continuous indicator derived from milk constituents' estimates, including fatty acid profiles, protein fractions, and acidity-related traits obtained from FT-MIR analysis of bulk tank milk. This indicator reflects the proportion of herbage in the cows' diet on a scale from 0 to 1.

In this study, we assess the robustness of a rule-based, derivative-driven AI pipeline designed to infer herbage consumption dynamics from FT-MIR–derived indicators. Derivative-based rules are used to extract discrete grazing periods from continuous time series. Using data from 2181 dairy farms in Wallonia for 2023 and 2024, we evaluate the influence of key algorithmic parameters, including smoothing intensity (4,7 and 10 days) and derivative order (4, 7 and 10 days), on inferred metrics. Results show that inferred grazing duration was moderately sensitive to parameter choices, with mean annual grazing duration across farms ranging from 193.5 to 222.3 days depending on algorithmic settings, while between-farm variability remained stable (SD ≈ 40–46 days).

This work highlights both the potential and the limitations of low-complexity, explainable algorithms such as this rule-based, derivative-driven AI pipeline. Although this study lacks precise validation based on in-farm data, it demonstrates clear potential as a screening tool that could be applied to large database to determine the associated number of grazing days, which may prove particularly useful for pasture-related labels.