Difford, G., Heringstad, B., Bakke, K., Gjuvsland, A., Gorjanc, G., Jenko, J. (2026) Phenotypic and genetic prediction of methane emissions from milk mid-infrared spectra in Norwegian Red dairy cattle, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 23655: https://doi.org/10.31274/wcgalp.23655 ©2026 Difford, et al. CC BY-NC-ND

Phenotypic and genetic prediction of methane emissions from milk mid-infrared spectra in Norwegian Red dairy cattle

G.F. Difford[1*], B. Heringstad[1,2], K.A. Bakke[2], A.B. Gjuvsland[2], G. Gorjanc[3] and J. Jenko[2]

[1]Norwegian University of Life Sciences (NMBU), Norway; [2]Geno SA, Norway; [3] The Roslin Institute, United Kingdom; [*]Corresponding author: gareth.difford@nmbu.no

Introduction

Milk mid-infrared (MIR) spectroscopy is routinely used in dairy recording and has been proposed as a low-cost indicator for difficult-to-measure traits such as enteric methane (CH_4) emissions (Negussie *et al.*, 2017). If successful, MIR-based indicators could enable large-scale phenotyping for CH_4 at low marginal cost, facilitating genetic evaluation and selection for reduced emissions.

However, evidence for robust and transferable MIR-based CH_4 prediction across cows and herds remains limited, particularly when validation schemes do not fully control for data leakage (Vanlierde *et al.*, 2015; Bresolin and Dórea, 2020; Shadpour *et al.*, 2022). Importantly, MIR spectra are heritable phenotypes (Rovere *et al.*, 2019; Zaalberg *et al.*, 2019) and may share additive genetic covariance with CH_4 production, meaning phenotypic prediction accuracy alone may not reflect their value for genetic improvement. The objective of this study was to evaluate the phenotypic robustness and genetic relevance of MIR-based CH_4 prediction in Norwegian commercial dairy herds.

Materials & Methods

Data sources. Methane emission data were collected using GreenFeed systems (C-Lock Inc., Rapid City, SD, USA) installed in 14 commercial dairy herds in Norway from the spring of 2020 until the fall of 2024. Milk MIR spectra were obtained from routine test-day milk recording conducted approximately every 30 days over the same period. Spectra were recorded using a Bentley FTS/FCM instrument (Bentley Instruments Inc., Chaska, MN, USA) and comprised 899 MIR wavenumbers. Additional cow information, including calving date, parity, and days in milk (DIM), was retrieved from the national dairy herd recording system.

Editing of methane phenotypes and MIR prediction. GreenFeed visit-level data were filtered to visits of at least 2 minutes in duration, and classified as "good data" based on head position and proximity criteria. Visit-level CH_4 estimates were aggregated into daily CH_4 emission rates (g/day) for each cow-day. MIR test-day records were merged with daily CH_4 phenotypes using a ±5-day window around each test day. A minimum of three CH4 phenotypes was retained within each test day window. For each MIR record, daily CH_4 values within the window were averaged to form a single matched phenotype, and the associated standard error (SE) of the mean was retained as a measure of phenotype reliability. The final dataset consisted of 10,218 matched MIR–CH_4 observations, derived from 13,602 MIR records and 693,830 daily CH_4 records, with approximately 75% overlap between MIR and CH_4 data. The final merged dataset comprised 1,506 Norwegian Red dairy cows.

MIR spectra were preprocessed using standard normal variate correction followed by Savitzky–Golay derivatives. Partial least squares (PLS) regression was used to predict daily CH_4 emissions from MIR spectra, either observations were weighted by the inverse of the CH_4 phenotype SE to account for differences in sampling intensity and within-window variability,

or as unweighted averages. The number of latent components was selected using the 1-SE rule based on cross-validated RMSE. Prediction performance was evaluated using RMSE, R^2, and MAPE under two schemes: (i) five-fold cross-validation with cow-level blocking (5CV_COW), and (ii) leave-one-herd-out cross-validation (LOHO_CV).

Genetic analyses. Genetic parameters were estimated for CH_4 production and for each of the 899 MIR wavenumbers using pedigree-based single-trait animal models. Pedigrees were traced back nine generations, resulting in 10,097 animals in the relationship matrix. Animal models included fixed effects of herd, parity, and fixed regressions on DIM and Wilmink term, and random effects including additive genetic effects, permanent environmental effects, and herd-test-day effects. Heritabilities of MIR wavenumbers and CH_4 production were estimated, and bivariate analyses were used to calculate genetic correlations between individual MIR wavenumbers and CH_4 production.

Results

Prediction accuracy declined substantially under LOHO_CV compared with 5CV_COW (Table 1). Weighting predictions based on the inverse of the standard error of the mean CH4 production during the MIR window had no limited or no benefit as compared to the unweighted mean phenotypes.

Table 1. Predictive performance of PLs regression models for estimating CH4 production across different crass-validation schemes.

Cross-validation scheme[1]	Component #[2]	R^2CV	RMSECV	MAPECV
5CV_COW unweighted	78	0.31	71.9g	14.7%
5CV_COW weighted	78	0.31	71.9g	14.7%
LOHO CV unweighted	10-26	0.05	91.1g	19.0%
LOHO CV weighted	10-34	0.04	89.7g	18.5%

[1] Cross-validation schemes either five-fold with cow-level blocking (5CV_COW) or leave-on-herd-out (LOHO_CV), weighting is done by the inverse of the standard error of mean CH4 production.

[2] Component # = number of components chosen in the folds, R^2 CV = coefficient of determination for cross validation prediction accuracy, RMSECV = Root mean square error of cross validation prediction accuracy, MAPECV = mean absolute percentage error of cross validation prediction accuracy.

Heritability estimates for individual MIR wavenumbers ranged from 0.00 to 0.30 with standard errors ranging from 0.02 to 0.08, indicating that substantial parts of the MIR spectrum contain additive genetic variation. Genetic correlations between MIR wavenumbers and CH_4 production ranged from −0.41 to 0.50 with standard errors in the range of 0.13 to 0.23, for MIR spectra with significant heritability estimates, revealing spectral regions with moderate genetic association with methane production.

Discussion

Several studies have explored MIR-based prediction of methane under experimental or semi-commercial conditions, often using relatively small datasets or limited validation schemes (Bresolin and Dórea, 2020; Shadpour *et al.*, 2022; McParland *et al.*, 2024). By contrast, the present study uses extensive GreenFeed and milk recording data across multiple commercial herds in Norway. With increased data recording, it becomes possible to assess more rigorously both phenotypic prediction performance and genetic relevance.

CH_4 production and MIR spectral wavenumbers are known to be heritable and repeatable traits, with phenotypic variation arising from a combination of additive genetic effects, permanent environmental effects, and short-term environmental and temporal influences (Wang,

Hulzebosch and Bovenhuis, 2016; Rovere *et al.*, 2019; Breukelen *et al.*, 2023). As a result, when prediction models are trained and evaluated purely on phenotypic records, it is often unclear which components of variation are driving apparent model performance. This is particularly relevant in datasets with repeated records per cow, where failure to adequately separate cows in the training and validation sets can lead to substantial inflation of prediction accuracy.

Earlier studies based on single records per cow and relatively small sample sizes (approximately 60–500 cows) have reported relatively high coefficients of determination (R^2 ranging from 0.57 to 0.79) (Dehareng *et al.*, 2012; Vanlierde *et al.*, 2015, 2018; Shadpour *et al.*, 2022). However, such estimates are difficult to compare directly with results from larger datasets containing repeated measurements, where information leakage across records must be carefully controlled.

We found that when 5CV_COW was performed, prediction accuracy was moderate ($R^2 = 0.31$). This result is more conservative than many previously reported estimates. Importantly, even with cow-level blocking, cows from the same herds may still appear in both training and validation sets, allowing herd-specific effects related to diet, management, calibration, or recording environment to contribute to model performance. When we therefore applied a leave-one-herd-out validation scheme, prediction accuracy declined substantially, with R^2 dropping to approximately 0.06 and corresponding increases in RMSE and MAPE (91 g/day and 19%, respectively). This between-herd validation scenario most closely reflects real-world deployment of MIR-based prediction equations in herds where no reference CH_4 data are available. The marked reduction in performance under this scheme demonstrates limited transferability of MIR-based phenotypic prediction of CH_4 emissions across herds and highlights the importance of validation design when evaluating such models.

Despite limited phenotypic transferability, genetic analyses revealed that MIR spectra capture heritable variation, consistent with previous findings in larger MIR studies. Heritability estimates for individual MIR wavenumbers were moderate, and genetic correlations between MIR wavenumber breeding values and CH_4 breeding values ranged from −0.41 to 0.50. Although the associated standard errors were relatively large (up to approximately 0.20), these results indicate that specific regions of the MIR spectrum share additive genetic covariance with CH_4 production. Consequently, the genetic signal contained in MIR spectra may be substantially stronger than suggested by phenotypic prediction accuracy alone.

These findings support the view that MIR spectra may be better suited as genetically informative indicator traits rather than as direct phenotypic predictors of methane emissions for management purposes. For breeding applications, this distinction is critical, as selection decisions are based on estimated breeding values rather than on raw phenotypic predictions. The presence of moderate genetic correlations suggests potential for exploiting MIR-derived information in genetic evaluation of methane emissions, particularly as reference populations and repeated recording continue to expand.

Further research is required to confirm the stability of these genetic relationships, reduce uncertainty in correlation estimates, and evaluate correlated responses with production traits.

Conclusions

Phenotypic prediction of methane emissions from MIR spectra was only moderately robust across herds, suggesting limited transferability of prediction equations under commercial conditions. However, MIR spectra captured heritable variation and exhibited moderate genetic correlations with methane production, demonstrating their potential as genetically informative indicator traits. These results emphasize the need to validate MIR-based methane predictions genetically before implementation in dairy cattle breeding programs.

Acknowledgements
The authors would like to acknowledge the support of two Research Council of Norway projects 309451 and 346741.

References
Bresolin T. and Dórea J. R. R. (2020) Front. Genet. 11:1–20. https://doi: 10.3389/fgene.2020.00923
Breukelen A. E. Van *et al*. (2023) J. Dairy Sci. 106:4121–4132. https://doi: 10.3168/jds.2022-22735
Dehareng F., *et al*. (2012) Animal, 6:1694–1701. https://doi: 10.1017/S1751731112000456
Negussie, E., *et al*. (2017) J. Dairy Sci. 100:2433–2453. https://doi: 10.3168/jds.2016-12030
Rovere, G., *et al.* (2019) J. Dairy Sci. 102:1354–1363. https://doi: 10.3168/jds.2018-15109
Shadpour, S., *et al*. (2022) J. Dairy Sci. 105:8272–8285. https://doi: 10.3168/jds.2021-21176
Vanlierde, A., *et al.* (2015) J. Dairy Sci. 98:5740–5747. https://doi: 10.3168/jds.2014-8436
Vanlierde, A., *et al.* (2018) J. Dairy Sci. 101:7618–7624. https://doi: 10.3168/jds.2018-14472
Wang Q., Hulzebosch A. and Bovenhuis H. (2016) J. Dairy Sci. 99:6793–6803. https://doi: 10.3168/jds.2015-10488
Zaalberg R.M. *et al*. (2019) J. Dairy Sci.102:503–510. https://doi: 10.3168/jds.2018-14464