Prediction of dry matter intake of dairy cows across multiple countries and production systems

A. Tedde,[1] C. Grelet,[2] P. N. Ho,[3] J. E. Pryce,[3,4] D. Hailemariam,[5] Z. Wang,[5] G. Plastow,[5] N. Gengler,[1] E. Froidmont,[2] F. Dehareng,[2] C. Bertozzi,[6] GplusE Consortium, H. Soyeurt[1]

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

[2]Walloon Agricultural Research Center (CRA-W), B-5030 Gembloux, Belgium

[3]Centre for AgriBioscience, AgriBio, Agriculture Victoria, Bundoora, Victoria 3083, Australia

[4]La Trobe University, 5 Ring Road, Bundoora, Vic. 3083, Australia

[5]Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB T6G 2P5, Canada

[6]Walloon Breeding Association, 5590 Ciney, Belgium

We predicted Dry Matter Intake of dairy cows using parity (PRT), weeks of lactation (WOL), milk yield (MY), milk mid-infrared (MIR) spectra, and MIR based predictions of bodyweight (BW) and fat, protein, lactose, and fatty acids contents in milk. The dataset comprised 10,711 samples of 534 dairy cows with a geographical diversity (Australia, Canada, Denmark, and Ireland), from first to third+ parity covering 44 weeks of lactation. When they were used individually, we found that the most contributing predictors were MIR spectra, bodyweight, and milk yield with cross-validation coefficient of determination (R^2_{CV}) > 20%, followed by fatty acids and parity with R^2_{CV} around 18%, while the weeks of lactation and the combination of fat, protein, and lactose were associated with R^2_{CV} of 5%. Using Partial Least

Square regression (PLS) with performances evaluated by cow-independent 10-fold cross-validation (CV) repeated ten times, we achieved the best Root Mean Square Error of cross-validation ($RMSE_{CV}$) of 3.24 ± 0.08 kg with equations MY + MIR + PRT, and MY + MIR + BW. Because the Australian cows' milk production and composition were significantly different and the milk yield was the second most important predictor in the PLS model, we did a second calibration procedure by grouping the data by providers. The $RMSE_{CV}$ varied from 2.73 ± 0.05 kg to 3.33 ± 0.08 kg, which is in line with the literature. The performances with out-of-sample RMSE employing the data outside the calibration set ranged from 3.49 to 7.68 kg, confirming the interest in combining the datasets across countries.