The Prediction of Dairy Cattle Methane Emissions on Commercial Farms: Machine Learning vs Statistical Analysis

The Prediction of Dairy Cattle Methane Emissions on Commercial Farms: Machine Learning vs Statistical Analysis

Stephen R Ross[1, 2], Huiru J Zheng[1], Haiying Wang[1], Tianhai Yan[2], Masoud Shirali[2, 3]

[1] School of Computing, Ulster University, Belfast, UK
[2] Agri Food and Biosciences Institute, Hillsborough, UK
[3] School of Biological Sciences, Queens University, Belfast UK

ross-s24@ulster.ac.uk

Methane (CH4) emissions produced by dairy cattle (DC) are a key contributor to global warming. Traditionally, Statistical (SA) methods have been employed in the prediction of DC CH4 emissions. However due to the smart farming revolution, the scale and variety of complex animal information now available for their prediction has grown exponentially, and within them are likely to exist non-linear relationships which these traditional SA models may struggle to capture. Instead, this research project believes that machine learning (ML) models are a more suitable alternative for the prediction of DC CH4 emissions, as they can capitalise on these inevitable non-linear relationships present within todays large, heterogenous datasets. We tested a traditional SA method, a Linear Mixed Effects (ME) model, against an original ML method, a Random Forest (RF) model, in the prediction of DC CH4 emissions (g/d) across 32 experiments. The RF model was able to outperform the ME model, achieving a Root Mean Square Prediction Error (RMSPE) and Concordance Correlation Coefficient (CCC) of 36.25 CH4 g/d and 0.85 respectively, compared to a ME models' 46.15 g/d and 0.77. The best results, however, were obtained when both the ME and RF models were combined within a Mixed Effects Random Forest (MERF), achieving a RMSPE of 34.92 g/d and CCC of 0.87. Therefore, whilst ML has the potential to outperform traditional SA methods in the prediction of DC CH4 emissions, instead of acting as direct competitors, it is a combination of the two approaches as complimentary teammates, which secures the best results.