

# Selection of the most appropriate predictors to improve the robustness of dairy cows' bodyweight prediction by using lactation stage, parity, milk yield, and milk mid-infrared spectra.

A. Tedde, C. Grelet, P. N. Ho, J. E. Pryce, D. Hailemariam, Z. Wang, N. Gengler, Y. Brostaux, E. Froidmont, I. Dufrasne, C. Bertozzi, F. Dehareng, M. A. Crowe, GpluSE Consortium, H. Soyeurt

From the dairy farmer's perspective, knowing the body weight (BW) of a cow at a specific moment or measuring its evolution through time is of great interest. Various methods exist to assess BW, but either they are too expensive and need costly and time-consuming maintenance to be in production at a large scale, or their application occurs only once during the animal lifetime. The current work was mainly inspired by the published equations in [1], in which BW was predicted using days in milk, parity, season, milk yield, and milk mid-infrared (MIR) spectrum. The study's main goal was to reduce the number of predictors while preserving models' performances or improving them through modelizations involving a nested calibration and validation pipeline. Minimizing the number of predictors brings to a less complex model that reduces the risk of overfitting. The most relevant variables were picked through feature selection techniques such as filter and wrapper methods combined with partial least squares regression models to predict BW. From 1,849 records collected in 9 herds comprising 360 Holstein cows, the best performing models achieved a root mean square error (RMSE) of herd-independent validation from  $52 \pm 2.34$  kg to  $56 \pm 3.16$  kg, including between 5 to 62 predictors. Amongst these models, 3 performed remarkably well in external validation, across a new and independent dataset (N=4,067), by featuring an RMSE of validation of 52 and 56 kg. The predicted BW available repeatedly through the lactation based on milk recording could open new perspectives in developing management and breeding tools.

[1 H. Soyeurt, E. Froidmont, I. Dufrasne, D. Hailemariam, Z. Wang, C. Bertozzi, F.G. Colinet, F. Dehareng and N. Gengler, "Contribution of milk mid-infrared spectrum to improve the accuracy of test-day body weight predicted from stage, lactation number, month of test and milk yield," *Livestock Science*, pp. 82-89, 2019.