Prediction of dry matter intake and gross feed efficiency using milk production and live weight in first-parity Holstein cows

Article *in* Tropical Animal Health and Production · September 2022

DOI: 10.1007/s11250-022-03275-8

3 authors:

Matome Madilindi
Agricultural Research Council, South Africa

Cuthbert Banga
Botswana University of Agriculture and Natural Resources

Oliver Tendayi Zishiri
University of KwaZulu-Natal

Tropical Animal Health and Production (2022) 54:278
https://doi.org/10.1007/s11250-022-03275-8

REGULAR ARTICLES

Prediction of dry matter intake and gross feed efficiency using milk production and live weight in first-parity Holstein cows

Matome A. Madilindi[1,2] · Cuthbert B. Banga[3,4,5] · Oliver T. Zishiri[1]

Received: 17 May 2022 / Accepted: 31 August 2022

Abstract

Direct measurement of dry matter intake (DMI) presents a major challenge in estimating gross feed efficiency (GFE) in dairy cattle. This challenge can, however, be resolved through the prediction of DMI and GFE from easy-to-measure traits such as milk production (i.e. milk yield, energy-corrected milk (ECM), butterfat, protein, lactose) and live weight (LW). The main objective of this study was, therefore, to investigate the feasibility of predicting dry matter intake and gross feed efficiency for first-parity Holstein cows using milk production traits and LW. Data comprised of 30 daily measurements of DMI and milk production traits, and 25 daily LW records of a group of 100 first-parity Holstein cows, fed a total mixed ration. Gross feed efficiency was calculated as kg ECM divided by kg DMI. The initial step was to estimate correlations of milk production traits and LW with DMI and GFE, to identify the best potential predictors of DMI and GFE. Subsequently, a forward stepwise regression analysis was used to develop models to predict DMI and GFE from LW and milk production traits, followed by within-herd validations. Means for DMI, butterfat yield (BFY) and LW were 21.91 ± 2.77 kg/day, 0.95 ± 0.14 kg/day and 572 ± 15.58 kg/day, respectively. Mean GFE was 1.32 ± 0.22. Dry matter intake had positive correlations with milk yield (MY) ($r = 0.32$, $p < 0.001$) and LW ($r = 0.76$, $p < 0.0001$) and an antagonistic association with butterfat percent (BFP) ($r = -0.55$, $p < 0.001$). On the other hand, GFE was positively associated with MY ($r = 0.36$, $p < 0.001$), BFP ($r = 0.53$, $p < 0.001$) and BFY ($r = 0.83$, $p < 0.0001$), and negatively correlated with LW ($r = -0.23$, $p > 0.05$). Dry matter intake was predicted reliably by a model comprising of only LW and MY ($R^2 = 0.79$; root mean squared error (RMSE) = 1.05 kg/day). A model that included BFY, MY and LW had the highest ability to predict GFE ($R^2 = 0.98$; RMSE = 0.05). Live weight and BFY were the main predictor traits for DMI and GFE, respectively. The best models for predicting DMI and GFE were as follows: DMI (kg/day) $= -54.21 - 0.192 \times$ MY (kg/day) $+ 0.146 \times$ LW (kg/day) and GFE (kg/day) $= 4.120 + 0.024 \times$ MY (kg/day) $+ 1.000 \times$ BFY (kg/day) $- 0.008 \times$ LW (kg/day). Thus, daily DMI (kg/day) and GFE can be reliably predicted from LW and milk production traits using these developed models in first-parity Holstein cows. This presents a big promise to generate large quantities of data of individual cow DMI and GFE, which can be used to implement genetic improvement of feed efficiency.

Keywords Correlation · Easy-to-measure traits · Feed intake · Feed efficiency · Stepwise regression

✉ Matome A. Madilindi
matomemadilindi@gmail.com

1 Discipline of Genetics, College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Private Bag X54001, Durban 4000, South Africa

2 ARC-Animal Production, Private Bag X2, Irene 0062, South Africa

3 Department of Animal Sciences, Faculty of Science, Tshwane University of Technology, Private Bag X680, Pretoria 0001, South Africa

4 Department of Agriculture and Animal Health, University of South Africa, Private Bag X6, Florida 1710, South Africa

5 Department of Animal Sciences, Faculty of Animal and Veterinary Sciences, Botswana University of Agriculture and Natural Resources, Private Bag 0027, Gaborone, Botswana

Introduction

Feed efficiency (FE) has large implications on animal production profitability and environmental sustainability; hence, it has become a common standard for monitoring the economic viability of milk production and environmental footprint (Vallimont et al. 2011; Pryce et al. 2015; Tempelman et al. 2015). Thus, extensive efforts are being made worldwide to include FE in dairy cattle breeding objectives. Such efforts include genome-wide association analyses for FE using international dairy FE research consortium data, aiming to implement marker-assisted selection (e.g. de Haas et al. 2015; Lu et al. 2018).

All FE traits such as gross feed efficiency and residual feed intake require the measurement of dry matter intake (DMI). Dry matter intake is a key variable in the calculation of an individual cow's feed efficiency. However, direct measurement of individual animal DMI remains a major challenge in assessing dairy feed efficiency. Measurement of individual DMI is generally difficult and expensive, especially for animals on a pasture-based feeding system. This has been a major factor hindering the inclusion of FE traits into selection objectives (Shetty et al. 2017; Wallén et al. 2018). It is generally difficult to obtain large numbers of individual animal records on DMI that are required to estimate accurate breeding values for FE (McParland et al. 2014).

Milk production traits and live weight (LW), which are cheap and easy to measure, could potentially be used as reliable predictors of DMI, since they can allow proper accounting for the amount of feed required for production and maintenance (Veerkamp 1998; Liinamo et al. 2012; VandeHaar et al. 2016). This can be achieved through developing models to predict DMI and gross feed efficiency (GFE, kg energy-corrected milk/kg DMI). Limited efforts have, however, been made to explore the possibility of developing such prediction models in dairy cattle, despite indications that DMI can be predicted reliably from these easy-to-measure traits (Madilindi et al. 2022). The development of reliable GFE prediction models can also limit the burden of generating DMI data prior to calculating GFE as its major determinant.

Models for predicting GFE can be developed through stepwise regression, by identifying traits that significantly account for variation in GFE. This technique selects independent/explanatory variables for multiple regression models based on their statistical significance (Smith 2018). Although it has often been criticised for the misapplication of single-step statistical tests to a multi-step procedure, stepwise regression has become popular in cases with many explanatory variables (Ngo 2012). The method efficiently chooses a relatively small number of explanatory variables from a vast array of possibilities. It is usually assumed that the more the number of possible predictors, the more useful is the stepwise regression (Smith 2018). Application of this technique presents a good opportunity to develop accurate models for predicting DMI and GFE using traits such as milk production (i.e. milk yield, energy-corrected milk, butterfat, protein, lactose) and live weight. This was demonstrated in a recent study that was carried out to predict DMI in multiparous Holstein cows (Liang et al. 2021).

The current study was carried out to investigate the correlations of milk production traits and LW with DMI and GFE, and subsequently develop the most suitable prediction models for daily DMI and GFE using milk production traits and live weight, in first-parity Holstein cows.

Materials and methods

Animal management and data collection

The study was conducted at Limpopo Dairies (PTY) LTD farm, Louis Trichardt, South Africa, from 2 November to 11 December 2020. The availability of a feed intake measurement system at this farm mainly motivated its choice for this study. Other dairy farms across South Africa hardly measure cow feed intake due to lack of weighing facilities. The Limpopo dairy farm had a herd with over 800 lactating Holstein cows. Measurements were recorded on a group comprising 100 first-parity cows. Animals were kept under an intensive production system, with conventional cubicle housing in free stall barns. During the experimental period, temperature in the area varied from 18 to 36 °C, with an average of 22 °C per day. Cows were fed a total mixed ration (TMR) ad libitum thrice per day (07h30 in the morning, 13h30 in the afternoon and 20h30 in the evening). Feed was disbursed by a feeding truck with an automated feeding scale and animals had free access to water. Residual feed was weighed and discarded daily before the morning feeding. Daily dry matter intake (kg/day) of the group was automatically calculated as feed provided minus feed refusal. All animals were milked thrice a day at 04h00, 13h00 and 20h00. Individual cow milk yield (kg/day) was automatically recorded using the DeLaval ALPRO System (Tumba, Sweden) and added to the group's milk production (MY, kg/day). Individual cow live weight was measured weekly using the DeLaval ALPRO System (Tumba, Sweden) weighing scale, after the morning milking session, before the morning feeding.

Sample collection

Daily milk samples were collected for the first-parity group during the afternoon milking session. Samples were collected from Monday to Friday, for a period of 6 weeks. All milk samples were conserved with bronopol (2-bromo-2-nitropropane-1,3-diol) and then stored at room temperature for <7 days. They were then sent to ARC-Elsenburg Analytical Services, Stellenbosch, South Africa, for determining the percentages of butterfat (BFP), protein (PROP) and lactose (LACP), using a CombiFoss™ F+(Foss, Hillerød, Denmark), according to manufacturer's protocols.

Traits computation

To standardise milk yield and its components (i.e. butterfat, protein and lactose), daily energy-corrected milk (ECM) yield was calculated according to the following formula (Kirchgeßner, 1997):

$$\text{ECM (kg/day)} = \text{milk yield (kg)} \times \frac{(0.39 \text{ x butterfat\%} + 0.24 \text{ x protein\%} + 0.17 \text{ x lactose\%})}{3.17}$$

Gross feed efficiency (GFE) was calculated as ECM (kg/day) divided by DMI (kg/day). Daily LW (kg/day) was calculated using linear interpolation between two weekly weight measurements. There were 25 measurements of LW and 30 records of daily MY, ECM, BFP, PROP, LACP, butterfat yield (BFY), protein yield (PROY), lactose yield (LACY), DMI and GFE, which were used for correlation analysis. Twenty-five (25) daily records of milk production traits, DMI and GFE as well as measurements of LW remained after deleting milk production traits and DMI and GFE records without corresponding LW measurements. Of these 25 records, 60% were randomly selected for a training data set that was used to develop the models, and the remaining 40% were used as a within-herd validation data set.

Data analysis

Correlation coefficients

Pearson correlation coefficients (*r*) amongst DMI, GFE, LW, MY, ECM, BFP, BFY, PROP, PROY, LACP and LACY were calculated using the PROC CORR procedure of the Statistical Analysis System (SAS) (version 9.4, SAS Institute, Cary, NC, USA).

Development of prediction models

A stepwise regression analysis was performed to select independent traits that met the 0.1500 significance level for entry into the model as well as those accounting for the highest coefficient of determination (R^2), in explaining variation in DMI and GFE. The analysis was conducted using the PROC REG procedures of SAS (version 9.4, SAS Institute, Cary, NC, USA), and the prediction models for DMI and GFE were fitted as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \varepsilon ,$$

where Y is the dependent trait [daily DMI (kg/day) or GFE], β_0 is the regression intercept, $\beta_1 - \beta_9$ are the regression coefficients, $X_1 - X_9$ are the independent traits (MY (kg/day), ECM (kg/day), BFP (%), PROP (%), LACP (%), BFY (kg/day), PROY (kg/day), LACY (kg/day) and LW (kg/day)), and ε is the residual error term.

Within-herd validation

A regression analysis was further performed between actual and predicted DMI and GFE records based on the validation dataset, using the PROC REG procedure of SAS (version 9.4, SAS Institute, Cary, NC, USA). The following fitting statistics were used to assess the robustness and accuracy of the developed prediction models: the R^2 between actual and predicted DMI and GFE values and the root mean squared error (RMSE) from plotting predicted versus actual DMI and GFE.

Results

Summary statistics

Descriptive statistics for DMI, GFE, milk production and LW are presented in Table 1. In this study, each cow consumed an average of 21.91 ± 2.58 kg/day dry matter of a total mixed ration, to produce mean ECM yield of 28.46 ± 3.91 kg/day, which resulted in a GFE of 1.32 ± 0.22 per animal per day. Animals produced MY, BFP and BFY that ranged from 21.78 to 44.73 kg/day, 1.36 to 3.36% per day and 0.51 to 1.27 kg/day, respectively. Each cow weighed an average weight of 572 ± 15.58 kg.

Table 1 Descriptive statistics for dry matter intake, gross feed efficiency, milk production and live weight of first-parity Holstein cows

Traits	*N*	Mean	Minimum	Maximum	SD
DMI (kg/day)	30	21.91	17.09	25.71	2.58
GFE	30	1.32	0.72	1.78	0.22
MY (kg/day)	30	34.26	21.78	44.73	4.29
ECM (kg/day)	30	28.46	17.94	37.92	3.91
BFP (%)	30	2.81	1.36	3.17	0.35
PROP (%)	30	2.96	1.91	3.39	0.23
LACP (%)	30	4.91	2.96	5.06	0.37
BFY (kg/day)	30	0.95	0.51	1.27	0.14
PROY (kg/day)	30	1.01	0.64	1.36	0.16
LACY (kg/day)	30	1.68	1.05	2.22	0.24
LW (kg/day)	25	572	545	595	15.58

DMI dry matter intake, *BFP* butterfat percent, *BFY* butterfat yield, *MY* milk yield, *ECM* energy-corrected milk, *GFE* gross feed efficiency, *PROP* protein percent, *PROY* protein yield, *LACP* lactose percent, *LACY* lactose yield, *LW* live weight, *N* number of observations, *SD* standard deviation

Table 2 Pearson's correlation coefficients of milk production traits and live weight with dry matter intake and gross feed efficiency, in first-parity Holstein cows

Traits	DMI	GFE
MY	0.32*	0.36*
ECM	0.07 ns	0.70***
BFP	−0.55**	0.53***
PROP	0.13 ns	0.35*
LACP	−0.19 ns	0.53**
BFY	−0.21 ns	0.83***
PROY	0.31 ns	0.49**
LACY	0.18 ns	0.61**
LW	0.76***	−0.23 ns

DMI dry matter intake (kg/day), *BFP* butterfat percent (%), *BFY* butterfat yield (kg/day), *MY* milk yield (kg/day), *ECM* energy-corrected milk (kg/day), *GFE* gross feed efficiency, *PROP* protein percent (%), *PROY* protein yield (kg/day), *LACP* lactose percent (%), *LACY* lactose yield (kg/day), *LW* live weight (kg/day), *ns* not significant at $p > 0.05$, *** significant at $p < 0.0001$, ** significant at $p < 0.01$, * significant at $p < 0.05$

Correlation analyses

Pearson's correlation coefficients of milk production traits and LW with DMI and GFE are presented in Table 2. DMI was positively associated with MY (0.32, $p < 0.05$) and LW (0.76, $p < 0.0001$), and negatively correlated with BFP (−0.55, $p < 0.01$). On the other hand, GFE had a desirable relationship with MY (0.36, $p < 0.05$), BFP (0.53, $p < 0.0001$) and BFY (0.83, $p < 0.0001$), but negatively correlated with LW (−0.23, $p > 0.05$).

Developed prediction models

The models that were developed for predicting DMI and GFE, using stepwise regression, are presented in Table 3. All the models developed were significant ($p < 0.0001$). Live weight (kg/day) was the best predictor of DMI (kg/day) ($p < 0.0001$), followed by MY (kg/day) ($p < 0.05$). The coefficient of determination significantly increased from 0.66 to 0.79 when MY was added to the DMI prediction model based on LW only. All the other traits did not make any significant contributions ($p > 0.05$); hence, the best model for predicting DMI included only LW and MY, with a corresponding RMSE of 1.05 kg/day. Butterfat yield was the best predictor of GFE ($p < 0.0001$), followed by LW ($p < 0.05$) and MY ($p < 0.001$). A model based on BFY only accounted for 87% of the variation in GFE (i.e. $R^2 = 0.87$). Addition of LW into this model increased R^2 to 0.91, and further inclusion of MY resulted in the best model ($R^2 = 0.98$), with a low corresponding RMSE (0.05).

Within-herd validation of the developed prediction models

Figure 1 is a graphical presentation of the regression of DMI predicted from the developed model (PDMI) on actual DMI, for the within-herd validation data set. There was a relatively moderate positive relationship ($R^2 = 0.49$) between PDMI and actual DMI, with a low root mean square error (RMSE) of 1.46 kg/day. The regression of predicted GFE on actual GFE is presented in Fig. 2. There was a strong relationship between predicted and actual GFE ($R^2 = 0.64$) and a low corresponding RMSE (0.13).

Discussion

Reliable data on individual cow DMI and GFE is key to achieving improvement in the efficiency of feed utilisation in dairy cattle. Availability of such data on a large scale would enable the estimation of accurate breeding values for feed efficiency traits, thereby facilitating their inclusion in the breeding objectives (McParland et al. 2014). Generating large amounts of DMI records through direct measurement is, however, challenging. In the current study, it was hypothesised that milk production traits and live weight, which are easy and cheap to measure, can be used as reliable predictors of DMI and GFE.

Table 3 Prediction models for DMI and GFE developed by stepwise regression

Explanatory traits	Model	R^2
DMI		
LW	$DMI = -38.09 + 0.106 \text{xLW}$	0.66
LW, MY	$DMI = -54.21 - 0.192 \text{xMY} + 0.146 \text{xLW}$	0.79
GFE		
BFY	$GFE = 0.018 + 1.335 \text{xBFY}$	0.87
BFY, LW	$GFE = 1.881 + 1.344 \text{xBFY} - 0.003 \text{xLW}$	0.91
MY, BFY, LW	$GFE = 4.120 + 0.024 \text{xMY} + 1.000 \text{xBFY} - 0.008 \text{xLW}$	0.98

DMI dry matter intake (kg/day), *BFY* butterfat yield (kg/day), *MY* milk yield (kg/day), *GFE* gross feed efficiency, *LW* live weight (kg/day). All models were significant at $p < 0.0001$; R^2 coefficient of determination

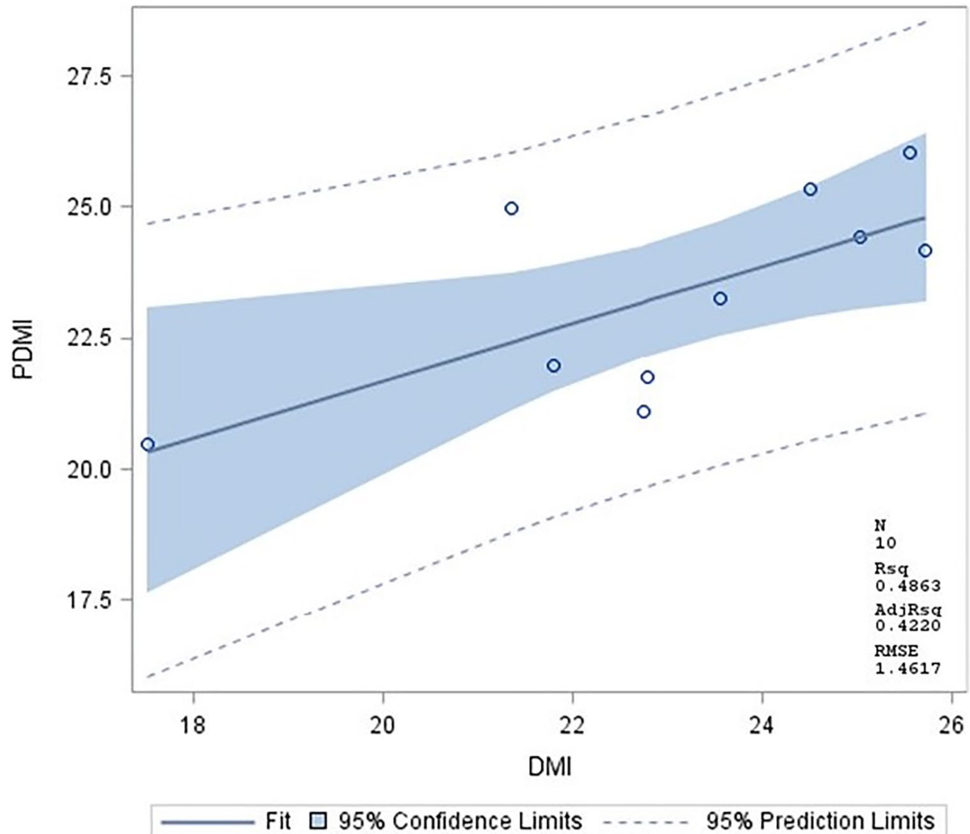

Fig. 1 Regression of PDMI on actual DMI in first-parity Holstein cows

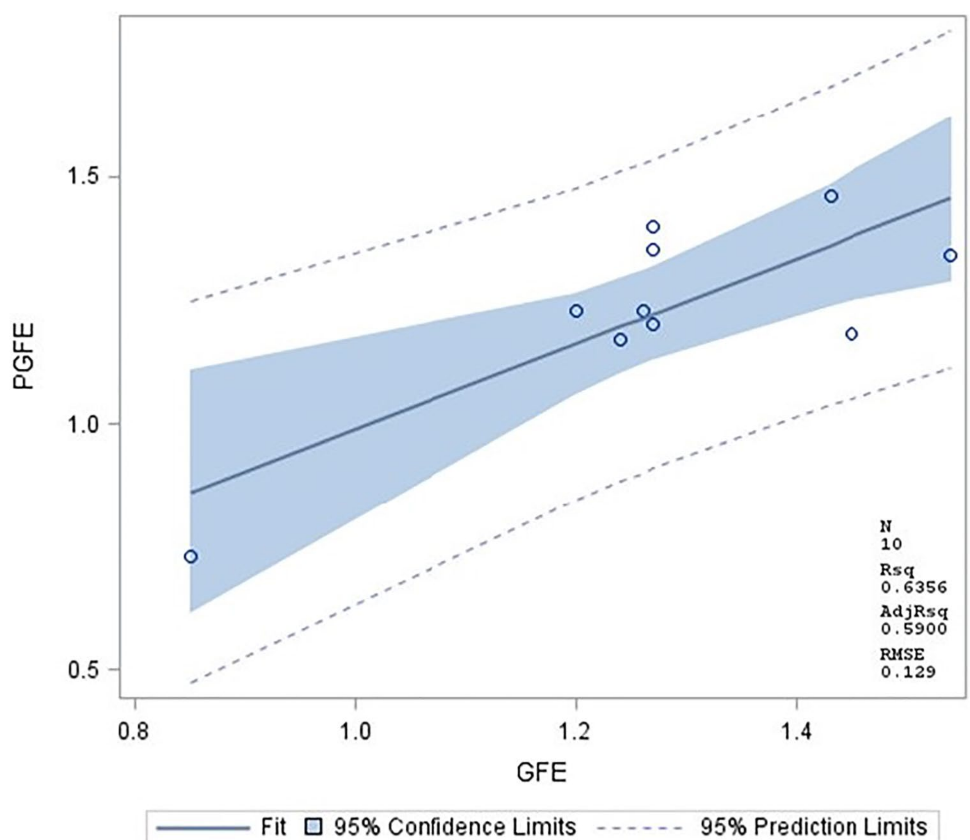

Fig. 2 Regression of PGFE on actual GFE in first-parity Holstein cows

Performance statistics for milk production, live weight, DMI and GFE

Each cow in the current study consumed an average of 21.91 kg/day dry matter of a total mixed ration, to produce mean energy-corrected milk and milk yield of 28.46 kg/day and 34.26 kg/day, respectively. This resulted in a gross feed efficiency of 1.32 per animal per day, which is within the expected range for Holstein cows in first lactation (Heinrichs and Ishler 2016). There is, however, scarcity of information on DMI from studies in a sub-tropical environment with which to compare the results of the current study. Our mean for DMI was, however, lower than that observed in Canadian first-parity Holstein cows (Beard 2018). This variation could be attributed to various factors including the different environmental conditions where animals were reared, the level of production and genetic merit of the cows, which may highly influence feed intake. A Holstein cow in the heat of the Limpopo Valley, where the current study was carried out, is expected to have much lower intake compared to one in freezing Canada. Maintenance requirements would also be higher for the cow in Limpopo, as it would need more energy to cool its body. Generally, the means for milk production, live weight, DMI and GFE found in this study are within the range for Holstein animals in first-parity (ICAR 2012; Poncheki et al. 2015; Heinrichs and Ishler 2016; Kratenmacher et al. 2019).

Correlations between milk production traits, live weight, DMI and GFE

A preliminary step in the development of prediction models was to quantify the phenotypic association of milk production traits and live weight with DMI and GFE, in order to determine those traits that could be the best candidates as predictors for DMI and GFE. Milk yield was moderately and positively associated with DMI (0.32), meaning that increased feed intake led to higher milk production, which is in agreement with findings from other studies (e.g. Ben Meir et al. 2018; Zhang et al. 2020; Liang et al. 2021). This is attributable to more nutrients being available for milk production, after meeting requirements for other physiological functions such as growth and maintenance (Erickson and Kalscheur, 2020). The association between MY and DMI demonstrates that MY could be a useful predictor for DMI. Accordingly, MY has been widely considered alone or together with other traits (e.g. butterfat, protein, lactose contents, LW, mid-infrared spectra of milk, parity and stage of lactation) in most prediction models for DMI in dairy cows (NRC 2001; Lindgren et al. 2001; Shetty et al. 2017; Lahart et al. 2019; Liang et al. 2021).

Butterfat percent had a moderate antagonistic association with DMI (−0.55), although a previous study reported a poor positive association between these traits in first-parity Holstein cows in Belgium (Zhang et al. 2020). This relationship implies that cows consuming more dry matter produced milk with less butterfat. The biological basis of this relationship is not clear. In extreme cases, too much feed in the cow rumen, particularly rations low in fibre, may cause digestibility problems, which may result in low milk butterfat production (Heinrichs and Jones 2016). This association between butterfat percent and DMI presented butterfat percent as a candidate predictor in developing our models for predicting DMI. Other studies have also demonstrated the importance of BFP in prediction models for DMI in dairy cows (e.g. Lahart et al. 2019).

Live weight had a strong positive association with DMI (0.76), in concurrence with findings from other studies on first-parity Holstein cows (e.g. Zhang et al. 2020). This relationship has, however, been reported to be moderate in multiparous cows (Zhang et al. 2020; Liang et al. 2021). Thus, heavier animals consume more feed due to their higher requirements for body maintenance, and this has been documented in many studies (Searle et al. 1982; Vallimont et al. 2011; Guinguina et al. 2019). Searle et al. (1982) indicated that LW is one of the factors most closely related to net energy for maintenance. The strong association between LW and DMI observed in the current study points to LW as a major candidate for predicting DMI. Consequently, LW has been widely considered in most prediction models of DMI in dairy cows (NRC 2001; Lahart et al. 2019; Martin et al. 2021; Liang et al. 2021).

The strong positive relationship observed between BFY and GFE (0.83) implies that cows producing higher quantities of butterfat were more efficient at converting feed into milk. There is, however, scarcity of information on the relationship between these two traits in the literature, with which to compare these findings. Thus, BFY also appeared as a promising predictor of GFE in the development of our prediction models.

A moderate positive association (0.36) was observed between MY and GFE. Comparable results have been reported in Holstein cows (Ben Meir et al. 2018); however, a previous study found a stronger relationship between these two traits (Spurlock et al. 2012). In addition, a long-term study may be beneficial to achieve a strong relationship. This association implies that as MY increases, corresponding gains in feed efficiency are achieved. Milk yield, therefore, came out as another strong potential predictor of GFE.

There was a low antagonistic relationship (−0.23) between LW and GFE, which was however not significant. This negative association is a well-documented phenomenon, and is attributable to the fact that larger cows demand more nutrients for body maintenance, resulting in less feed being available for milk production (Linn et al., 2009; Vallimont et al., 2011; Ben Meir et al., 2018; Guinguina et al., 2019). Thus, LW could contribute towards the prediction of

GFE. A recent study by Guinguina et al. (2019) found the inclusion of LW to be useful in models for predicting GFE.

Developed prediction models for DMI and GFE

Reliable prediction of DMI and/or GFE from easy-to-measure traits could assist in generating large quantities of data for the estimation of accurate breeding values. Such predictions can be obtained from basal linear models, with milk production and live weight as independent variables (VandeHaar et al. 2016). These easy-to-measure traits are known to greatly influence feed efficiency, as they are important drivers of feed intake (VandeHaar et al. 2016). It is unclear which trait, between DMI and GFE, can be predicted more reliably than the other from milk production and live weight. In the current study, stepwise regression analyses were performed to develop models for predicting DMI and GFE using milk production traits and live weight, as independent variables, in first-parity Holstein cows.

Dry matter intake

Live weight (kg/day) was the best predictor of DMI (kg/day), followed by MY (kg/day) in the present study. This confirms the correlation results of our preliminary analysis, which found these two traits to be good potential predictors of DMI. Combining MY and LW, we have achieved better accuracy of prediction, compared to a model with either of the traits only. Previous studies have similarly found MY and LW to be highly correlated with DMI and, thus, included them in prediction models for DMI (Holter et al. 1997; NRC 2001; Lahart et al. 2019; Liang et al. 2021; Martin et al. 2021). These two variables are the major determinants of the cow's total nutrient requirements; hence, they have large influence on DMI. Our best model for predicting DMI had a greater prediction power (R^2 of 0.79) prior validations, than the one recently developed from multiparous Holstein data in China (R^2 of 0.46) (Liang et al. 2021). A high prediction accuracy (R^2 of 0.71) for DMI was also obtained from a model including MY, LW and mid-infrared (MIR) spectra data, in a multiparous American Holstein cattle population (Dórea et al. 2018). Differences in the methods used to measure DMI, model development approaches, milk production traits considered, parity and stage of lactation may be responsible for the disparity in accuracy of prediction between studies. Based on the magnitude of the R^2 value, our model may be considered sufficiently reliable for application to obtain large quantities of DMI data at low cost. High accuracy of phenotypes may be dispensed with as a requirement for obtaining accurate estimated breeding values (EBVs) if large quantities of phenotypic records are obtainable (Calus et al. 2013; McParland and Berry 2016). Given the fact that a significantly high number of the South African Holstein cattle population is under national milk recording and improvement scheme (NMRIS 2020), the models developed in the current study provide an opportunity to generate large quantities of DMI data, which can be utilised to produce high accuracy EBVs for feed efficiency. There is, however, a need to determine the genetic variability of the predicted DMI, so as to determine the potential to improve it through selection.

Gross feed efficiency

In the present study, butterfat yield (kg/day) was the best predictor of GFE, followed by LW (kg/day) and MY (kg/day). This supports the results of our preliminary analysis (i.e. correlations), which established these three traits to be good potential predictors of GFE. A model including BFY, MY and LW achieved the best accuracy of prediction, compared to one with only one of these traits. These traits have also been found to be associated with GFE and, therefore, good predictor traits for GFE in several other studies (Linn et al. 2009; Vallimont et al. 2011; Ben Meir et al. 2018; Guinguina et al. 2019). Our best model for predicting GFE had an exceptionally stronger prediction power ($R^2 = 0.98$) compared to the previously developed model from multiparous Holstein data ($R^2 = 0.76$) (Guinguina et al. 2019). Beard (2018) developed models with much lower prediction ability ($R^2 = 0.45$) using primiparous Canadian Holstein data, including milk yield, milk components and live weight only. The disparity in prediction power amongst the different studies may be attributed to variation in methods used to measure DMI, parities considered and approaches used to develop the models. The stage of lactation at which DMI is measured may also influence accuracy of prediction (Lahart et al. 2019). There are limited studies on primiparous cows in the literature. Given its high prediction power (i.e. high coefficient of determination), our best model could be applied to generate large quantities of reliable GFE data at low cost, which can be used to estimate accurate EBVs for GFE. It is, however, necessary to first determine the genetic variability of this predicted trait, in order to determine the extent to which it is under genetic control.

Within-herd validation of DMI and GFE prediction models

An assessment of the robustness and accuracy of the models developed for predicting DMI and GFE from milk production and live weight was carried through within-herd validation of predicted data. There were no external data available to carry out such an analysis.

Validation of DMI prediction model

Validation of the model for predicting DMI that included MY and LW only yielded a moderate R^2 and low RMSE, indicating modest robustness and accuracy. This prediction power was, however, relatively higher compared to values observed by Liang et al. (2021) in a study on multiparous Holstein cattle. Interestingly though, the model developed by Liang et al. (2021) included dry matter intake of the first 2 h after feeding, in addition to MY and LW. On the other hand, Lahart et al. (2019) developed models that predicted DMI with milk production traits (MY, fat percent and protein percent) and LW as well as stage of lactation and parity of grazing Irish Holstein–Friesian and Jersey cross-bred cattle. Adding milk MIR spectra data to these prediction models resulted in a slightly improved prediction power (Lahart et al. 2019). Similarly, Dórea et al. (2018) predicted DMI from milk yield, body weight and days in milk of multiparous American Holstein cattle with a strong prediction power, which improved with the addition of milk MIR spectra data. Models developed from the different studies are bound to vary in their ability to predict DMI largely due to differences in the predictor traits included, and methods used to measure DMI and validate predicted DMI, as well as factors such as breed, parity, stage of lactation and production system (Dórea et al. 2018; Lahart et al. 2019; Liang et al. 2021). Utilisation of additional easy-to-measure traits such as milk MIR spectra and days in milk may be useful in improving the accuracy of predicting DMI, as indicated by some previous studies (Shetty et al. 2017; Dórea et al. 2018; Wallén et al. 2018; Lahart et al. 2019). It might also be worthwhile to explore other prediction and validation methods such as the partial least squares approach and artificial neural networks (Felipe et al. 2015; Dórea et al. 2018).

Validation of GFE prediction model

Studies on the prediction of GFE are generally scarce in the literature. The model that we developed for predicting GFE from BFY, LW and MY had a reasonably strong prediction ability, as indicated by a high R^2 and small RMSE. This was confirmed by the results of within-herd validation, despite a slight inconsistence in the relative magnitude of the R^2 and RMSE values. Guinguina et al. (2019) obtained comparable results, using energy-corrected milk, live weight and estimated dry matter digestibility to predict GFE. In another study on primiparous Canadian Holstein cattle, prediction of weekly feed intake conversion efficiency from milk yield, milk components and live weight yielded a moderate prediction ability, with a much higher RMSE than in the current study (Beard 2018). There is scope to improve the model for predicting GFE developed in the current study by utilising data from novel technologies, such as milk MIR spectra. Extensive research has demonstrated that milk MIR spectra data can considerably improve the performance of models for predicting feed efficiency traits such DMI, residual feed intake and energy intake in dairy cows (McParland et al. 2014; Shetty et al. 2017; Dórea et al. 2018; Wallén et al. 2018; Lahart et al. 2019). The application of other model development approaches, such as partial least squares and artificial neural networks, also warrants investigation for better prediction of GFE (Felipe et al. 2015; Dórea et al. 2018).

Limitations of the study and recommended future work

The current study developed models for predicting DMI and GFE utilising data of first-parity Holstein cows, and achieved reasonable prediction accuracies, as determined by within-herd validation. External (across-herd) validation, which examines whether a prediction model can function outside of the dataset used to create the model, was, however, not performed in this study. Such validation is essential in ensuring that a model is robust enough to achieve the same power of prediction in other populations (Shetty et al. 2017; Lahart et al. 2019). It is also not clear if the models developed in the current study can be reliably extrapolated to multi-parity cows. Thus, further studies are required to determine if the models developed in this study are applicable in other herds, multi-parity cows, breeds and/or other dairy production systems. Such knowledge is a prerequisite to the wide application of these models in different environments. Furthermore, it is important to determine the genetic variation of the predicted DMI and GFE phenotypes, in order to assess their potential for improvement through selection. Further research is also warranted to determine whether there are genes or parts of the genome that are associated with these predicted traits. Such knowledge may assist in implementing marker-assisted selection, by identifying animals that utilise feed efficiently through DNA analysis.

Conclusion

Results of the current study suggest that daily DMI and GFE can be predicted reliably for first-parity Holstein cows, using models comprising of milk production traits and live weight. The prediction models developed in the study may be used to generate large quantities of records of individual cow DMI and GFE, predicted from these easy-to-measure and routinely recorded traits. Provided these predicted traits exhibit considerable genetic variation, such data could be used to estimate accurate EBVs, thus facilitating genetic improvement of feed efficiency through selection. Further research is, however, required to validate these prediction models with external data, as well as extending the analyses to multi-parity cows.

Acknowledgements The authors thankfully acknowledge the Limpopo Dairy (PTY) LTD, for permitting us to conduct a TMR feeding experiment that involved collection of milk samples in their farm and providing the data thereof at no cost. This has been accomplished through the support of Limpopo Academy staff, namely Mr. Peter Els and Miss Tshedza N. Manyuha, as well as directorate of on-farm dairy operation under management of Mr. A.J. Fourie. We would also like to thank Mr. A.E. Muntswu and Mr. R. Groenveld of ARC National Milk Recording and Improvement Scheme, for advices on collection of daily milk samples during the experiment.

Author contribution MAM designed the study, conducted the experiment, collected milk samples, analysed data and wrote the manuscript. CBB and OTZ assisted with the designing of the study and the inputs in the manuscript. All the authors read, edited and approved the final manuscript.

Funding This project received the financial support from the National Research Foundation – Competitive Funding for Rated Researchers, through Agricultural Research Council (ARC) (Project No: P02000135) and National Research Foundation—Doctoral Scholarship (Grant Number: 113928 of the first author) of South Africa. These funds also covered the costs for milk sample analyses at ARC-Elsenburg Analytical Services, Stellenbosch, South Africa.

Data availability The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval The experiment involving the recording of dry matter intake (DMI), milk yield measurement, collection of milk samples and weighing of animals complied with the guidelines of the respective national legislations on animal experimentation and care of animals under the study. This study was granted an animal ethics approval by both Animal Ethics Committee of Agricultural Research Council (APAEC [2020/08]) and Animal Ethics of the University of KwaZulu-Natal (AREC/033/020D).

Conflict of interest The authors declare no competing interests.

References

Beard, S.C., 2018. Evaluating the use of mid-infrared spectroscopy as an indicator of feed efficiency. MSc Thesis. The University of Guelph.

Ben Meir, Y.A., Nikbachat, M., Fortnik, Y., Jacoby, S., Levit, H., Adin, G., Zinder, M.C., Shabtay, A., Gershon, E. and Zachut, M., 2018. Eating behavior, milk production, rumination, and digestibility characteristics of high-and low-efficiency lactating cows fed a low-roughage diet. Journal of Dairy Science, 101, 10973-10984.

Calus, M.P., De Haas, Y., Pszczola, M. and Veerkamp, R., 2013. Predicted accuracy of and response to genomic selection for new traits in dairy cattle. Animal, 7, 183-191.

de Haas, Y., Pryce, J.E., Calus, M.P.L., Wall, E., Berry, D.P., Løvendahl, P., Krattenmacher, N.,Miglior, F., Weigel, K., Spurlock, D., Macdonald, K.A., Hulsegge, B. and Veerkamp, R.F., 2015. Genomic prediction of dry matter intake in dairy cattle from an international data setconsisting of research herds in Europe, North America, and Australasia. Journal of Dairy Science, 98, 6522-6534.

Dórea, J.R.R., Rosa, G.J.M., Weld, K.A. and Armentano, L.E., 2018. Mining data from milk infrared spectroscopy to improve feed intake predictions in lactating dairy cows. Journal of Dairy Science, 101, 5878-5889. https://doi.org/10.3168/jds.2017-13997

Erickson, P.S. and Kalscheur, K. F., 2020. Nutrition and feeding of dairy cattle. F. W. Bazer, G. C.Lamb & G. Wu (Eds.), Animal agriculture. Academic Press, 157–180.

Felipe, V.P.S, Silva, M.A., Valente, B.D. and Rosa, G.J.M., 2015. Using multiple regression, Bayesian networks and artificial neural networks for prediction of total egg production in European quails based on earlier expressed phenotypes. Poultry Science, 94, 772-780.

Guinguina A., Ahvenjärvi, S., Prestløkken, E., Lund, P. and Huhtanen, P., 2019. Predicting feed intake and feed efficiency in lactating dairy cows using digesta marker techniques. Animal, 13 (10), 2277-2288. https://doi.org/10.1017/S1751731119000247

Heinrichs, J. and Ishler, V.A., 2016. Feed Efficiency in Lactating Cows and Relationship to Income Over Feed Costs. https://extension.psu.edu/feed-efficiency-in-lactating-cows and-relationship-to-income-over costs#:~:text=Feed%20efficiency%20(FE%3B%20sometimes%20called,pound%20o%20dr%20matter%20consumed. Accessed: 24 March 2022.

Heinrichs, J. and Jones. C.M., 2016. Milk components: understanding milk fat and protein variation in your dairy herd. https://extension.psu.edu/milk-components understanding-milk fat-and-protein-variation-in-your-dairy-herd. Accessed: 23 March 2022.

Holter, J.B., West, J.W. and McGilliard, M.L., 1997. Predicting ad libitum dry matter intake and yield of Holstein cows. Journal of Dairy Science, 80, 2188-2199. https://doi.org/10.3168/jds.2015-10012.

International Committee for Animal Recording (ICAR), 2012. International agreement of recording practices. Rome, International Committee for Animal Recording (available at www.icar.org/Documents/Rules%20and%20regulations/ Guidelines/Guidelines_2012.pdf).

Kirchgeßner, M., 1997. Tierernährung. 10th ed. DLG-Verlag, Frankfurt, Germany.

Krattenmacher, N., Thaller, G. and Tetens, J., 2019. Analysis of the genetic architecture of energy balance and its major determinants dry matter intake and energy-corrected milk yield in primiparous Holstein cows. Journal of Dairy Science, 102, 3241–3253. https://doi.org/10.3168/jds.2015-10012.

Lahart, B., McParland, S., Kennedy E., Boland T., Condon T., Williams M., Galvin N, McCarthy, B. and Buckley F., 2019. Predicting the dry matter intake of grazing dairy cows using infrared reflectance spectroscopy analysis. Journal of Dairy Science, 102, 8907-8918. https://doi.org/10.3168/jds.2019-16363

Liang, S., Wu, C., Peng, W., Liu, J.-X. and Sun, H.-Z., 2021. Predicting Daily Dry Matter Intake Using Feed Intake of First Two Hours after Feeding in Mid and Late Lactation DairyCows with Fed Ration Three Times Per Day. Animals, (11), 104, 1-11. https://doi.org/10.3390/ani11010104

Liinamo, A.E., Mantysaari, P. and Mantysaari, E.A., 2012. Short communication: Genetic parameters for feed intake, production, and extent of negative energy balance in Nordic Red dairy cattle. Journal of Dairy Science, 95, 6788-6794.

Lindgren, E., Murphy, M. and Andersson, T., 2001. Värdering av foder. LantmännenFoderutveckling AB, Nötfor. Almqvist and Wiksell. Uppsala, Sweden.

Linn, J, Raeth-Knight, M. and Litherland, N., 2009. Role of feed (dairy) efficiency in dairy management. In Proceedings of the 44th Pacific Northwest Animal Nutrition Conference, October 2009, Boise, ID, USA, 167–176.

Lu, Y., Vandehaar, M.J., Spurlock, D.M., Weigel, K.A., Armentano, L.E., Connor, E.E. and Tempelman, R.J., 2018. Genome-wide association analyses based on a multiple-trait approach for modeling feed efficiency. Journal of Dairy Science, 101(4), 3140-3154. https://doi.org/10.3168/jds.2017-13364

Madilindi, M.A., Zishiri, O.T., Dube, B., and Banga, C.B., 2022. Technological advances in genetic improvement of feed efficiency in dairy cattle – A review. Livestock Science, 258,104871, 1-11. https://doi.org/10.1016/j.livsci.2022.104871

Martin, M.J., Dórea, J.R.R., Borchers, M.R., Wallace, R.L., Bertics, S.J., DeNise, S.K., Weigel, K.A. and White, H.M., 2021. Comparison of methods to predict feed intake and residual feed intake using behavioral and metabolite data in addition to classical performance variables. Journal of Dairy Science, 104, 8765-8782. https://doi.org/10.3168/jds.2020-20051

McParland, S., Lewis, E., Kennedy, E., Moore, S.G., McCarthy, B., Butler, S.T. and Berry, D.P., 2014. Mid-infrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows. Journal of Dairy Science, 97, 5863-5871.

McParland, S., and Berry, D.P., 2016. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. Journal of Dairy Science, 99, 4056-4070.

Ngo, T.H.D., 2012. The steps to follow in a multiple regression analysis. In Proceedings of the SAS Global forum, La Puente, CA, USA, 22–25 April 2012, 1–12.

National Milk Recording and Improvement Scheme (NMRIS), 2020. National Milk Recording and Improvement Scheme. Annual Milk Cattle Bulletin, 23, 1-39.

National Research Council (NRC), 2001. Nutrient Requirements of Dairy Cattle. 7th rev. ed., National Academies Press: Washington, DC, USA, ISBN 0309069971.

Poncheki, J.K., Canha, M.L.S, Viechnieski, S.L. and de Almeid, R., 2015. Analysis of daily body weight of dairy cows in early lactation and associations with productive and reproductive performance. Brazilian Journal of Animal Science, 44(5), 187-192. https://doi.org/10.1590/S1806-92902015000500004

Pryce, J.E., Gonzalez-Recio, O., Nieuwhof, G., Wales, W.J., Coffey, M.P., Hayes, B.J. and Goddard, M.E., 2015. Hot topic: Definition and implementation of a breeding value for feed efficiency in dairy cows. Journal of Dairy Science, 98, 7340–50. https://doi.org/10.3168/jds.2015-9621

Searle, T., Graham N.M. and Donnelly, J., 1982. The effect of plane of nutrition on the body composition of two breeds of wearier sheep fed a high protein diet. The Journal of Agricultural Science, 98, 241-245.

Shetty, N., Lovendahl, P., Lund, M.S. and Buitenhuis, A.J., 2017. Prediction and validation of residualfeed intake and dry matter intake in Danish lactating dairy cows using mid infrared spectroscopy of milk. Journal of Dairy Science, 1 00, 253-264. https://doi.org/10.3168/jds.2016-11609

Smith, G., 2018. Step away from stepwise. Journal of Big Data 5, 32. https://doi.org/10.1186/s40537018-0143-6. Accessed: 15 March 2022.

Spurlock, D.M., Dekkers, J.C.M., Fernando, R., Koltes, D.A. and Wolc, A., 2012. Genetic parameters for energy balance, feed efficiency, and related traits in Holstein cattle. Journal of Dairy Science, 95, 5393-5402.

Tempelman, R.J., Spurlock, D.M., Coffey, M., Veerkamp, R.F., Armentano, L.E., Weigel, K.A., de Haas, Y., Staples, C.R., Connor, E.E., Lu, Y. and VandeHaar, M.J., 2015. Heterogeneity in genetic and nongenetic variation and energy sink relationships for residual feed intake across research stations and countries. Journal of Dairy Science, 98, 2013-2026.

Vallimont, J.E., Dechow, C.D., Daubert, J.M., Dekleva M.W., Blum, J.W., Barlieb, C.M., Liu, W., Varga, G.A., Heinrichs, A.J. and Baumrucker, C.R., 2011. Heritability of gross feed efficiency and associations with yield, intake, residual intake, body weight, and body condition score in11 commercial Pennsylvania tie stalls. Journal of Dairy Science, 94, 2108-2113.

VandeHaar, M.J., Armentano, L.E., Weigel, K., Spurlock, D.M., Tempelman, R.J. and Veeramp, R.F, 2016. Harnessing the genetics of the modern dairy cow to continue improvements in feed efficiency. Journal of Dairy Science, 99, 4941-4954.

Veerkamp, R.F., 1998. Selection for economic efficiency of dairy cattle using information on live weight and feed intake: A review. Journal of Dairy Science, 81, 1109-1119.

Wallén, S.E., Prestløkken, E., Meuwissen, T.H.E., Mcparland, S. and Berry. D.P., 2018. Milk midinfrared spectral data as a tool to predict feed intake in lactating Norwegian Reddairy cows, Journal of Dairy Science, 101, 1-12. https://doi.org/10.3168/jds.2017-13874.

Zhang, L, Gengler, N., Dehareng, F., Colinet, F., Froidmont, E. and Soyeurt, H., 2020. Can We Observe Expected Behaviors at Large and Individual Scales for Feed Efficiency Related Traits Predicted Partly from Milk Mid-Infrared Spectra? Animals, 10 (873), 1 -13.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.