6. Milk Analysis Workshop 1 Securing Value from Milk Analysis

Title presentation

Large-scale phenotyping from milk MIR spectra: challenges to obtain reliable predictions

Author(s)

C. Grelet, P. Dardenne, H. Soyeurt, J.A. Fernandez, N. Gengler, A. Vanlierde, F. Stevens, F. Dehareng

Institution for which the first author of this abstract is working

Walloon Agricultural Research Centre, B-5030 Gembloux, Belgium

Gembloux Agro-Bio Tech, ULiège, B-5030 Gembloux, Belgium

Gembloux Agro-Bio Tech, ULiÃ¨ge, B-5030 Gembloux, Belgium

Abstract

In the recent years, the research aiming to predict new phenotypes from the FT-MIR analysis of milk was very active. Models were developed to predict phenotypes such as fine milk composition, cow environmental impact and health or technological properties of milk. Those models could be of great interest in order to perform genetic studies, herd or industrial process management as they could allow generating large amount of data at large scale and with reasonable cost. To achieve this, it is nonetheless necessary to insure that the models provide reliable predictions when applied on the large diversity of spectral data met on real field conditions. The robustness of models -its capacity to be “all terrain” and provide good results in various conditions- is therefore essential to ensure reliability of predictions. Robustness could be estimated by evaluating the error in external validation (RMSEP), the reproducibility of predictions between instruments and the ability of the calibration dataset to cover the variability of routine field data. However, in current literature, the model robustness is often omitted and the factors affecting robustness are not documented. Additionally, only a limited number of models is used in routine and faces the large variability of real field conditions. Finally, there is a lack of methods and indicators to evaluate the quality of models and predictions before use for genetic studies, herd or industrial process management. The objective of this work is consequently to evaluate the impact of different factors influencing robustness on prediction quality. The impact of model complexity, spectral regions used and inclusion of variability in the model regarding the spectral and the reference data are investigated. Results show that accuracy of predictions (RMSEP) can be improved from 8% to 52%. Methods and indicators to evaluate the quality of models and predictions generated in routine conditions are discussed as well. The obtained results encourage for development of guidelines and good practices, as well as for international collaborations in order to constitute large and robust datasets and enable the use of models in routine conditions.