


**OPEN ACCESS**
**EDITED BY**

 Hojjat Asadollahpour Nanaei,  
 Université Toulouse III Paul Sabatier,  
 France

**REVIEWED BY**

 Vicente Eliezer Vega-Murillo,  
 Universidad Veracruzana, Mexico  
 Matome Madilindi,  
 Agricultural Research Council, South  
 Africa

**\*CORRESPONDENCE**

 A. Chegini,  
 ✉ arash.chegini@luke.fi

 RECEIVED 23 February 2026  
 REVISED 13 March 2026  
 ACCEPTED 19 March 2026  
 PUBLISHED 15 April 2026

**CITATION**

 Chegini A, Negussie E, Kokkonen T and  
 Lidauer MH (2026) Different feed  
 efficiency modeling approaches for the  
 prediction of genomic breeding values in  
 lactating dairy cows.  
*Front. Genet.* 17:1815864.  
 doi: 10.3389/fgene.2026.1815864

**COPYRIGHT**

 © 2026 Chegini, Negussie, Kokkonen and  
 Lidauer. This is an open-access article  
 distributed under the terms of the [Creative  
 Commons Attribution License \(CC BY\)](#).  
 The use, distribution or reproduction in  
 other forums is permitted, provided the  
 original author(s) and the copyright  
 owner(s) are credited and that the original  
 publication in this journal is cited, in  
 accordance with accepted academic  
 practice. No use, distribution or  
 reproduction is permitted which does not  
 comply with these terms.

# Different feed efficiency modeling approaches for the prediction of genomic breeding values in lactating dairy cows

 A. Chegini<sup>1\*</sup>, E. Negussie<sup>1</sup>, T. Kokkonen<sup>2</sup> and M. H. Lidauer<sup>1</sup>
<sup>1</sup>Natural Resources Institute Finland (Luke), Jokioinen, Finland, <sup>2</sup>Department of Agricultural Sciences, University of Helsinki, Helsinki, Finland

Feed is the main cost of production in dairy farming. Any improvement in feed efficiency (FE) would increase marginal profit and sustainability and mitigate the environmental impact of dairy farming. In this study, we applied single-step genomic best linear unbiased prediction to different feed-efficiency metrics using records collected from Nordic Red dairy cattle (RDC). The main objective was to compare different metrics in terms of their effectiveness in selecting more feed-efficient animals. Weekly observations ( $n = 22,071$ ) of dry-matter intake records from 791 RDC cows collected from 1998 to 2021 were used in this study. The pedigree consisted of 5,604 individuals, of which 1,489 animals were genotyped. Different modeling approaches, including conventional residual feed intake (RFI), regression on expected feed intake (ReFI), two multi-trait residual feed efficiency indices (RFI<sub>Index</sub> and RZFE), and energy conversion efficiency (ECE) were analyzed. For the ReFI approach, two alternatives for predicting the expected feed intake, namely, a prediction equation tailored to the RDC data and a prediction equation based on Holstein dairy cow data proposed by the National Academies of Sciences, Engineering, and Medicine (NRC 2021), were compared. First, a BLUP model was developed, and the necessary variance components were estimated for each approach. Then, pedigree-based and genomic-enhanced breeding values (PEBV and GEBV, respectively) were estimated using either reduced or full datasets. For model validation, PEBV and GEBV estimated using the full dataset were regressed on PEBV and GEBV estimated using the reduced dataset, respectively, to measure bias, dispersion, and prediction accuracy (PAC). The heritability estimates of different residual metrics ranged from 0.23 for RFI to 0.30 for ReFI<sub>NRC2021</sub>, and the repeatability estimates ranged from 0.48 to 0.52. The estimated heritability and repeatability of ECE were 0.23 and 0.56, respectively. For all metrics, the use of genomic information increased PAC. However, there were discrepancies between the metrics in terms of the magnitude of PAC, with the PAC being the highest for ReFI<sub>RDC</sub> and the lowest for RFI<sub>Index</sub>. Similarly, ReFI<sub>RDC</sub> had the lowest bias, while the highest bias was estimated for RFI<sub>Index</sub>. In addition, RZFE and ReFI<sub>RDC</sub> showed lower dispersion. The correlations between GEBV of the residual metrics and the GEBV of ECE were lowest for RFI<sub>NRC2021</sub> and RFI and highest for ReFI<sub>RDC</sub>. Among the metrics compared, ReFI<sub>RDC</sub> and RFI<sub>Index</sub> showed the highest effectiveness in selecting efficient cows. This indicates that the use of appropriate partial regression coefficients and the type of modeling are vital in breeding programs aimed at enhancing FE.

**KEYWORDS**

feed efficiency, prediction accuracy, regression on expected feed intake, residual feed intake, single-step genomic best linear unbiased prediction

## Introduction

Feed cost is the main cost of production in dairy farming; in some production systems, it accounts for up to 75 percent of total production costs (Alqaisi et al., 2011). Therefore, any improvement in the efficiency of feed utilization would improve the profitability of dairy farmers and reduce the emission of greenhouse gases (Negussie et al., 2017). Indirect improvements in feed efficiency (FE) have been observed over the last decades as a consequence of genetic selection for milk yield. This is because by increasing milk yield, maintenance requirement is diluted, and consequently, a higher proportion of energy consumption is allocated to production (Capper et al., 2008). However, it has been stated that further improvement in FE through increased milk yield is difficult as there is uncertainty whether the dilution of maintenance evolves linearly or curvilinearly (VandeHaar and St-Pierre, 2006; Bach et al., 2020). In addition, further selection for increased milk yield can impair fertility and the health status of the animal (Pryce et al., 2004; Oltenacu and Broom, 2010).

Different studies analyzed the possibility of improving FE in different species (i.e., pig, chicken, turkey, and beef), especially in dairy cattle in the last decade (Pryce et al., 2012; Berry and Crowley, 2013; Abdalla et al., 2019; Islam et al., 2020; Wang et al., 2022; Lidauer et al., 2023). Different metrics have been studied to improve FE, including feed conversion ratio, gross energy efficiency, and energy conversion efficiency (ECE) (Liinamo et al., 2015). Although they are easier to calculate and interpret, the ratio traits were not found to be proper metrics for FE because selection for their improvement may impose inequivalent selection pressure on the component traits (Gunsett, 1984). Moreover, they are highly correlated with their components and can increase error variance in the statistical analysis (Berry and Crowley, 2013). Among the suggested metrics, residual feed intake (RFI) is the most studied and well-known metric to improve FE, and it is calculated by subtracting the expected dry matter intake (eDMI) from the actual dry matter intake (DMI) (Koch et al., 1963). Conventionally, eDMI is calculated by regressing DMI on energy sinks, mainly energy corrected milk (ECM), metabolic body weight (MBW), body weight (BW) loss, and BW gain.

Kennedy et al. (1993) stated that RFI is correlated with energy sinks and has some disadvantages; therefore, they proposed a metric that is uncorrelated with energy sinks, known as the genetic RFI (gRFI). However, it was found that gRFI has similar features to RFI in finding feed-efficient animals. This is because the estimated partial phenotypic regression coefficients that model the eDMI in gRFI resemble the estimated partial regression coefficients of RFI (Lu et al., 2015; Lidauer et al., 2023). Lu et al. (2015) stated that the calculation of RFI is based on the assumption that the recording of energy sinks is without measurement errors and that regression coefficient estimates are not confounded. However, both assumptions may not be met and could yield illogical partial regression coefficients for the calculation of eDMI. For instance, the partial regression coefficient for ECM was estimated to be below 0.30 in several studies (Tempelman et al., 2015; Li et al., 2017; Lidauer et al., 2023). Considering the metabolizable energy (ME) required for producing 1 kg of ECM (4,765 kJ/kg; National

Academies of Sciences, Engineering, and Medicine, 2021), this means that the energy content of the feed should be at least 15,883 kJ/kg DMI because otherwise the cows' efficiency of using energy for lactation would be greater than 100 percent. To address this discrepancy, Lidauer et al. (2023) suggested calculating eDMI and deriving partial regression coefficients from energy requirement formulations that were estimated in nutritional studies. They have also demonstrated that regressing DMI on eDMI, i.e., applying the so-called "regression on expected feed intake (ReFI)" metric, would lead to the selection of more feed-efficient animals.

The objectives of this study were to 1) estimate variance components and the pedigree-based and genomic-enhanced breeding values (PEBV and GEBV) for five different feed utilization efficiency metrics and 2) compare the different metrics in terms of their effectiveness and accuracy in selecting feed-efficient animals by applying forward validation and calculating correlations of GEBV among the five different metrics with GEBV and adjusted phenotypes of ECE, which serves as a measurement of the realized overall FE of a lactating cow.

## Materials and methods

### Data

The data used in this study were extracted from the FE research database of the Natural Resources Institute Finland (Luke). All the animals enrolled in the feed-intake recording were primiparous Nordic Red dairy cattle (RDC), the majority of which were recruited from the VikingGenetics nucleus breeding herd and are, therefore, genetically superior cows of the Nordic RDC population. Recording was carried out at four different research farms: two in Jokioinen and one each in Helsinki and Maaninka. The cows at the latter two farms were housed in free-stall barns, while those at the Jokioinen herd were housed in a tie-stall barn until 2009 and in a free-stall barn thereafter.

All cows included in the data were fed *ad libitum*, and the diets were based on grass silage and a home-blend concentrate mixture, which contained an average of 48% concentrate in DM. For the chemical analyses of the diets, samples of grass silages were taken twice a week and combined into 4-week samples, and concentrate samples were taken once a week and combined into 6-week samples. Milk samples were analyzed once a week until lactation week 8 and monthly until the end of lactation using infrared analyzers (MilkoScan FT 6000, Foss, Hillerød, Denmark) at the Valio Ltd. milk laboratory (Seinäjoki, Finland). Data collection started 5 days after calving and continued throughout the whole lactation period, up to a maximum of 305 days in milk (DIM). Feed intake was not recorded during the pasture period; therefore, cows that calved at the end of the calving period (September to December) had gaps in their feed-intake records. The BW observations used in this study were calculated by fitting cow-specific random regression functions to the BW measurements (Mäntysaari and Mäntysaari, 2015) and predicting daily BW and daily BW change observations from the fitted functions.

The raw data consisted of 138,365 weekly and daily observations, from which 22,071 weekly averages for 791 individuals were formed and used for this study. Phenotypic records were available from

TABLE 1 Descriptive statistics of the phenotypic records.

Trait <sup>a</sup>	n	Mean	S.D.	Min	Max	CV (%)
DMI (kg)	22,071	19.4	3.05	8.40	30.96	15.70
MEI (MJ/d)	22,060	212.7	30.49	103.53	328.20	14.33
ECM (kg)	22,032	29.0	4.46	11.58	45.85	15.40
MBW	22,063	119.7	8.91	85.40	156.72	7.45
BWL	4,261	0.35	0.37	0.00	1.70	105.71
BWG	17,267	0.35	0.23	0.00	1.29	65.71
BCS	22,050	3.17	0.33	1.75	4.85	10.36
eDMI <sub>RDC</sub>	21,513	20.0	2.28	7.12	29.98	11.38
eDMI <sub>NRC2021</sub>	22,003	20.8	1.94	12.49	29.35	9.31
eDMI <sub>FTT</sub>	21,524	19.6	1.97	8.35	27.50	10.06
ECE (kg/MJ)	21,954	0.137	0.022	0.060	0.218	15.66

<sup>a</sup>DMI, dry matter intake; MEI, metabolizable energy intake; ECM, energy-corrected milk; MBW, metabolic body weight ( $BW_{0.75}$ ); BWL, body weight loss; BWG, body weight gain; BCS, body condition score; eDMI<sub>RDC</sub>, expected DMI calculated using partial regression coefficients tailored to the RDC data; eDMI<sub>NRC2021</sub>, expected DMI calculated using the DMI predicting equation from National Academies of Sciences, Engineering, and Medicine (2021); eDMI<sub>FTT</sub>, expected DMI calculated by fitting a single-step residual feed intake model; ECE, ECM/metabolizable energy intake (energy conversion efficiency).

1998 to 2021. Descriptive statistics for the different traits are presented in Table 1. The formula stated by Sjaunja et al. (1990) was used to calculate ECM, i.e.,  $ECM = \text{milk yield} \times [(38.30 \times \text{fat content} + 24.20 \times \text{protein content} + 16.54 \times \text{lactose content} + 20.7) / 3,140]$ . The ECE observations were calculated by dividing the ECM observations by the metabolizable energy intake observations. The distribution of phenotypic records within 30-day intervals throughout lactation is shown in Figure 1.

Cows with observations were the daughters of 309 sires and 666 dams, of which 289 sires and 229 dams were genotyped, and from 2004 onwards, 431 cows with FE observations were genotyped. In total, 1,489 animals (762 male and 727 female) were genotyped using the Illumina BovineSNP50 BeadChip (Illumina, San Diego, CA). Using the same edition criteria as the Nordic Cattle Genetic Evaluation (NAV), 46,914 SNP markers remained. The pedigree was pruned for cows with observations by including up to five generations, resulting in a total of 5,604 animals. Animals with unknown parents were assigned to 10 unknown parent groups according to their birth year and origin to account for genetic differences because all cows in the Jokioinen herd were recruited from the RDC nucleus breeding program and were closely related to the RDC elite bull population.

### Statistical analysis

Five different feed-utilization efficiency metrics were studied.

### Regression on the expected feed intake

The first two metrics were based on the ReFI approach (Equation 1) and were modeled using a random regression model as follows:

$$DMI_{himwp} = \alpha_h \times eDMI_{iw} + \beta_m \times eDMI_{iw} + \gamma_{PEi} \times eDMI_{iw} + \delta_{Ai} \times eDMI_{iw} + e_{himwp} \tag{1}$$

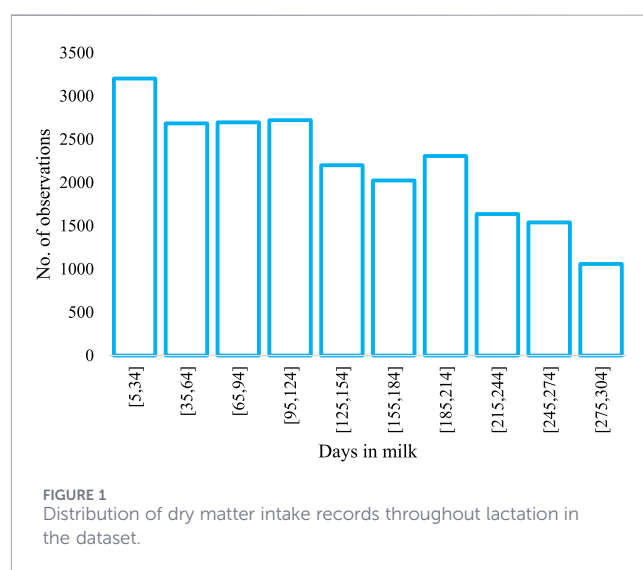


FIGURE 1 Distribution of dry matter intake records throughout lactation in the dataset.

where  $DMI_{himwp}$  is a DMI observation of cow  $i$  in lactation week  $w$  and lactation period class  $p$  recorded in the herd  $\times$  production-year (HPY)  $h$  and herd  $\times$  trial  $\times$  test month (HTM)  $m$  with 791, 43, 5, 28 and 261 effect levels, respectively;  $\alpha_h$  is a fixed regression coefficient nested within HPY;  $\beta_m$  is a random regression coefficient nested within HTM [ $\beta \sim N(0, I \times \sigma_\beta^2)$ , where  $I$  is an identity matrix and  $\sigma_\beta^2$  is the variance of the HTM coefficients];  $\gamma_{PEi}$  is the random regression coefficient for the animal permanent environmental effect [ $\gamma \sim N(0, I \times \sigma_\gamma^2)$ , where  $I$  is an identity matrix and  $\sigma_\gamma^2$  is the variance of the permanent environmental regression coefficients];  $\delta_{Ai}$  is the random regression coefficient for the animal additive genetic effect with [ $\delta \sim N(0, A \times \sigma_\delta^2)$ ] or [ $\delta \sim N(0, H \times \sigma_\delta^2)$ ], where  $\sigma_\delta^2$  is the variance of the additive genetic regression coefficient,  $A$  is the pedigree-based relationship matrix, and  $H$  is the combined pedigree and genomic relationship matrix that will be explained later, respectively; and  $e_{himwp}$  is the residual

random effect [ $e \sim N(\mathbf{0}, \mathbf{R})$ , where  $\mathbf{R} = \sum_{p=1}^5 \oplus \mathbf{I}_{(p)}\sigma_{e_{(p)}}^2$ ] and considering five lactation period classes (lactation month 1, 2, 3, 4–8, and 9–11).

For calculating the expected dry-matter intake ( $eDMI_{iw}$ ) for cow  $i$  that was in lactation week  $w$ , we applied two energy requirement formulations. The first formulation was developed based on the RDC FE data (Lidauer et al., 2023) and considers four energy sinks:

$$eDMI_{iw} \text{ (kg/d)} = 0.456 \times ECM_{iw} + 0.0508 \times MBW_{iw} - 3.25 \times BWL_{iw} + 3.25 \times BWG_{iw},$$

where BWL and BWG are the BW loss and BW gain, respectively. The second formulation was recommended by the National Academies of Sciences, Engineering, and Medicine (2021) for predicting DMI in Holstein cows and was tailored for first parity cows:

$$eDMI_{iw} \text{ (kg/d)} = [3.7 + 0.305 \times MilkE_{iw} + 0.022 \times BW_{iw} - 0.689 \times BCS_{iw}] \times [1 - (0.212) \times e^{(-0.053 \times DIM)}],$$

where MilkE is milk energy (Mcal/d) and BCS is the body condition score. From here onwards, **ReFI<sub>RDC</sub>** refers to the ReFI metric based on the first formulation, and **ReFI<sub>NRC2021</sub>** refers to the ReFI metric based on the second formulation.

### Residual feed intake index

As an alternative to the conventional RFI, a breeding value (BV) for RFI based on an index was developed (**RFI<sub>Index</sub>**). First, the breeding values were predicted from a multi-trait (i.e., Equation 2) model that included five traits, namely, DMI, ECM, MBW, BWL, and BWG, as follows:

$$DMI_{t:ihmlwp} = \sum_{k=1}^5 \gamma_{t:lk} \phi_{iwk} + HPY_{th} + HTM_{tm} + pe_{ti} + u_{ti} + e_{tihmlwp}, \tag{2}$$

where  $DMI_{t:ihmlwp}$  is the observation for trait  $t$  of cow  $i$  made in the lactation week  $w$  in HPY  $h$  within the contemporary group  $m$ ;  $\sum_{k=1}^5 \gamma_{t:lk} \phi_{iwk}$  is the fixed regression function on lactation week  $w$  nested within herd  $l$ , where  $\phi_{iwk}$  is the  $k$ th coefficient of a fourth-order Legendre polynomial;  $HTM_{tm}$  is the random effect of HTM  $m$  on trait  $t$  with  $[\text{var}(HTM) \sim \text{MVN}(\mathbf{0}, \mathbf{I} \otimes \mathbf{M})]$ , where  $\mathbf{I}$  is an identity matrix and  $\mathbf{M}$  is a  $5 \times 5$  matrix of HTM (co)variances] and  $\otimes$  is the Kronecker product;  $pe_{ti}$  and  $u_{ti}$  are the animal permanent environmental and animal additive genetic effects, respectively, for trait  $t$  with  $[\text{var}(pe) \sim \text{MVN}(\mathbf{0}, \mathbf{I} \otimes \mathbf{P})]$ , where  $\mathbf{I}$  is an identity matrix,  $\mathbf{P}$  is a  $5 \times 5$  matrix of permanent environmental (co)variances], and with  $[\text{var}(u) \sim \text{MVN}(\mathbf{0}, \mathbf{A} \otimes \mathbf{V}_a)]$  or  $[\text{var}(u) \sim \text{MVN}(\mathbf{0}, \mathbf{H} \otimes \mathbf{V}_a)]$ , where  $\mathbf{A}$  is the pedigree-based relationship matrix or  $\mathbf{H}$  is the combined single-step relationship matrix, respectively, and  $\mathbf{V}_a$  is a  $5 \times 5$  matrix of animal additive genetic (co)variances]; and  $e_{tihmlwp}$  is the random residual effect with  $[\text{var}(e) \sim \text{MVN}(\mathbf{0}, \mathbf{R})]$ , where  $\mathbf{R}$  is the (co)variance matrix for the residual effects,  $\mathbf{R} = \sum_{p=1}^5 \oplus \mathbf{I}_{(p)}\mathbf{R}_{(p)}$ , considering five lactation period classes].

In the second step, for each animal  $i$ , a pedigree-based or single-step based BV index, (i.e.,  $RFI_{Index}$ ; Equation 3), was derived by

applying the same index weights as used for  $ReFI_{RDC}$  to an animal's estimated multi-trait BV:

$$RFI_{Indexi} = BV_{DMIi} - (0.456 \times BV_{ECMi} + 0.0508 \times BV_{MBWi} - 3.25 \times BV_{BWLi} + 3.25 \times BV_{BWGi}), \tag{3}$$

where  $BV_{ti}$  is either the estimated PEBV or GEBV for the trait  $t$  and animal  $i$  and the index weight values are the same values as the energy requirement values applied for the  $ReFI_{RDC}$  metric.

### RZFeedEfficiency

The recently published German FE index for Holstein dairy cows (Equation 4) differs from the  $RFI_{Index}$  metrics in that it does not account for the feed requirement for maintenance; therefore, it resembles a simplified version of the feed saved index proposed by Pryce et al. (2015). The German index is called RZFeedEfficiency (RZFE; Abdalla et al., 2024), and in this study, it was calculated for an animal  $i$  by the same procedure as described for the  $RFI_{Index}$  but by applying weights that are used in the German index, as follows:

$$RZFE_i = (0.4 \times BV_{ECMi} - 4.5 \times BV_{BWL_i} + 4.5 \times BV_{BWGi}) - BV_{DMI_i}, \tag{4}$$

wherein, in addition to the differences with  $RFI_{Index}$ , the index weights are also reversed, and thus, higher positive RZFE values are favorable and indicate higher efficiency.

### Residual feed intake

RFI is the fifth metric (Equation 5), which is a common form of a one-step RFI model, and it was applied as follows:

$$DMI_{himwp} = c1 \times ECM_{iw} + c2 \times MBW_{iw} + c3 \times BWL_{iw} + c4 \times BWG_{iw} + HPY_h + HTM_m + pe_i + u_i + e_{himwp}, \tag{5}$$

where  $DMI_{himwp}$  is a DMI observation of cow  $i$  in lactation week  $w$  recorded in HPY  $h$  and HTM  $m$ ;  $c1$ – $c4$  are the partial regression coefficients for the regression of DMI on the energy sinks. The model effects had the same number of effect levels as given in model (1). The HPY effect was modeled as a fixed effect, and the random  $HTM_m$ ,  $pe_i$ ,  $u_i$ , and  $e_{himwp}$  effects were modeled in the same way as in model (2) but with a univariate analysis.

### Energy conversion efficiency

It is an ECE trait used for benchmarking the five studied feed-utilization efficiency metrics, where ECE does not involve any partial regression coefficients and represents a clearly established FE trait. The model used for the analysis of ECE (Equation 6) included the same fixed and random effects as in model (5), with the exception that the energy sinks in model (5) were replaced with a fixed effect for lactation week:

$$ECE_{himwp} = LWK_w + HPY_h + HTM_m + pe_i + u_i + e_{himwp}, \tag{6}$$

where  $ECE_{himwp}$  is the ECE observation of cow  $i$  in lactation week  $w$  recorded in HPY  $h$  and HTM  $m$ ,  $LWK_w$  is the fixed effect of lactation

week  $w$ , and all other effects in the model were specified in the same way as in model (5) but with a univariate analysis.

### Estimation of variance components and the prediction of breeding values

The required variance components were estimated using pedigree BLUP models and a Monte Carlo expectation-maximization REML algorithm (Matilainen et al., 2012). The heritabilities and variances for the  $RFI_{index}$  and RZFE metrics were derived based on the variance components estimated for model (2) and the applied index weights of the metrics. The models developed were used to predict the necessary PEBV and GEBV for forward validation, as described in the following chapter (see the Model validation section). For the estimation of GEBV, we carried out single-step GBLUP analyses by substituting the inverse of the numerator relationship matrix  $\mathbf{A}^{-1}$  with the inverse of the single-step relationship matrix  $\mathbf{H}^{-1}$  in the BLUP models (Aguilar et al., 2010; Christensen and Lund, 2010), i.e.,  $\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$ , where  $\mathbf{G}^{-1}$  is the inverse of the weighted genomic relationship matrix,  $\mathbf{G} = \mathbf{G}^* \times 0.9 + \mathbf{A}_{22} \times 0.1$ , and  $\mathbf{A}_{22}^{-1}$  is the inverse of the numerator relationship matrix for the genotyped animals. Method 1 in VanRaden (2008) was used to construct genomic relationship matrix  $\mathbf{G}^*$ . The Relax2 program (Strandén and Vuori, 2006) was used for pruning and validating the pedigree, the Htginv\_v7 program (Strandén and Mäntysaari, 2018) was used to construct the  $\mathbf{H}^{-1}$  matrix, and MiX99 software suite (Pitkänen et al., 2022) was used for estimating BV.

### Model validation

To assess the predictive ability of the different metrics, the linear regression forward validation method described by Legarra and Reverter (2018) (LR) was used. In this method, PEBV and GEBV from the full dataset ( $PEBV_f$  and  $GEBV_f$ ) are regressed on PEBV and GEBV from the reduced dataset ( $PEBV_r$  and  $GEBV_r$ ), respectively, as follows:  $(P)GEBV_f = b_0 + b_1 \times (P)GEBV_r$ , where  $b_0$  and  $b_1$  are the bias (intercept) and dispersion (regression coefficient), respectively, and the correlation between  $(P)GEBV_f$  and  $(P)GEBV_r$  is the prediction accuracy (PAC). The 88 youngest genotyped animals, born between 2017 and 2019, which had records in the full dataset but neither records nor daughters with records in the reduced dataset, were included in the validation group.

The Delta method was applied to calculate the standard errors of heritabilities ( $h^2$ ) and repeatabilities ( $r$ ). In addition, the ordinary nonparametric bootstrap method (R Development Core Team, 2012) was applied to calculate the standard errors of estimates for different metrics to assess the consistency of the estimates. Furthermore, Pearson and Spearman correlations between GEBV of animals with phenotypes for the different metrics were calculated. The ECE phenotypes were adjusted for all the effects estimated with model (6) except for the animal additive genetic effect to obtain adjusted phenotypes that are equivalent to ECE yield deviation, and the correlations between these adjusted ECE phenotypes and the cows' GEBV for the different metrics were calculated.

## Results and discussion

### Summary statistics and partial regression coefficients

Descriptive statistics for the phenotypes of the studied feed efficiency data and the applied alternative eDMI values are provided in Table 1. On average, a cow consumed 19.4 kg DM/d containing 212.7 MJ ME and produced 29.0 kg ECM with 4.43% milk fat and 3.63% milk protein. The average energy density of the diets was 10.94 MJ ME per kg DM, which is consistent with reports from previous studies (Mäntysaari et al., 2012; Karlsson et al., 2022). The averages of eDMI using different energy requirement formulations were relatively similar. When applying the RFI model, the average of eDMI was the closest to the average of the observed DMI. In contrast, the formulation recommended in the National Academies of Sciences, Engineering, and Medicine (2021) overestimated the DMI by 1.4 kg. The differences are due to differences in the regression coefficients of energy sinks between different formulations. For instance, a lower requirement coefficient has been allocated to ECM (0.229 vs. 0.276 for RFI) and a higher requirement coefficient to MBW (0.108 vs. 0.068 for RFI) in the National Academies of Sciences, Engineering, and Medicine (2021) compared to that of the other formulations. In addition, the National Academies of Sciences, Engineering, and Medicine (2021) has a requirement coefficient for BCS in the formula ( $-0.689$ ), which makes it somewhat different from the other formulations. Moreover, the prediction formula for DMI presented in the National Academies of Sciences, Engineering, and Medicine (2021) was developed for Holstein cows and used observations from both primiparous and multiparous cows, which may have different metabolism and energy requirements. The SD of the eDMI was the largest for the formulation that was developed for the RDC FE data, which is most likely because, among all studied formulations, it had the largest energy requirement coefficient for ECM.

The partial regression coefficients for the different metrics are summarized in Table 2. To make the partial regression coefficients of the different metrics comparable, we converted the coefficients of the eDMI prediction equation from the National Academies of Sciences, Engineering, and Medicine (2021) to the ECM and MBW scales. Assuming that 1 kg of ECM has 0.75 Mcal NEL, it leads to a  $ReFI_{NRC2021}$  coefficient of 0.229 kg DM per kg ECM. The average MBW in our data was  $119.7 \text{ kg}^{0.75}$ , which corresponds to a BW of 590 kg. Using the coefficient of 0.022 kg DM/kg BW from equation 2-1 in the National Academies of Sciences, Engineering, and Medicine (2021) resulted in 0.108 as the coefficient for MBW. However, apart from the National Academies of Sciences, Engineering, and Medicine (2021) formulation applied in this study, the National Academies of Sciences, Engineering, and Medicine (2021) proposed increasing the energy requirement for maintenance by 25% compared to NRC (2001), which would result in a coefficient of 0.121 for MBW.

When comparing the partial regression coefficients of the formulations that were developed by considering nutritional studies ( $ReFI_{RDC}$ ,  $RFI_{Index}$ , and RZFE) with those that are

TABLE 2 Partial regression coefficients used in different metrics.

Metric	ECM	MBW	BWL	BWG	BCS
ReFI <sub>RDC</sub>	0.456	0.051	-3.25	3.25	—
ReFI <sub>NRC2021</sub> <sup>a</sup>	0.229 <sup>b</sup>	0.108 <sup>c</sup>	—	—	-0.689
RFI <sub>Index</sub>	0.456	0.051	-3.25	3.25	—
RZFE	0.400	—	-4.50	4.50	—
RFI	0.276	0.068	-3.80	1.44	—

ECM, energy-corrected milk; MBW, metabolic body weight ( $BW^{0.75}$ ); BWL, body weight loss; BWG, body weight gain; BCS, body condition score; ReFI<sub>RDC</sub> = regression on expected feed intake using expected DMI from tailored coefficients; ReFI<sub>NRC2021</sub> = regression on expected feed intake using expected DMI from NASEM (2021); RFI<sub>Index</sub> = multi-trait residual feed intake; RZFE = German feed efficiency index; RFI = single step residual feed intake.

<sup>a</sup>Expected DMI for this scenario was calculated using the DMI predicting equation 2-1 from the National Academies of Sciences, Engineering, and Medicine (2021) with coefficients that correspond to the values given here.

<sup>b</sup>This was calculated by assuming that 1 kg of ECM has 0.752 Mcal NEL according to equation 2-1 in the National Academies of Sciences, Engineering, and Medicine (2021).

<sup>c</sup>Average MBW in our data was 119.7, which corresponds to a BW of 590 kg. Using the coefficient from equation 2-1 resulted in 0.108 as the coefficient for MBW.

TABLE 3 Estimates of the genetic, residual, permanent environmental ( $Var_G$ ,  $Var_R$ , and  $Var_{PE}$ , respectively), variances, heritability ( $h^2$ ), and repeatability ( $r$ ) for regression of feed intake calculated with different criteria using a repeatability animal model.

Metric	$Var_{PE}$	$Var_G$	$Var_R^*$	$h^2$ S.E.	$r$ S.E.
ReFI <sub>RDC</sub>	0.00333	0.00303	2.502	0.24 <sub>0.20</sub>	0.50 <sub>0.10</sub>
ReFI <sub>NRC2021</sub>	0.00244	0.00334	2.348	0.30 <sub>0.21</sub>	0.52 <sub>0.11</sub>
RFI <sub>Index</sub>	1.109	1.137	2.166	0.26 <sub>NA</sub>	0.51 <sub>NA</sub>
RZFE	1.078	1.409	2.330	0.29 <sub>NA</sub>	0.52 <sub>NA</sub>
RFI	1.127	1.043	2.369	0.23 <sub>0.20</sub>	0.48 <sub>0.11</sub>
ECE	0.000118	0.000081	0.000157	0.23 <sub>0.21</sub>	0.56 <sub>0.09</sub>

$Var_{PE}$ , permanent environmental variance;  $Var_G$ , additive genetic variance;  $Var_R$ , residual variance;  $h^2$ , heritability;  $r$ , repeatability; ReFI<sub>RDC</sub> = regression on expected feed intake using expected DMI from tailored coefficients; ReFI<sub>NRC2021</sub> = regression on expected feed intake using expected DMI from NASEM (2021); RFI<sub>Index</sub> = multi-trait residual feed intake; RZFE = German feed efficiency index; RFI = single step residual feed intake; ECE = energy conversion efficiency.

\*There were five lactation periods, and the weighted average of the residual variances for each approach were calculated and subsequently used to calculate heritabilities; NA, not available.

based on partial regression analyses (ReFI<sub>NRC2021</sub> and RFI), it can be seen that formulations based on partial regression analyses provide lower regression coefficients for ECM and higher regression coefficients for MBW. The partial regression coefficients obtained by applying the one-step RFI model were 0.276 for ECM and 0.068 for MBW, while the corresponding values based on RDC data were 0.456 and 0.0508, respectively. The requirement formulations for ReFI<sub>RDC</sub> and RZFE differ substantially in the allocation of feed toward milk production and growth. For ReFI<sub>RDC</sub>, the coefficient for BWG is 7.1 times larger than that for ECM. In contrast, for RZFE, the coefficient for BWG is 11.3 times larger than that for ECM, implying that the lower conversion efficiency of DMI to body tissue has been considered in the German RZFE metric.

## Variance components

Variance components,  $h^2$ , and repeatability for the five feed-utilization efficiency metrics and for ECE are provided in Table 3. Combined residual variance estimates, calculated as the weighted averages of the residual estimates for the five lactation periods, were used to facilitate comparisons. Among the five metrics, the lowest residual variance was observed for RFI<sub>Index</sub>, followed by

similar estimates for RZFE, ReFI<sub>NRC2021</sub>, and RFI, with the highest estimate for ReFI<sub>RDC</sub>. Nevertheless, heritabilities were on the same level for RFI, ReFI<sub>RDC</sub>, and RFI<sub>Index</sub> (0.23, 0.24, and 0.26, respectively) and slightly higher for RZFE and ReFI<sub>NRC2021</sub> (0.29 and 0.30, respectively). This may be due to the estimated genetic variance for ReFI<sub>RDC</sub> ( $0.00303 \times 20^2 = 1.21 \text{ kg}^2$ ), which was higher than those estimated for RFI and RFI<sub>Index</sub>. We obtained the highest additive genetic variance estimate for ReFI<sub>NRC2021</sub> ( $0.00334 \times 20.8^2 = 1.45 \text{ kg}^2$ ), followed by that for RZFE ( $1.41 \text{ kg}^2$ ). The estimated repeatabilities were rather similar for ReFI<sub>RDC</sub>, ReFI<sub>NRC2021</sub>, RFI<sub>Index</sub>, and RZFE (0.50–0.52) but slightly lower for RFI (0.48). Due to the size of the available data, the heritability and reliability estimates were associated with large standard errors.

Reports on  $h^2$  of RFI in dairy cows from different countries agreed well with the estimates from the present study, which ranged from 0.16 to 0.38 (Pryce et al., 2012; Lu et al., 2015; Stephansen et al., 2022). Similar values of  $h^2$  have been reported for Holstein dairy heifers (0.24) (Khanal et al., 2023), Nellore cattle (0.21) (Brunes et al., 2023), and other species such as turkeys (0.15) (Abdalla et al., 2019) and pigs (0.30) (Wang et al., 2022). The magnitude of the estimated genetic variance and  $h^2$  indicates that sufficient genetic variation for FE exists, and these traits can be improved through selection.

TABLE 4 Results of forward validation by the Legarra–Reverter procedure (regression of PEBV and GEBV predicted using the full dataset on PEBV and GEBV predicted using the reduced dataset, respectively).

Metric <sup>a</sup>	Forward validation parameter			
	Model <sup>b</sup>	b0	b1	corr
ReFI <sub>RDC</sub>	BLUP	0.007 <sub>0.002</sub>	0.724 <sub>0.137</sub>	0.41 <sub>0.065</sub>
	ssGBLUP	-0.001 <sub>0.002</sub>	0.638 <sub>0.109</sub>	0.53 <sub>0.069</sub>
ReFI <sub>NRC2021</sub>	BLUP	-0.003 <sub>0.003</sub>	0.514 <sub>0.165</sub>	0.30 <sub>0.053</sub>
	ssGBLUP	0.003 <sub>0.002</sub>	0.518 <sub>0.117</sub>	0.43 <sub>0.072</sub>
RFI <sub>Index</sub>	BLUP	-0.171 <sub>0.001</sub>	0.479 <sub>0.001</sub>	0.17 <sub>0.030</sub>
	ssGBLUP	-0.244 <sub>0.001</sub>	0.454 <sub>0.002</sub>	0.33 <sub>0.055</sub>
RZFE	BLUP	0.234 <sub>0.001</sub>	0.951 <sub>0.001</sub>	0.30 <sub>0.062</sub>
	ssGBLUP	0.232 <sub>0.001</sub>	0.735 <sub>0.004</sub>	0.40 <sub>0.060</sub>
RFI	BLUP	-0.149 <sub>0.037</sub>	0.547 <sub>0.168</sub>	0.30 <sub>0.052</sub>
	ssGBLUP	-0.097 <sub>0.035</sub>	0.605 <sub>0.116</sub>	0.47 <sub>0.072</sub>
ECE	BLUP	-0.0004 <sub>0.0004</sub>	0.552 <sub>0.333</sub>	0.18 <sub>0.030</sub>
	ssGBLUP	-0.0007 <sub>0.0004</sub>	0.747 <sub>0.161</sub>	0.43 <sub>0.061</sub>

<sup>a</sup>ReFI<sub>RDC</sub> = regression on expected feed intake using expected DMI from tailored coefficients; ReFI<sub>NRC2021</sub> = regression on expected feed intake using expected DMI from NASEM (2021);

RFI<sub>Index</sub> = multi-trait residual feed intake; RZFE = German feed efficiency index; RFI = single step residual feed intake; ECE = energy conversion efficiency.

<sup>b</sup>BLUP and ssGBLUP are regressions of PEBV and GEBV using the full dataset on PEBV and GEBV using the reduced dataset, respectively.

The reason for the higher  $h^2$  estimate for ReFI<sub>NRC2021</sub> compared to that for ReFI<sub>RDC</sub> is unclear. The energy requirement formulation applied for ReFI<sub>NRC2021</sub> allocates less DMI to ECM and more DMI toward BW compared to the energy requirement formulation for ReFI<sub>RDC</sub>. If the difference between the observed and expected DMI is less influenced by the actual difference in FE and more affected by inappropriate coefficients, it may inflate heritability as the difference may capture variations caused by sink traits rather than by variations in FE. The higher  $h^2$  estimate for RZFE may be due to the fact that RZFE is an index trait composed of a metabolic efficiency component, similar to the other metrics in this study, and of a maintenance component that accounts for the size of the cow.

The ECE trait, which was used for validation purposes in this study, had a relatively similar  $h^2$  to other FE metrics. Due to the limited availability of both DMI records and assays of ME content, estimates of ECE are scarce in the literature. Performing a longitudinal analysis, Hurley et al. (2017) estimated  $h^2$  for ECE to be between 0.06 in DIM 60 to 0.28 in DIM 250. Liinamo et al. (2015) reported lactational  $h^2$  of ECE to be 0.16 for RDC cows. The estimated  $h^2$  value for gross feed efficiency was 0.32 for Holstein cows at Iowa State University's Research Farm (Spurlock et al., 2012). In another study using a larger dataset, Vallimont et al. (2011) reported an  $h^2$  value of 0.14 for DMI efficiency, which was the ratio of fat-corrected milk to DMI.

## Validation of metrics

Forward validation results obtained using the LR procedure are presented in Table 4. The results show that the biases obtained from pedigree- and genomic-based prediction models differed, indicating lower biases for ReFI<sub>RDC</sub>, ReFI<sub>NRC2021</sub>, and ECE and

higher biases for the index metrics RFI<sub>Index</sub> and RZFE. Among the different metrics, RFI<sub>Index</sub>, RZFE, and RFI showed higher prediction bias for both pedigree- and genomic-based predictions, ranging between 0.09 and 0.22 genetic standard deviations. The dispersion estimates for the pedigree- and genomic-based prediction models were lower than 1.0. Among the genomic models, the lowest dispersion estimate (0.454) was obtained for RFI<sub>Index</sub> and the highest (0.951) was for RZFE. In addition, dispersions estimated using genomic predictions were generally associated with lower standard errors, indicating greater stability of the estimates. Correlations between GEBV<sub>f</sub> and GEBV<sub>r</sub> for different metrics ranged from 0.30 to 0.53 and were significantly higher than the correlations between PEBV<sub>f</sub> and PEBV<sub>r</sub> (from 0.17 to 0.41; Table 4). Among the metrics, ReFI<sub>RDC</sub> yielded the highest PAC, followed by RFI. Although ReFI<sub>RDC</sub> and ReFI<sub>NRC2021</sub> had the same modeling, ReFI<sub>RDC</sub> yielded 23% higher PAC. This can be attributed to the applied regression coefficients in the calculation of eDMI, which differed considerably (Table 2). The applied regression coefficients for eDMI in ReFI<sub>RDC</sub> are derived from nutritional studies and reflect the DMI requirement for ECM, MBW, and BW change. On the contrary, the regression coefficients applied in calculating eDMI for ReFI<sub>NRC2021</sub> were obtained from the regression analyses reported by the National Academies of Sciences, Engineering, and Medicine (2021). Partial regression analyses ensure good model fit in modeling the total energy requirement. Nevertheless, a partial coefficient for a specific energy sink might be suboptimal, which reduces the accuracy of the individual eDMI values. For instance, for ReFI<sub>NRC2021</sub>, the applied coefficient for ECM underestimates the expected feed requirement for ECM. Furthermore, the coefficient of variation of ECM is higher than that of MBW, which enables ReFI<sub>RDC</sub> to

TABLE 5 Pearson (above diagonal) and Spearman (below diagonal) correlations (S.E. are subscripted) between GEBV of the animals with phenotype (n = 791) for different metrics of feed-utilization efficiency.

Metric <sup>a</sup>	ReFI <sub>RDC</sub>	ReFI <sub>NRC2021</sub>	RFI <sub>Index</sub>	RZFE	RFI	ECE
ReFI <sub>RDC</sub>		0.84 <sub>0.02</sub>	0.82 <sub>0.02</sub>	-0.75 <sub>0.03</sub>	0.93 <sub>0.01</sub>	-0.70 <sub>0.02</sub>
ReFI <sub>NRC2021</sub>	0.80 <sub>0.02</sub>		0.84 <sub>0.02</sub>	-0.69 <sub>0.03</sub>	0.92 <sub>0.02</sub>	-0.54 <sub>0.03</sub>
RFI <sub>Index</sub>	0.77 <sub>0.03</sub>	0.78 <sub>0.03</sub>		-0.92 <sub>0.02</sub>	0.88 <sub>0.02</sub>	-0.69 <sub>0.03</sub>
RZFE	0.67 <sub>0.03</sub>	0.60 <sub>0.03</sub>	0.88 <sub>0.02</sub>		-0.84 <sub>0.02</sub>	0.62 <sub>0.03</sub>
RFI	0.91 <sub>0.01</sub>	0.90 <sub>0.02</sub>	0.85 <sub>0.02</sub>	0.78 <sub>0.03</sub>		-0.58 <sub>0.03</sub>
ECE	0.66 <sub>0.03</sub>	0.50 <sub>0.03</sub>	0.65 <sub>0.03</sub>	0.55 <sub>0.03</sub>	0.54 <sub>0.03</sub>	

<sup>a</sup>ReFIRDC = regression on expected feed intake using expected DMI from tailored coefficients; ReFINRC2021 = regression on expected feed intake using expected DMI from NASEM (2021); RFI<sub>Index</sub> = multi-trait residual feed intake; RZFE = German feed efficiency index; RFI = single step residual feed intake; ECE = energy conversion efficiency.

capture a greater proportion of total variation in the composite trait and, thus, better stabilize the training and validation GEBV and ensure that they are on the same scale. On the other hand, ReFI<sub>RDC</sub> includes BWL and BWG, which are measured more accurately compared to the subjective nature of BCS in the National Academies of Sciences, Engineering, and Medicine (2021). In addition, for ECE, the correlation between GEBV<sub>f</sub> and GEBV<sub>r</sub> was considerably higher than that between PEBV<sub>f</sub> and PEBV<sub>r</sub> (0.43 vs. 0.18). Generally, FE metrics with single-step modeling led to higher PAC.

The results indicate that including genomic information in the evaluation of animals leads to higher PAC. Using different approaches (GBLUP and BayesA), Pryce et al. (2012) reported the average PAC for RFI values of 0.37 and 0.31 in Australian and New Zealand dairy heifers, respectively. In addition, Li et al. (2020) estimated an average PAC of 0.34 for RFI for all phenotyped animals using single-step GBLUP in their study. In another study, Abdalla et al. (2019) reported improved dispersion and higher PAC using single-step GBLUP compared to PBLUP for RFI in turkey. The highest attainable PAC using genomic selection may vary depending on the specific application and data used. In general, it can range from approximately 0.4 to 0.9, implying that between 40% and 90% of the genetic variance in a trait can be explained by genomic data.

Despite the LR forward prediction method allowing the assessment of the quality of predicted BVs for the metrics compared in this study, it does not allow the assessment of the metrics' potential for selecting the overall feed-efficient animals. Therefore, we assess the correlations between the GEBV of the different metrics with genomic breeding values for ECE. In summary, ECE is easy to calculate and understand and is the actual FE trait that dairy farmers want to improve. Pearson and Spearman correlations between GEBV of animals for different residual metrics and ECE are shown in Table 5. The GEBV for ReFI<sub>RDC</sub> and RFI<sub>Index</sub> had the highest Pearson correlations with the GEBV for ECE (-0.70 and -0.69, respectively) and the highest Spearman correlations, whereas the GEBV for RFI<sub>NRC2021</sub> and RFI had the lowest Pearson correlation with GEBV for ECE (-0.54 and -0.58, respectively). This implies that the selection of individuals with lower GEBV for ReFI<sub>RDC</sub> or RFI<sub>Index</sub> would best improve the overall FE compared to selection based on GEBV, which is based on the other studied metrics. Contrary

TABLE 6 Correlations (S.E. are subscripted) between ECE phenotypes adjusted for all effects in the model except the animal additive genetic effect and GEBV of different metrics (n = 791).

Metric <sup>a</sup>	Correlation
ReFI <sub>RDC</sub>	-0.64 <sub>0.03</sub>
ReFI <sub>NRC2021</sub>	-0.47 <sub>0.03</sub>
RFI <sub>Index</sub>	-0.58 <sub>0.03</sub>
RZFE	0.52 <sub>0.03</sub>
RFI	-0.49 <sub>0.03</sub>
ECE	0.88 <sub>0.02</sub>

<sup>a</sup>ReFIRDC = regression on expected feed intake using expected DMI from tailored coefficients; ReFINRC2021 = regression on expected feed intake using expected DMI from NASEM (2021); RFI<sub>Index</sub> = multi-trait residual feed intake; RZFE = German feed efficiency index; RFI = single step residual feed intake; ECE = energy conversion efficiency.

to our expectation, the correlation between ReFI<sub>RDC</sub> and RFI<sub>Index</sub> was smaller than that between ReFI<sub>RDC</sub> and RFI (0.82 vs. 0.93), and it deviated from unity, although both used the same set of partial regression coefficients. This is because of the modeling, as ReFI<sub>RDC</sub> and RFI were performed in a single-step manner, while for RFI<sub>Index</sub>, we first predicted multi-trait (G)EBVs and then combined them.

The latter findings are also supported by the correlations between ECE phenotypes that are adjusted for all effects in the model except for the animal additive genetic effect and the GEBV of different metrics (Table 6). Among the studied metrics, the animals' GEBV for ReFI<sub>RDC</sub> had the highest correlation (-0.64), and GEBV for RFI<sub>NRC2021</sub> (-0.47) had the lowest correlation with the adjusted phenotypes of ECE. The reason for the correlation of ECE being higher with ReFI<sub>RDC</sub> compared to that with RZFE requires further analysis. It could be expected that RZFE has a higher correlation with ECE because the RZFE index includes the DMI needed for maintenance as an additional component, and thus, it should better resemble ECE. One reason could be that the applied energy requirement coefficient for ECM in the RZFE metric is rather low (0.4) and corresponds to a very high efficiency of using ME for lactation of  $k_l = 0.72$ , which most likely has not been achieved by the cows. On the other hand, ReFI<sub>RDC</sub> is based on a random regression approach and, therefore, has the same capability as ECE to account for the

multiplicative effect of efficiency factors (Lidauer et al., 2023). GEV for  $RFI_{NRC2021}$ , showing a relatively low correlation with the adjusted ECE observations ( $-0.47$ ), is most likely due to the very low energy requirement coefficient for ECM, which might be in poor agreement with the DMI requirement for production. Therefore, considering an animal's specific energy requirements for various energy sinks, along with the choice of an appropriate model, is essential when aiming to maximize the genetic gains for FE in a breeding program. Although all the observations in this study were from primiparous cows, as the calculations were based on the unit of energy sinks, it can be expected that similar results would also be obtained for multiparous cows.

## Conclusion

Different FE metrics based on various models and energy requirement systems were compared in this study. Genomic information was also included to assess the extent of improvement in prediction accuracy that could potentially be achieved by implementing single-step GBLUP. The results indicated that both differences in energy requirement systems and differences in the modeling process would lead to different rankings of animals varying in their feed-utilization efficiency. When using the same model,  $ReFI_{RDC}$  outperformed  $ReFI_{NRC2021}$ , yielding 23% higher prediction accuracy, which underscores the importance of choosing proper partial regression coefficients (or predicting equations) to calculate the expected DMI. On the other hand, when using the same partial regression coefficients but different modeling processes,  $ReFI_{RDC}$  led to higher prediction accuracy and showed a slightly higher correlation with ECE compared with  $RFI_{Index}$ . The results indicate that modification in the modeling of FE traits would lead to a significant genetic gain. It also indicates that considerable improvement in FE could be achieved even when a small number of phenotyped and genotyped animals are available. However, a larger reference population is needed to achieve higher prediction reliabilities and accelerate the genetic improvement for FE.

## Data availability statement

The dataset used in this study (obtained from two research farms in Jokioinen, one in Helsinki and one in Maaninka) belong to Natural Resources Institute (LUKE) and hence, are not publicly available. Data are however available upon required agreement with LUKE and should be directed to the last author Martin H. Lidauer, martin.lidauer@luke.fi.

## Ethics statement

The requirement of ethical approval was waived by the National Ethics Committee of Finland for the studies involving animals because collecting the feed efficiency data at the four research herds complied with the ethical standards and did not require approval by the National Ethics Committee. The studies were

conducted in accordance with the local legislation and institutional requirements.

## Author contributions

AC: Formal analysis, Validation, Visualization, Writing – original draft. EN: Data curation, Writing – review and editing. TK: Writing – review and editing. ML: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review and editing.

## Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was funded by the Finnish Ministry of Agriculture (453/03.01.02/2018), Luke, and the dairy farmers' owned companies and corporations, VikingGenetics Ltd. (Denmark), Valio Ltd. (Finland), Nordic Cattle Genetic Evaluation NAV (Denmark), and Faba co-op (Finland). The funders had no role in the study design, data collection, analysis or interpretation of data, the writing of the manuscript, or the decision to submit it for publication.

## Acknowledgements

The authors acknowledge the Finnish Ministry of Agriculture, VikingGenetics, Valio, NAV, and Faba for financially supporting this study.

## Conflict of interest

The authors declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Abdalla, E. E. A., Schenkel, F. S., Emamgholi Begli, H., Willems, O. W., van As, P., Vanderhout, R., et al. (2019). Single-step methodology for genomic evaluation in turkeys (*meleagris gallopavo*). *Front. Genet.* 10, 1248. doi:10.3389/fgene.2019.01248
- Abdalla, E., Polman, L., Rensing, S., Liu, Z., Alkhoder, H., Reents, R., et al. (2024). Breeding for feed efficiency in German holsteins: the new RZFeedEfficiency. *Proc 2024 interbull mtg. Interbull Bull.* 60, 47–53.
- Aguiar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., and Lawlor, T. J. (2010). Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of holstein final score. *J. Dairy Sci.* 93, 743–752. doi:10.3168/jds.2009-2730
- Alqaisi, O., Asaah, O., and Hemme, T. (2011). Global view on feed cost and feed efficiency on dairy farms. *All About Feed* 2, 1–5.
- Bach, A., Terré, M., and Vidal, M. (2020). Symposium review: decomposing efficiency of milk production and maximizing profit. *J. Dairy Sci.* 103 (6), 5709–5725. doi:10.3168/jds.2019-17304
- Berry, D. P., and Crowley, J. J. (2013). Genetics of feed efficiency in dairy and beef cattle. *J. Anim. Sci.* 91, 1594–1613. doi:10.2527/jas.2012-5862
- Brunes, L. C., de Faria, C. U., Magnabosco, C. U., Lobo, R. B., Peripolli, E., Aguiar, I., et al. (2023). Genomic prediction ability and genetic parameters for residual feed intake calculated using different approaches and their associations with growth, reproductive, and carcass traits in nellore cattle. *J. Appl. Genet.* 64 (1), 159–167. doi:10.1007/s13353-022-00734-8
- Capper, J. L., Castañeda-Gutiérrez, E., Cady, R. A., and Bauman, D. E. (2008). The environmental impact of recombinant bovine somatotropin (rbST) use in dairy production. *Proc. Natl. Acad. Sci. U. S. A.* 105, 9668–9673. doi:10.1073/pnas.0802446105
- Christensen, O. F., and Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genet. Sel. Evol.* 42, 2. doi:10.1186/1297-9686-42-2
- Gunsett, F. C. (1984). Linear index selection to improve traits defined as ratios. *J. Anim. Sci.* 59, 1185–1193. doi:10.2527/jas1984.5951185x
- Hurley, A. M., López-Villalobos, N., McParland, S., Lewis, E., Kennedy, E., O'Donovan, M., et al. (2017). Genetics of alternative definitions of feed efficiency in grazing lactating dairy cows. *J. Dairy Sci.* 100, 5501–5514. doi:10.3168/jds.2016-12314
- Islam, M. S., Jensen, J., Løvendahl, P., Karlskov-Mortensen, P., and Shirali, M. (2020). Bayesian estimation of genetic variance and response to selection on linear or ratio traits of feed efficiency in dairy cattle. *J. Dairy Sci.* 103, 9150–9166. doi:10.3168/jds.2019-17137
- Karlsson, J., Danielsson, R., Åkerlind, M., and Holtenius, K. (2022). Full-lactation performance of multiparous dairy cows with differing residual feed intake. *PLoS One.* 17 (8), e0273420. doi:10.1371/journal.pone.0273420
- Kennedy, B. W., Vanderwerf, J. H. J., and Meuwissen, T. H. E. (1993). Genetic and statistical properties of residual feed-intake. *J. Anim. Sci.* 71, 3239–3250. doi:10.2527/1993.71123239x
- Khanal, P., Johnson, J., Gouveia, G., Ross, P., and Deeb, N. (2023). Genomic evaluation of feed efficiency in US holstein heifers. *J. Dairy Sci.* 106 (10), 6986–6994. doi:10.3168/jds.2023-23258
- Koch, R. M., Swiger, L. A., Chambers, D., and Gregory, K. E. (1963). Efficiency of feed use in beef cattle. *J. Anim. Sci.* 22, 486–494. doi:10.2527/jas1963.222486x
- Legarra, A., and Reverter, A. (2018). Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genet. Sel. Evol.* 50 (1), 53. doi:10.1186/s12711-018-0426-6
- Li, B., Berglund, B., Fikse, W. F., Lassen, J., Lidauer, M. H., Mäntysaari, P., et al. (2017). Neglect of lactation stage leads to naive assessment of residual feed intake in dairy cattle. *J. Dairy Sci.* 100, 9076–9084. doi:10.3168/jds.2017-12775
- Li, B., VanRaden, P. M., Guduk, E., O'Connell, J. R., Null, D. J., Connor, E. E., et al. (2020). Genomic prediction of residual feed intake in US holstein dairy cattle. *J. Dairy Sci.* 103 (3), 2477–2486. doi:10.3168/jds.2019-17332
- Lidauer, M. H., Negussie, E., Mäntysaari, E. A., Mäntysaari, P., Kajava, S., Kokkonen, T., et al. (2023). Estimating breeding values for feed efficiency in dairy cattle by regression on expected feed intake. *Animal* 17 (9), 100917. doi:10.1016/j.animal.2023.100917
- Liinamo, A.-E., Mäntysaari, P., Lidauer, M. H., and Mäntysaari, E. A. (2015). Genetic parameters for residual energy intake and energy conversion efficiency in nordic red dairy cattle. *Acta Agric. Scand. A Anim. Sci. Sect. A* 65 (2), 63–72. doi:10.1080/09064702.2015.1070897
- Lu, Y., Vandehaar, M. J., Spurlock, D. M., Weigel, K. A., Armentano, L. E., Staples, C. R., et al. (2015). An alternative approach to modeling genetic merit of feed efficiency in dairy cattle. *J. Dairy Sci.* 98 (9), 6535–6551. doi:10.3168/jds.2015-9414
- Luke (2017). Finnish feed tables. Available online at: [https://portal.mtt.fi/portal/page/portal/Rehutaulukot/feedtables\\_english](https://portal.mtt.fi/portal/page/portal/Rehutaulukot/feedtables_english) (Accessed February 10, 2023).
- Mäntysaari, P., and Mäntysaari, E. A. (2015). Modeling of daily body weights and body weight changes of nordic red cows. *J. Dairy Sci.* 98 (10), 6992–7002. doi:10.3168/jds.2015-9541
- Mäntysaari, P., Liinamo, A. E., and Mäntysaari, E. A. (2012). Energy efficiency and its relationship with milk, body, and intake traits and energy status among primiparous nordic red dairy cattle. *J. Dairy Sci.* 95 (6), 3200–3211. doi:10.3168/jds.2011-4685
- Matilainen, K., Mäntysaari, E. A., Lidauer, M. H., Strandén, I., and Thompson, R. (2012). Employing a Monte Carlo algorithm in expectation maximization restricted maximum likelihood estimation of the linear mixed model. *J. Anim. Breed. Genet.* 129, 457–468. doi:10.1111/j.1439-0388.2012.01000.x
- National Academies of Sciences, Engineering, and Medicine (2021). *Nutrient requirements of dairy cattle*. Eighth Revised Edition. Washington, DC: The National Academies Press.
- Negussie, E., de Haas, Y., Dehareng, F., Dewhurst, R. J., Dijkstra, J., Gengler, N., et al. (2017). Invited review: large-scale indirect measurements for enteric methane emissions in dairy cattle: a review of proxies and their potential for use in management and breeding decisions. *J. Dairy Sci.* 100, 2433–2453. doi:10.3168/jds.2016-12030
- NRC (2001). *Nutrient requirements of dairy cattle*. 7th Edn. National Academies Press.
- Oltenuacu, P. A., and Broom, D. M. (2010). The impact of genetic selection for increased milk yield on the welfare of dairy cows. *Anim. Welf.* 19, 39–49. doi:10.1017/S0962728600002220
- Pitkänen, T., Gao, H., Kudinov, A., Taskinen, M., Mäntysaari, E. A., Lidauer, M. H., et al. (2022). "From data to genomic breeding values with the MiX99 software suite," in *Proceedings, 12th world congress of genetics applied to livestock production* (Rotterdam, Netherlands), 1534–1537.
- Pryce, J. E., Royal, M. D., Garnsworthy, P. C., and Mao, I. L. (2004). Fertility in the high-producing dairy cow. *Livest. Prod. Sci.* 86, 125–135. doi:10.1016/S0301-6226(03)00145-3
- Pryce, J. E., Arias, J., Bowman, P. J., Davis, S. R., Macdonald, K. A., Waghorn, G. C., et al. (2012). Accuracy of genomic predictions of residual feed intake and 250 day bodyweight in growing heifers using 625,000 SNP markers. *J. Dairy Sci.* 95, 2108–2119. doi:10.3168/jds.2011-4628
- Pryce, J. E., Gonzalez-Recio, O., Nieuwhof, G., Wales, W. J., Coffey, M. P., Hayes, B. J., et al. (2015). Definition and implementation of a breeding value for feed efficiency in dairy cows. *J. Dairy Sci.* 98, 7340–7350. doi:10.3168/jds.2015-9621
- R Core Development Team (2012). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available online at: <http://www.R-project.org/> (Accessed February 11, 2025).
- Sjaunja, L. O., Baevre, L., Junkkarinen, L., Pedersen, J., and Setälä, J. (1990). "A nordic proposal for an energy corrected milk (ECM) formula," in *Proceedings of the 27th biennial session of the international committee of animal recording* (Paris, France), 156–157.
- 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. *J. Dairy Sci.* 95, 5393–5402. doi:10.3168/jds.2012-5407
- Stephansen, R. B., Martin, P., Gredler-Grandl, B., Manzanilla Pech, C. I. V., Sahana, G., Ducrocq, V., et al. (2022). "Novel genetic parameters to improve gRFI in dairy cattle using big data from multiple lactations and countries," in *Proceedings of the 12th World Congress on Genetics Applied to Livestock Production (WCGALP)*, 3–8 Jul 2022 (Rotterdam, Netherlands), 340–343.
- Strandén, I., and Vuori, K. (2006). "Relax2: pedigree analysis program," in *Proceedings of the 8th world congress on genetics applied to livestock production (WCGALP)*, 13–18 August 2006 (Brazil: Belo Horizonte), 27–30.
- Strandén, I., and Mäntysaari, E. A. (2018). *HGinv program*. Natural Resources Institute Finland (LUKE).
- Tempelman, R. J., Spurlock, D. M., Coffey, M., Veerkamp, R. F., Armentano, L. E., Weigel, K. A., et al. (2015). Heterogeneity in genetic and nongenetic variation and energy sink relationships for residual feed intake across research stations and countries. *J. Dairy Sci.* 98, 2013–2026. doi:10.3168/jds.2014.8510
- Vallimont, J. E., Dechow, C. D., Daubert, J. M., Dekleva, M. W., Blum, J. W., Barlieb, C. M., et al. (2011). Heritability of gross feed efficiency and associations with yield, intake, residual intake, body weight, and body condition score in 11 commercial Pennsylvania tie stalls. *J. Dairy Sci.* 94 (4), 2108–2113. doi:10.3168/jds.2010-3888
- VandeHaar, M. J., and St-Pierre, N. (2006). Major advances in nutrition: relevance to the sustainability of the dairy industry. *J. Dairy Sci.* 89, 1280–1291. doi:10.3168/jds.S0022-0302(06)72196-8
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi:10.3168/jds.2007-0980
- Wang, Y., Diao, C., Kang, H., Hao, W., Mrode, R., Chen, J., et al. (2022). A random regression model based on a single-step method for improving the genomic prediction accuracy of residual feed intake in pigs. *Front. Genet.* 12, 769849. doi:10.3389/fgene.2021.769849