Prediction of grass-based diet from indirect traits using milk MIR-based predictors to assess the feeding typology of farms

H. Soyeurt[1], C. Gerards[1], C. Nickmilder[1], S. Franceshini[1], F. Dehareng[3], D. Veselko[4], C. Bertozzi[5], J. Bindelle[1], N. Gengler[1], A. Marvuglia[6], A. Bayram[6,7] and A. Tedde[1,2]

[1]TERRA, Gembloux Agro-Bio Tech, University of Liège, Passage des déportés 2, 5030 Gembloux, Belgium
Corresponding Author: hsoyeurt@uliege.be
[2]FNRS, Rue d'Egmont 5, 1000 Brussels, Belgium
[3]Walloon Agricultural Research Centre, Valorisation of Agricultural Products, Chaussée de Namur 24, 5030 Gembloux, Belgium
[4]Comité du lait, Route de Herve 104, 4651 Battice, Belgium
[5]Walloon Breeding Association, Rue des Champs Elysées 4,5590 Ciney, Belgium
[6]Luxembourg Institute of Science and Technology (LIST), Rue du Brill 41,4422 Belvaux, Luxembourg
[7]Computational Sciences, Faculty of Science, Technology and Medicine, University of Luxembourg, Avenue de l'Université 2, 4365 Esch-sur-Alzette, Luxembourg

Abstract

Grassland provides access to a local and low-cost food resource usable by livestock. Today, for some brands of dairy products, one of the specifications asked the farmers is to put their cows on pasture for a minimum period or feed their cows mainly with a grass-based diet. But this constraint is rarely verified. However, this kind of feeding has a fingerprint in the milk. So, this work aims to develop a method predicting indirectly the level of grass in the cows' diet using features estimated from milk mid-infrared (MIR) spectrometry. More than 3 million records were collected between 2011 and 2021 on 2,449 farms. Those included the fat, protein, lactose, and urea contents estimated by the spectrometers and 34 MIR-based predictors reflecting the milk fatty acid composition, the protein fraction, the minerals, and the lactoferrin content. As no grazing calendars and detailed feed composition were available at large scale, the innovative aspect of this work will consist of estimating the grass-based diet using a trait defined from the month of analysis of bulk milk. Indeed, in Wallonia, nearly all herds are on pasture between the beginning of sprint and the beginning of autumn. The data were collected between 2019 and 2020. The calibration set used 80% of the farms chosen randomly (345,223 records). The remaining ones were used to validate the model (85.069 records). The direct prediction of test month from a Partial Least Square discriminant analysis (PLS-DA) was poor (0.40). However, a hierarchical clustering applied to those obtained predictions revealed a GRASS group composed of predictions done in June, July, August, and September. So, a second PLS-DA was realized to discriminate the GRASS group. A better accuracy was good (0.88). Then, the probability of belonging to the GRASS modality was used to observe the feeding typology of the farm. The evolution of this probability through the year was the one expected with an increasing value from April and a constant decrease after summer. The interval between this increase and decrease could be used to count the number of days spent by cows on pasture and the intensity of this change can be studied discriminate different feeding strategy. In conclusion, using the MIR analysis done for the milk payment by dairy companies, it is feasible to detect the presence of grass in the cow diet and estimate potentially the number of days spent by the cows on pasture during a year.

Keywords: Grass, milk, mid-infrared.

Introduction

In the Southern part of Belgium, grassland represents a large part of the useful agricultural surface. Permanent grasslands represent 42.1% of this surface (309,180 ha) and temporary grasslands represent 5% (State of the Walloon Environment, 2020).

Grassland has an important place in our landscape and fulfils many roles. But grassland provides also access to a lower cost feed resource, requiring very little supplementation for ruminants (Kolver, 2003). Therefore, producing milk from pasture makes sense. Many specifications for milk and protected designation of origin products require farmers to have a minimum period of grazing for their cows. For example, in France, CANDIA recommends an average of 150 days of grazing per year and for at least 6 hours per day. The "Lait de pâturage" label is based on the same rules but also requires a minimum grazing area of 10 ares per animal. Finally, the "Grand Pâturage" milk requires its members to have cows that have access to at least 180 days of pasture per year. In Belgium, the "MARGUERITE HAPPY COW" chain advocates feeding cows with a minimum of 70% grass in the ration and at least 180 days of grazing, for a cattle load of 4 animals per ha (Servais, 2015).

However, even though these specifications exist, there is no real verification of these conditions. Thus, recently, the Chronopâture tool was developed to automatically count the number of days that the herd spends grazing, using a GPS collar. Unfortunately, such equipment can hardly be developed on a large scale, as it depends on the choice of the breeder. An alternative could be to observe more specifically the variations of milk composition according to the feed. For example, Coppa *et al* (2021) demonstrated that it was possible to predict the composition of the cow's ration and to identify PDO practices using mid-infrared spectrum (MIR) analysis of milk, but they did not attempt to predict the number of days on pasture. This will be the aim of this work with the innovative aspect to use an indirect reference to predict the grass-based diet.

Materials and methods

The data come from the FuturoSpectre agreement linking Gembloux Agro-Bio Tech (University of Liège, Gembloux, Belgium), the Walloon Agricultural Research Centre (CRA-W, Gembloux, Belgium), the milk laboratory 'Comité du Lait' (Battice, Belgium), and the Walloon Breeding Association (AWé, Ciney, Belgium). Milk samples were collected in the bulk tank of 2,868 Walloon dairy farms between 2019 and 2020. The dataset contained 430,292 observations. All milk spectral data were generated by MilkoScan FT6000 MIR spectrometers (Foss Electric A/S, Hillerød, Denmark). Equations were applied on the recorded spectral data to estimate 38 traits related to the milk fatty acid (FA) profile, composition of mineral and protein as well as the content of lactoferrin and beta-hydroxy butyrate. All equations have a prediction R^2 higher than 65%. The FA contents initially predicted by MIR in g per dl of milk were converted into g per 100 g of fat using the fat content provided by the spectrometer.

The prediction of grass-based diet is not directly possible as we do not have data on grazing management. However, in the Walloon Region of Belgium, there is grazing between the beginning of spring and the beginning of autumn. So, the hypothesis formulated in this work is that there is an indirect link between the month of the milk analysis and the presence on pasture. This month of test was so predicted using Partial Least Square (PLS) regression. To allow an external validation of the developed prediction equations, the initial dataset was divided into two parts: 80% training set and 20% validation set. The former includes 2,287 farms (345,223 records) while the latter includes 581 farms (85.069 records). To ensure the data independence, the partition was performed with the constraint that the same farm could not be found in both sets. Based on the obtained predictions for the month, a hierarchical clustering was performed to highlight a "GRASS" group and a "NOGRASS" group. Then a second prediction model was performed to predict the GRASS modality. The prediction quality of the model was assessed by calculating the accuracy.

Results and discussion

The accuracy obtained for the prediction of the month of test based on the used features was poor and equal to 0.394 ± 0.002. The accuracy of the external validation was of 0.398. However, a strong confusion existed between some months (Figure 1). Based on the dendrogram created after the hierarchical clustering done on the predicted values, it appeared clearly 2 groups. The first group that covered June, July, August, and September can be attributed to predicted of grass-based diet (i.e., GRASS modality). Therefore, a second PLS regression was built to predict the GRASS modality. The obtained prediction was, as expected, strongly better with an accuracy of 0.874 ± 0.002. The validation accuracy was similar (0.876). This confirms the findings of Frizzarin *et al*. (2021) and Coppa *et al*. (2021) who mentioned that PLS-DA has very good ability to predict cow diet from the MIR spectrum.

The feeding fingerprint explains this capacity to discriminate the grass-based diet in milk spectral data. According to Chilliard *et al*. (2000), fresh grass is the richest food in C18:3 fatty acids. Therefore, its ingestion by cows will confer a particular composition to the milk. A literature review done by Elgersma *et al*. (2006) revealed

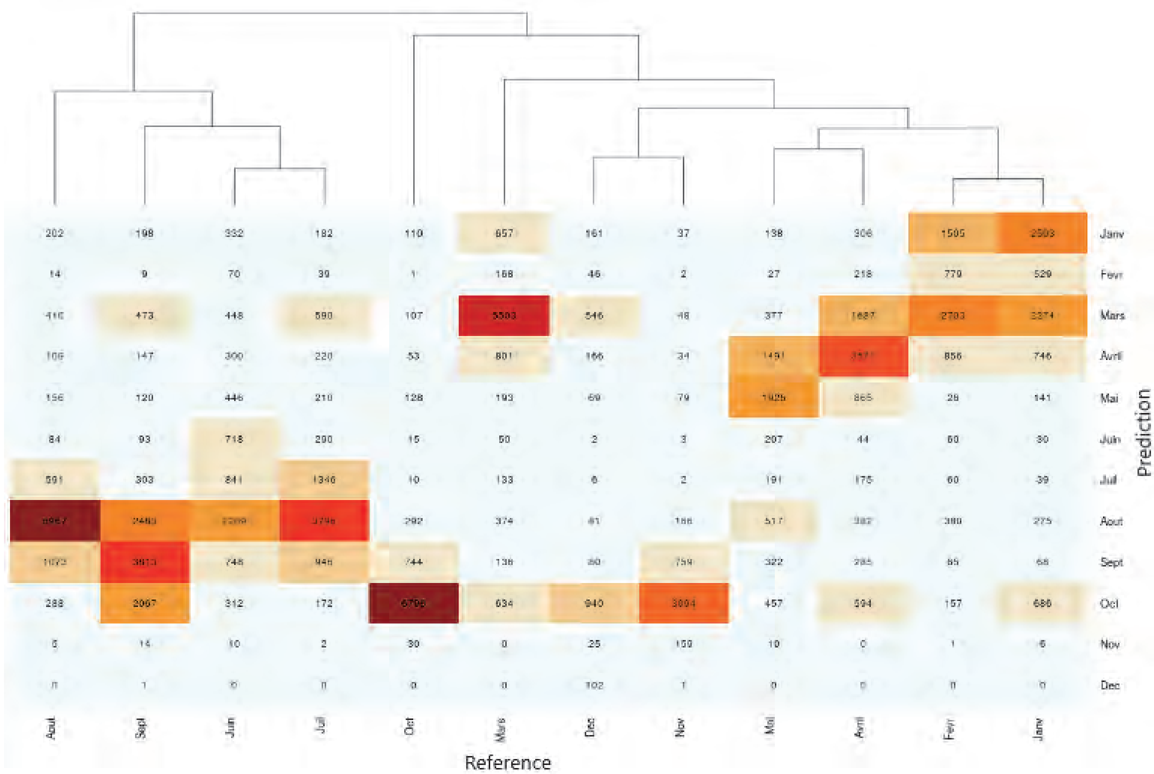

Figure 1. Dendrogram and confusion matrix between predicted and observed months.

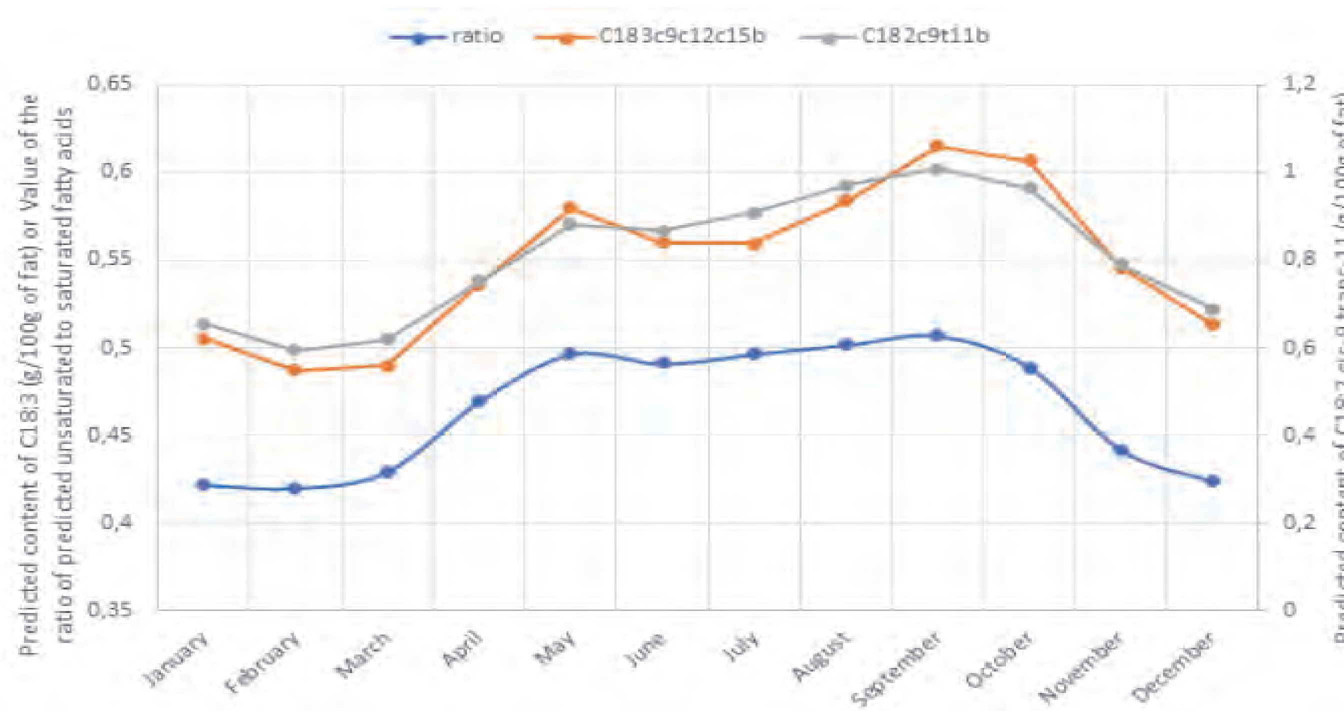

Figure 2. Yearly evolution of the ratio of predicted unsaturated to saturated fatty acids and the predicted content of C1 cis-9,trans-11 and C18:3 cis-9, cis-12,cis-15 in milk fat.

that the milk of cows fed with fresh green forage has a higher ratio of unsaturated to saturated fatty acids and contains a higher proportion of polyunsaturated fatty acids and conjugated fatty acids (CLA) than the milk of cows fed with a total mixed ration. Even if the prediction accuracy for those FA traits by FT-MIR spectroscopy is moderate, the same trends were observed from our predicted FA values during the grazing period (Figure 2). Moreover, Frelich *et al*. (2012) observed an increase in long-chain FAs (LCFA) and a decrease in short- (SCFA) and medium-chain FAs (MCFA) in the milk fat when the feeding of cows is based on fresh grass. Again, this was also confirmed by our predicted FA values (Figure 3).

The presence of grass-based diet is not a binary variable. Indeed, there is a feeding transition between winter and summer diets. The composition of the milk will therefore evolve gradually. Therefore, the use of a binary prediction does not make sense.

Thus, it is much more relevant to look at the evolution of the prediction probability related to the presence of grass-based diet. Figure 4 represents the distribution of the GRASS PLS probability following the week of test in 2019. The annual evolution of these probability values has an expected trend with an increasing probability from April, which represents the passage to pasture, and then a decrease when the winter arrived, signifying the return of cows to the barn and a total mixed ration as feeding. So, the interval between the increase and decrease of the probability of belonging to the GRASS modality could be used to calculate the number of days spent by dairy cows on pasture.

8:2

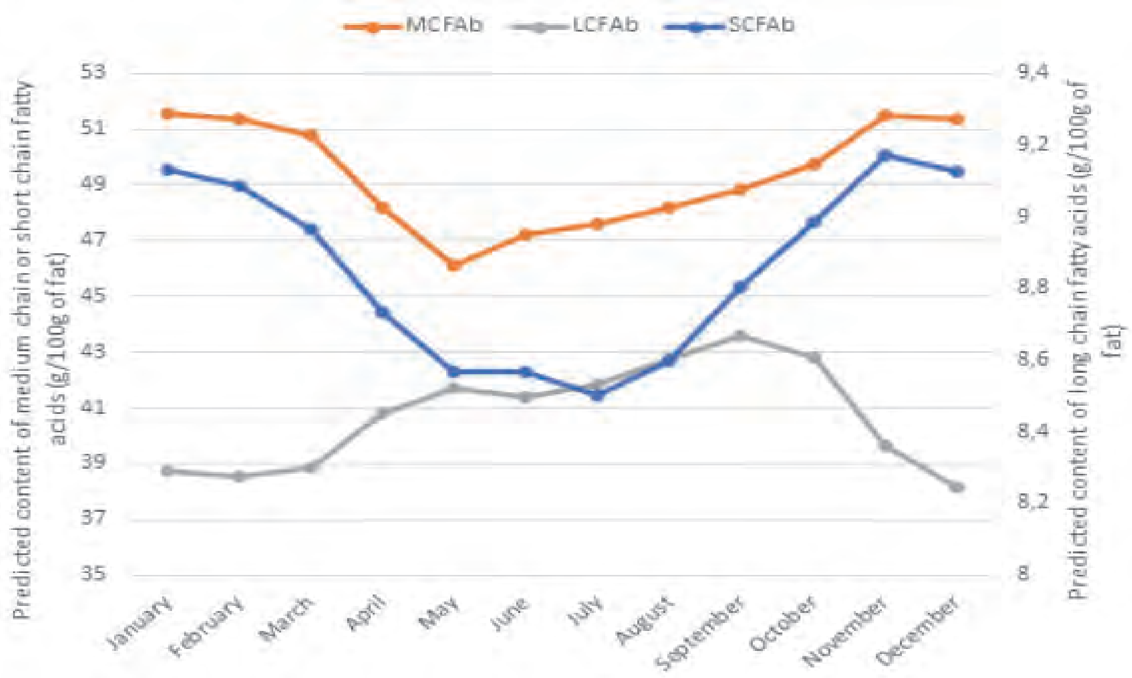

Figure 3. Yearly evolution of the predicted content of short, medium, and long chain fatty acids in milk fat.

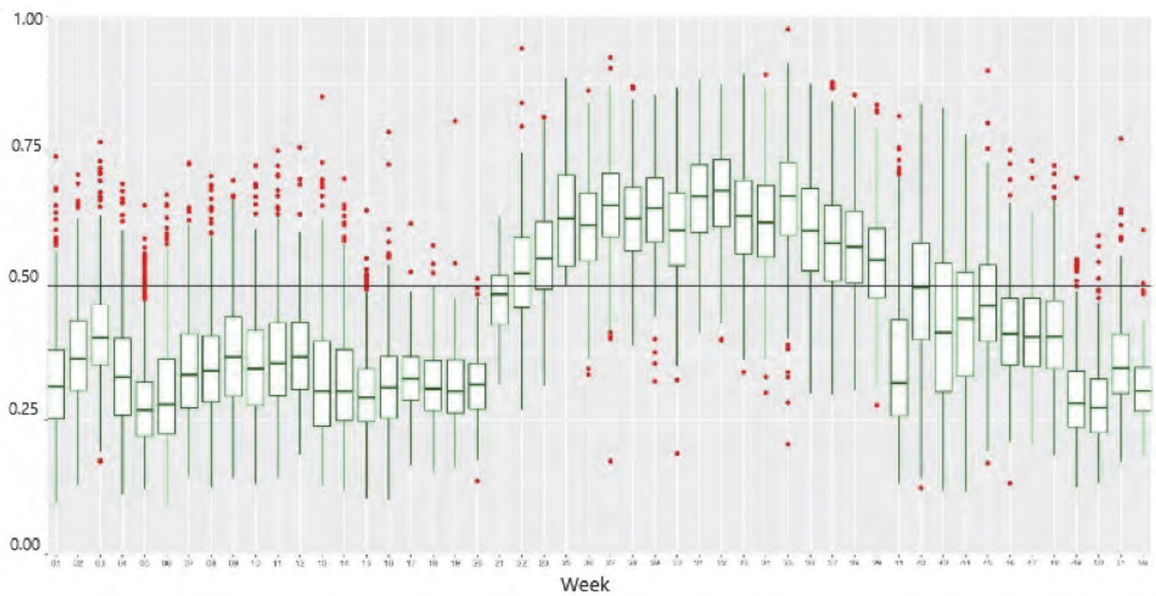

Figure 4. Boxplot of probabilities of belonging to the GRASS group in 2019 according to weeks (N = 101,442).

Conclusion

The objective of this work was to contribute to the development of a tool for counting days on pasture for dairy cows on a given farm automatically. To achieve this objective, model was built to predict the presence of grass-based diet from the spectral data provided by the analysis of milk collected in the bulk tank. As no grazing calendar was available, this innovative aspect of this study was to propose the development of a model predicting an indirect trait, the test month. The month confusion during the prediction revealed the possibility to create a GRASS group. So, the second model was built to predict this group with a good accuracy. The evolution of the PLS probability had the expected trend and suggest the possibility to count the number of days spent on pasture. This measurement as based on the milk mid-infrared spectrometry presents the advantage to be cheap and available for all farms delivered their milk to a dairy company. This calculation is also interesting as the frequency of data acquisition is high. Indeed, the bulk tank milk samples are analysed every 1 or 4 days in each farms delivered milk to the dairy company.

Acknowledgment

This work was supported by the Luxembourg National Research Fund (INTER/FNRS/18/12987586/SimBa) and by the Fonds de la Recherche Scientifique-FNRS under Grant n° T.0221.19.

References

Chilliard, Y., Ferlay, A., Mansbridge, R.M., Doreau, M., 2000. Ruminant milk fat plasticity:nutritional control of saturated, polyunsaturated, trans and conjugated fatty acids. Ann. Zootech. 49, 181–205.

Coppa M., Revello-Chion A., Giaccone D., Ferlay A., Tabacco E. and Borreani G., 2014. Comparison of near and medium infrared spectroscopy to predict fatty acid composition on fresh and thawed milk. Food Chem. 150, 49–57.

Frizzarin M., O'Callaghan T.F., Murphy T.B., Hennessy D. and Casa A., 2021. Application of machine-learning methods to milk mid-infrared spectra for discrimination of cow milk from pasture or total mixed ration diets. J. Dairy Sci. 104, 12394-12402.

Elgersma, A., Tamminga, S., Ellen, G., 2006. Modifying milk composition through forage. Anim. Feed Sci. Technol. 131, 207–225.

Frelich, J., Šlachta, M., Hanuš, O., Špièka, J., Samková, E., Zapletal, P., 2012. Seasonal variation in fatty acid composition of cow milk in relation to the feeding system. Animal Science Papers and Reports 30, 219-229.

Kolver, E.S., 2003. Nutritional limitations to increased production on pasture-based systems. Proc. Nutr. Soc. 62, 291–300.

Servais L. (Awé asbl), 2015. Marguerite happy cow une nouvelle filière laitière durable. Wallonie Elev. - Mai 2015 51–52.