



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

Ross, S., Zheng, H., Wang, H., Yan, T., & Shirali, M. (2023). *The Prediction of Dairy Cattle Methane Emissions on Commercial Farms: Machine Learning vs Statistical Analysis*. Abstract from QUB AFBI Alliance Conference, Belfast, United Kingdom.

[Link to publication record in Ulster University Research Portal](#)

Publication Status:

Unpublished: 13/11/2023

General rights

The copyright and moral rights to the output are retained by the output author(s), unless otherwise stated by the document licence.

Unless otherwise stated, users are permitted to download a copy of the output for personal study or non-commercial research and are permitted to freely distribute the URL of the output. They are not permitted to alter, reproduce, distribute or make any commercial use of the output without obtaining the permission of the author(s).

If the document is licenced under Creative Commons, the rights of users of the documents can be found at <https://creativecommons.org/share-your-work/licenses/>.

Take down policy

The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk

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

Stephen R Ross^{1,2}, Huiru J Zheng¹, Haiying Wang¹, Tianhai Yan², 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 today's large, heterogeneous 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 complementary teammates, which secures the best results.