Artificial neural networks outperform partial least squares models for fatty acids estimation using FT-MIR spectrometry on bovine milk: a first step towards transfer learning on small ruminants.

E. Cabaraux[1,2], I. Alexakis[1], F. Sébastien[1], O. Christophe[3], C. Grelet[3], M. Calmels[4], C. Lecomte[5], C. Bertozzi[6], D. Veselko[7], F. Dehareng[3], H. Soyeurt[1]

[1] TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, Passage des déportés [2], 5030 Gembloux, Belgium, 2 Scientific Research Fund, FRS - FNRS, Rue d'Egmont 5, 5000 Bruxelles, Belgium, [3] Walloon Agricultural Research Centre, Chaussée de Namur 24, 5030 Gembloux, Belgium, [4] Seenovia, Boulevard des Loges 141, 53940 Saint-Berthevin, France, [5] Eliance, Rue de Bercy 149, 75595 Paris, France, [6] Walloon Breeding Association, ue des Champs Elysées 4, 5590 Ciney, Belgium, [7] Comité Du Lait, Route de Herve 104, 4651 Battice, Belgium

Fourier transform mid-infrared (FT-MIR) spectrometry in milk can improve production, milk quality and animal welfare among others. Nevertheless, this technology remains underused in dairy ewes and goats, despite it becoming increasingly common in the dairy cattle sector. To reduce the cost of developing equations, transfer learning could be used to adapt FT-MIR models from dairy cattle to dairy sheep and goats. However, this transfer requires artificial neural networks (ANN) but these bovine ANN models do not yet exist. This study aims to develop a multi-layer perceptron ANN for fatty acids to see if we can reach at least the same performance as the ones obtained with commonly built partial least squares (PLS) regressions. The ANN models developed for bovine milk were then directly applied to goat milk to evaluate the potential of transductive learning, the simplest transfer learning method. The database (n=1942) contained 42 individual and groups of fatty acids measured by gas chromatography and their associated standardized bovine milk spectra. A total of 80% was used to train models and 20% used for validation. Validation R^2 values ranged from 0.1223 to 0.9890 for PLS models and from 0.2050 to 0.9886 for ANN models. Compared with PLS models, ANN reduced prediction errors (RMSE) by 0.08 to 47.27% for 39 fatty acids, while an increase in errors ranging from 1.74 to 3.38% was observed for 3 fatty acids. These results highlight an overall improvement in the predictive performance of neural networks compared to PLS models on the validation dataset. After applying the built ANN models on the goat dataset (n=345), only 10 fatty acids have a positive R^2 from 0.1228 to 0.9410, one fatty acid showed an R^2 equal to zero, while the remaining 28 fatty acids have aberrant R^2 values, highlighting the lack of model generalization across species. In conclusion, these results highlight an overall improvement in the predictive performance of ANN compared to PLS models to estimate fatty acids by FT-MIR spectrometry for dairy cattle. Furthermore, transductive transfer learning confirms that models created specifically for dairy cattle cannot be applied as they are to small ruminants and opens the door to testing more sophisticated transfer learning methods.