



J. Dairy Sci. 102:1–11  
<https://doi.org/10.3168/jds.2018-15197>

© 2019, The Authors. Published by FASS Inc. and Elsevier Inc. on behalf of the American Dairy Science Association®.  
This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Linkage between predictive transmitting ability of a genetic index, potential milk production, and a dynamic model

E. Ruelle,<sup>1\*</sup> L. Delaby,<sup>2</sup> and L. Shalloo<sup>1</sup>

<sup>1</sup>Teagasc, Animal and Grassland Research and Innovation Centre, Moorepark, P61C996 Fermoy, Co. Cork, Ireland

<sup>2</sup>INRA-Agrocampus-Ouest, UMR 1348, Physiologie, Environnement et Génétique pour l'Animal et les Systèmes d'Élevage, Domaine de la Prise, 35590 Saint Gilles, France

### ABSTRACT

With the increased use of information and communication technology-based tools and devices across traditional desktop computers and smartphones, models and decision-support systems are becoming more accessible for farmers to improve the decision-making process at the farm level. However, despite the focus of research and industry providers to develop tools that are easy to adopt by the end user, milk-production prediction models require substantial parameterization information for accurate milk production simulations. For these models to be useful at an individual animal level, they require the potential milk yield of the individual animals (and possibly potential fat and protein yields) to be captured and parameterized within the model to allow accurate simulations of the interaction of the animal with the system. The focus of this study was to link 3 predicted transmitting ability (PTA) traits from the Economic Breeding Index (PTA for milk yield, fat, and protein) with potential index parameters for milk, fat, and protein required as inputs to a herd-based dynamic milk model. We compiled a data set of 1,904 lactations that included different experiments conducted at 2 closed sites during a 14-yr period (2003–2016). The treatments implied different stocking rates, concentrate supplementation levels, calving dates, and genetic potential. The first step, using 75% of the data randomly selected, was to link the milk, fat, and protein yields achieved within each lactation to their respective PTA value, stocking rate, parity, and concentrate supplementation level. The equations generated were transformed to correspond to inputs to the pasture-based herd dynamic milk model. The equations created were used in conjunction with the model to predict milk, fat, and protein production. Then, using the remaining 25% data of the data set, the simulations were compared against the actual milk

produced during the experiments. When the model was tested, it was capable of predicting the lactation milk, fat, and protein yield with a relative prediction error of <10% at the herd level and <13% at the individual animal level.

**Key words:** dairy cow, model, genetic index, link

### INTRODUCTION

With the abolition of the European Union's milk quota regimen and the increase in milk price volatility, farmers need to be more precise around their daily management decisions. To provide information, it is important to be able to predict, with precision, the effect of a management change on the animal, the herd, and the pasture in a pasture-based system. Management of dairy cows at pasture requires the execution of several different decisions on a daily basis, with many decisions being based on intuition rather than through the application of hard facts (Hanrahan et al., 2017). To avoid risk, key decisions tend to be based on past experience, which does not always lead to the most profitable or innovative solutions. Dynamic and mechanistic models can help farmers to predict the impact of a management decision on farm and thus minimize risks while increasing innovation. Decisions could be, for example, the effect of feeding silage in the early stage of a drought period or the actual benefit of feeding concentrate at different stages of lactation. The numbers of models developed are increasing internationally. Those models can have different goals, such as evaluating technology and management strategies on farm (DAFOSYM; Rotz et al., 1989, 1999), evaluating the environmental and economic sustainability of a farm (DairyWise; Schils et al., 2007), (Mélodie; Chardon et al., 2012), or predicting the biophysical and economic performance of grazing systems (e-Dairy; Baudracco et al., 2013). However, most of the published models are used only (if used at all) by researchers because of the complexity of the inputs required. If a model is to be used by a farmer or an advisor, the number of inputs must be limited and

Received June 12, 2018.

Accepted November 28, 2018.

\*Corresponding author: [elodie.ruelle@teagasc.ie](mailto:elodie.ruelle@teagasc.ie)

easily accessible on farm, and the outputs of the model must have strong utility to the end user.

When predicting milk production, simulation models generally have 2 options: either predict milk production based on previous milk records or predict milk production based on the genetic potential of the animal, which is independent of actual historical milk production (Baudracco et al., 2011; Ruelle et al., 2016). The first method is the easiest in terms of on-farm parameterization and is used in several models (Shalloo et al., 2004). However, that approach does not allow the utilization of those models as innovation tools to test changes in management in a dynamic way because the milk yields included are based on specific management of the farm at that particular time. An example might include changing grass allowance or feed supplementation levels, with the objective to determine the influence on the performance of individual animals. The second solution (potential milk yield linked to an arbitrary genetic potential) is used in other models, such as the pasture-based herd dynamic milk (**PBHD**M) model (Ruelle et al., 2015) or e-Dairy (Baudracco et al., 2013). This type of parameterization permits the model to complete simulations with the potential of the animal being independent of the feed levels offered, allowing the user to run simulations of different feeding levels for the same animal potential. Linking the actual genetic index of an animal to the input required in a model for milk production simulation would be the most efficient approach to parameterize the models. Furthermore, once a mechanistic model is used and a relationship between the genetic potential of an individual animal and the potential milk yield of that animal is included in the model, a whole range of different management options can be simulated in a realistic and accurate fashion, ultimately greatly increasing the utility of the model as a whole.

The objective of this paper was to develop a link between the genetic indices of an individual animal to the potential milk production of that animal to deploy within a dynamic mechanistic model. In this study, the genetic selection indices used were PTA of the Economic Breeding Index (**E**BI; Berry et al., 2007), and the model of interest was the **PBHD**M (Ruelle et al., 2015). The structure of this study would allow the same approach to be developed for other models and indices used in countries around the world.

## MATERIALS AND METHODS

The focus of this study was to develop equations to link the PTA milk (kg), fat (kg), and protein (kg) (from the **E**BI) of the animals with the inputs required by

the **PBHD**M model. This was completed in a 2-step process. The first step was to establish, using linear regression analysis, the link between the PTA (for kg of milk, fat, and protein), parity, stocking rate (**S**R), and concentrate supplementation level, and total actual milk, fat, and protein yields of the individual animals using 75% of a data set composed of 1,904 lactations. The second step involved creating a direct link between the total yields predicted in step 1 and the inputs of the model in terms of the development of a potential index for milk (**theoMYmax**), protein (**theoProt**), and fat (**theoFat**) production, which is deployed within the model.

For evaluation, using the **PBHD**M model on the remaining 25% of the data set, the actual yields produced by the animal were compared with the corresponding simulated yields predicted by the model.

## **E**BI

The **E**BI is a breeding index launched in Ireland in February 2001 to identify genetically superior animals to increase profitability within Irish dairy herds (Veerkamp et al., 2002). The **E**BI and its subindices are described in detail by Berry et al. (2007). The **E**BI is expressed as expected profit (€) per lactation of the progeny of the animal in question, and the genetic merit values of the component traits are, therefore, expressed as PTA values. At present, 7 subindices are included in the **E**BI (ICBF, 2017, 2018): production (32% contribution to the overall **E**BI value), fertility (35%), calving (10%), beef (8%), maintenance (7%), management (4%), and health (4%). The production subindex is made up of 3 PTA representing the milk (economic weight of -€0.09/kg; contribution of 9.9%), protein (economic weight of €5.58/kg; contribution of 18.6%), and fat (economic weight of €2.08/kg; contribution of 3.5%) yields (in kg). These 3 PTA are used in the rest of the study and are linked to the inputs of the model and are referred to herein as **PTA<sub>MYkg</sub>**, **PTA<sub>prot</sub>**, and **PTA<sub>fat</sub>**, respectively.

## **PBHD**M

The **PBHD**M model (Ruelle et al., 2015) is a dynamic, stochastic agent-based model developed in C++. The **PBHD**M model comprises the “herd dynamic milk” model (Ruelle et al., 2016) adapted for grazing and management conditions and the “Moorepark St Gilles grass growth” model (MoSt GG; Ruelle et al., 2018). The model simulates all the main aspects of the life of an animal from birth to culling and death through several different submodels. The submodels of the **PB**

HDM model include fertility, intake, animal growth, BCS change, and milk production. The model predicts the production of standard milk at 4.0% fat and 3.1% protein (Faverdin et al., 2011; Equation [1]). The actual milk produced is converted to standard milk for comparison purposes based on the following equation (Faverdin et al., 2011):

$$\text{Standard MY} = \frac{MY \times [0.44 + 0.0055 \times (FC - 40) + 0.0033 \times (PC - 31)]}{0.44}, \quad [1]$$

where  $MY$  is milk yield (kg),  $FC$  is the fat content of the milk (g/kg), and  $PC$  is the protein content of the milk (g/kg). The model also predicts the daily fat and protein concentration.

The simulation of milk production per day is calculated based on an interaction between energy and protein intake by the cow, BCS change, and the individual animal's potential milk yield. The potential milk yield is the milk that would be produced if the cow was in an energy balance equal to 0 without any change in body condition. Each cow has a pool of BCS that can be mobilized during lactation. This pool is calculated at calving and depends on the parity of the animal, her  $\text{theoMY}_{\text{max}}$ , and her BCS at calving (Delaby et al., 2010). If the energy intake available for milk allows a lower production than her potential, the cow will mobilize reserves (BCS loss if the BCS pool is not empty), which will allow her to produce more milk than possible through feed alone. If the energy intake available for milk allows a higher milk production than the cow's theoretical potential, part of this energy is used to increase the body reserve of the cow (BCS gain) and part will produce additional milk. The model simulates each individual animal's intake at grazing and depends on animal characteristics but also on supplementation and grass availability and quality, with a decrease in animal intake occurring during the defoliation process as the animal grazes through the grass profile. The grass growth in the model is predicted using the MoSt GG model (Ruelle et al., 2018), which simulates grass growth with a daily time step, taking into account soil water and soil N dynamics.

The 3 inputs of the model corresponding to the genetics of the animal are  $\text{theoMY}_{\text{max}}$ ,  $\text{theoProt}$ , and  $\text{theoFat}$ . The  $\text{theoMY}_{\text{max}}$  is the potential index of the animal for standard milk (which corresponds to the theoretical maximum standard milk production of the cow at her peak day if her demand could be matched by her intake), and  $\text{theoFat}$  and  $\text{theoProt}$  are the fat and protein potential indices, respectively, corresponding to

the lactation-average fat and protein percentages of the animal.

### Description of the Data Set

Data from 2 Teagasc experimental farms were collected and collated. The first included data collected from grazing experiments undertaken at Teagasc Curtins Farm, Animal and Grassland Research and Innovation center, Moorepark, Fermoy, Co. Cork, Ireland (52.17N, 8.27W) from 2003 to 2013. The data set is composed of 11 years of grazing experiments with a range of concentrate supplementation from 350 to 1,500 kg/cow per year and a range of SR from 2.47 to 3.92 cow/ha (Table 1). The full descriptions of the different experiments from which the data were sourced were published previously (Horan et al., 2005, McCarthy et al., 2007, 2014; Coleman et al., 2010). The second data set was from the Next Generation Herd, which was assembled in and available since 2012. Maiden heifers, in-calf heifers, and heifer calves were sourced from commercial dairy herds and from within Teagasc dairy herds. The herd is based at the Dairygold Research Farm in Kilworth, Co. Cork, Ireland (52.17N, 8.24W). There are 2 distinct EBI groups: 90 elite cows, corresponding to the top 5% EBI females in the country, and 45 national average EBI females. The animals were subjected to 2 different concentrate supplementations (300 and 1,000 kg per cow per lactation) and 2 different postgrazing height strategies (target of either 3.25 or 4.25 cm) at an SR of 2.75 cow/ha. Data from 2013 to 2016 were used in this study.

The data available for each animal lactation were daily milk yield, weekly fat and protein yields, EBI (and its subindex), parity, and their linkages to the different concentrate and SR groups. Descriptions of the different EBI traits of the animals in the data set are presented in Table 2. Data edits included the removal of lactations with >30% of milk records missing and lactations that were shorter than 200 d, resulting in 1,904 individual lactation records being available across the 2 studies (Table 1).

### Transformation of the Data

For each lactation, the total production of milk, fat, and protein was calculated by summing the daily recorded production within the experiments. In this study, lactation length was mainly affected by management, experiment type, and date of calving. To overcome lactation length effect, milk, fat, and protein yields were transformed to their 305-d equivalents (Bonaiti et al., 1990; IDELE, 1999):

$$MY305 = MY \times \frac{385}{\text{lactation length} + 80}; \quad [2]$$

$$\text{Protein305} = \text{Protein} \times \frac{MY305}{MY}; \quad [3]$$

$$\text{Fat305} = \text{Fat} \times \frac{MY305}{MY}. \quad [4]$$

### Statistical Analyses (Step 1)

Using the statistical linear model in R (R Core Team, 2017), we established relationships between the equivalent adult MY305 and  $PTA_{MYkg}$  (kg), Fat305 and  $PTA_{fat}$  (kg), and Protein305 and  $PTA_{prot}$  (kg), SR,  $SR^2$ , concentrate supplementation, year, and parity. Parity was grouped into parities 1, 2, 3, and 4+:

$$Y_i = \beta_0 + \beta_1 \times PTA_i + \beta_2 \times SR_i + \beta_3 \times SR_i^2 + \beta_4 \times C_i [5] \\ + \beta_5 \times year + \beta_6 \times parity + \varepsilon_i,$$

where  $Y_i$  corresponds to MY305, Fat305, or Protein305 (kg), as defined in equations [2], [3], and [4], for lactation  $i$ ;  $PTA_i$  is the corresponding  $PTA_{MYkg}$ ,  $PTA_{fat}$ , or  $PTA_{prot}$  (kg);  $SR_i$  and  $SR_i^2$  are stocking rate (cow/ha) and its square, respectively;  $C_i$  is concentrate supplementation (kg/lactation) of lactation  $i$ ;  $year$  is the year corresponding to the lactation; and  $parity$  is the parity of the animals (1, 2, 3, or 4+). The parameters  $\beta$  are the regression coefficients and  $\varepsilon_i$  is the error term. Data were randomly distributed into 2 subsets containing either 75 or 25% of the original data set. The 75% data set was used to create the statistical model and the

remaining 25% was used to validate the model outputs. The randomization was generated within treatment and subtreatment to ensure a homogeneous analysis of SR and concentrate supplementation in the 2 groups. The equations resulting from the linear regression are presented in Table 3.

### Calculation of the Potential Index for Milk, Fat, and Protein (Step 2)

Using the equation developed in step 1 (Table 3), the potential indices were developed using the following steps.

**Step 2.1.** The inputs to the model in terms of milk production correspond to the potential peak milk yield of a mature animal without any gain or loss of BCS (Ruelle et al., 2016). Consequently, the milk yield predicted by the regression has to be a potential milk yield only linked to the  $PTA_{MYkg}$ , which is why the SR, concentrate, year, and parity values in the equation had to be replaced by fixed values. The SR and concentrate parameters were generated based on an iterative process where the error in the model was minimized in the evaluation process using the 75% of the data that was used to create the model. The error was calculated when comparing the milk yield predicted by the PBHDM model to the actual milk yield in a fashion similar to that of the evaluation process. The range of values used was 1.5 to 3.0 cow/ha for SR and 200 to 2,000 kg/cow for concentrate. The values leading to the lowest error in the evaluation process were an SR of 2.3 cow/ha and a concentrate supplementation of 1,000 kg/cow. Parity was set to 4+ and the coefficient used for the year was the average of their 14 regression

**Table 1.** Repartition of animals across the different years, stocking rates (2.47 to 3.92 cows/ha), and concentrate supplementation (from 300 to 1,500 kg/cow per year)

Year	2.47		2.51	2.64	2.74	2.75		2.8	3.28	3.92	Total
	350	1,500	450	500	350	300	1,000	1,200	450	450	
2003	29	30			26						85
2004	41	42			41						124
2005	42	41			41						124
2006				65				60			125
2007				62				59			121
2008				66				65			131
2009			45						45	45	135
2010			46						45	45	136
2011			46						45	45	136
2012			45						45	45	135
2013			40			87	44		44	41	256
2014						88	44				132
2015						87	44				131
2016						88	45				133
Total	112	113	222	193	108	350	177	184	224	221	1,904

**Table 2.** Diversity of the milk, fat, and protein PTA of the animals used in the study from the Curtins Research Farm and the Next Generation Herd

Item	Curtins Research Farm					Next Generation Herd						
	PTA <sup>1</sup>		Herd production			PTA production trait (kg)			Herd production			
	PTA <sub>MYkg</sub> (kg)	PTA <sub>fat</sub> (kg)	PTA <sub>prot</sub> (kg)	Milk 305 <sup>2</sup> (kg)	Fat (%)	Protein (%)	PTA <sub>MYkg</sub> (kg)	PTA <sub>fat</sub> (kg)	PTA <sub>prot</sub> (kg)	Milk 305 (kg)	Fat (%)	Protein (%)
Mean	113	10.4	8.4	6,707	4.5	3.3	3	6.2	3.6	6,525	4.4	3.7
SD	156	4.6	3.9	1,028	0.5	0.4	146	4.8	3.9	953	0.4	0.2
Minimum	-384	-5.5	-2.5	4,232	3.4	2.4	-391	-8.7	-8.4	4,159	3.3	3.1
Maximum	476	31	18.5	10,626	6.3	4.4	422	18.5	13.9	9,311	5.7	4.2

<sup>1</sup>PTA<sub>MYkg</sub> = PTA for milk production; PTA<sub>fat</sub> = PTA for fat production; PTA<sub>prot</sub> = PTA for protein production.

<sup>2</sup>Milk 305 = recalculated milk production of the animal at 305 d using equation [2].

coefficients. Yields represent the potential yields of the mature animal without any effect of actual feeding level (equation [6]):

$$MY = 7,169 + 3.19 \times PTA_{MYkg}, \quad [6]$$

where 7,169 corresponds to

$$\begin{aligned} &7,744 \text{ (intersect)} - 1,673 \times 2.3 \text{ (SR)} + 237 \\ &\times 2.3^2 \text{ (SR}^2\text{)} + 0.82 \times 1,000 \text{ (concentrate)} \\ &+ 1,771 \text{ (parity)} - 572 \text{ (year)} \text{ (Table 3)}. \end{aligned}$$

**Step 2.2.** The inputs to the model in terms of protein and fat content correspond to the average percentage for the lactation; the protein and fat yields predicted through the regression were divided by the milk yield predicted through the regression to create the potential index theoFat and theoProt concentrations (equations [7] and [8]):

$$theoFat = \frac{2,972 + 35.2 \times PTA_{fat}}{MY} \times 100, \quad [7]$$

where 2,972 corresponds to

$$\begin{aligned} &3,192 \text{ (intersect)} - 774 \times 2.3 \text{ (SR)} + 114 \times 2.3^2 \text{ (SR}^2\text{)} \\ &+ 0.28 \times 1,000 \text{ (concentrate)} + 866 \text{ (parity)} \\ &- 189 \text{ (year)} \text{ (Table 3); and} \end{aligned}$$

$$theoProt = \frac{2,489 + 30.7 \times PTA_{prot}}{MY} \times 100, \quad [8]$$

where 2,489 corresponds to

$$\begin{aligned} &2,740 \text{ (intersect)} - 701 \times 2.3 \text{ (SR)} + 97 \times 2.3^2 \text{ (SR}^2\text{)} \\ &+ 0.32 \times 1,000 \text{ (concentrate)} + 643 \text{ (parity)} \\ &- 115 \text{ (year)} \text{ (Table 3)}. \end{aligned}$$

**Step 2.3.** The potential index theoMYmax corresponds to theoretical standard daily milk yield at peak for the animal; the milk yield calculated through the equation was first transformed into standard milk using equation [1] and theoFat and theoProt as the percentages of fat and protein, respectively. Then, because theoMYmax corresponds to a potential daily milk yield at peak of lactation, the standard theoretical milk yield calculated was divided by 232 (equation [9]; based on Faverdin et al., 2011), optimized during the iterative process determining the optimal SR and concentrate replacement value):

**Table 3.** Coefficients of the linear regression models for the prediction of 305-d milk, fat, and protein yields depending on the corresponding PTA, parity, stocking rate (SR), and concentrate supplementation (C)

Item	305-d milk yield (kg)	305-d fat yield (100 × kg)	305-d protein yield (100 × kg)
Intersect	7,744	3,192	2,740
PTA <sup>1</sup>	3.19	35.2	30.7
SR	-1,673	-774	-701
SR <sup>2</sup> (squared coefficient of SR)	237	114	97
C	0.82	0.28	0.32
Parity			
Parity 1	0	0	0
Parity 2	921	424	362
Parity 3	1,486	710	562
Parity 4+	1,771	866	643
Year <sup>2</sup>			
R <sup>2</sup>	0.83	0.80	0.82

<sup>1</sup>PTA from the calculation of the Economic Breeding Index (EBI); the PTA used in the regression are PTA<sub>MYkg</sub>, PTA<sub>fat</sub>, or PTA<sub>prot</sub> (PTA for milk, fat, or protein production for the milk yield, fat yield, or protein yield regression).

<sup>2</sup>Year was significant in the model with an average value of -572 (0 to -1,308; minimum and maximum) for milk yield, -189 (0 to -488) for fat yield, and -115 (+135 to 396) for protein yield, with 2003 as a reference.

$$\frac{\text{theoMYmax} = MY \times 0.44 + 0.0055 \times (\text{theoFat} - 40) + 0.0033 \times (\text{theoProt} - 31)}{0.44} \times \frac{1}{232}.$$

[9]

### Evaluation of Equations Created Using the PBHDM Model

The 25% of the data set not used in the creation of the statistical models was used to evaluate the simulation model. This means that each of the experimental conditions was replicated in the model, including all animal information, farm information, and climate information over the 14-yr period (2003–2016). For each animal, their theoMYmax, theoProt, and theoFat were calculated using equations [9], [8], and [7], respectively. Animals were grouped by concentrate supplementation, SR, and year, leading to 38 different scenarios. The weather data for each year, used in the grass growth submodel, was collected at Teagasc, Animal and Grassland Research and Innovation Centre, Moorepark, Fermoy, Co. Cork, approximately 1 km from Curtins and Kilworth farms. Stocking rate was recreated by adapting the size of the paddocks to the number of animals within each scenario (18 paddocks were used) being modeled, and concentrate supplementation was simulated to match the actual concentrate supplementation strategy. For simplification, BW of the animals at the start of the simulation was set at 550 kg and BCS at 3.25, which corresponds to the average input data for the different studies included in the analysis.

The quality of the concentrate was assumed to be 1.03 UFL (unité fourragère lait) and 120 PDI (protéine digestible dans l'intestin); the quality of the grass was set at 0.97 UFL, 100 PDI, and 1 FV (fill value; McCarthy et al., 2013). The UFL system corresponds to the French system for nutrition of dairy animals and is fully presented in Faverdin et al. (2011). Once the simulation was completed, the milk production of each animal within the model was transformed to a 305-d equivalent milk production using equation [2] to allow comparison with the 305-d milk production of the actual animals and to delete any lactation length effect.

### Statistic Evaluation

The statistical analysis used to evaluate the accuracy of the yields predicted by the model compared with the actual yields on 25% of the data set measured the mean square prediction error (**MSPE**), the relative prediction error (**RPE**), and the concordance correlation coefficient (**CCC**). The MSPE is the sum of 3 components; namely, the mean bias  $(M_m - P_m)^2$ , the line variation  $S_P^2(1 - b)^2$  (the deviation of the slope of the regression line of  $P$  regressed on  $M$ ), and the random variation about the line,  $S_M^2(1 - R^2)$  (Bibby and Toutenburg, 1977). Each is expressed as a proportion of the total MSPE (Bibby and Toutenburg, 1977):

$$\begin{aligned} MSPE &= \frac{\sum (M - P)^2}{n} \\ &= (M_m - P_m)^2 + S_P^2(1 - b)^2 + S_M^2(1 - R^2), \end{aligned} \quad [10]$$

where  $n$  is the number of measured ( $M$ ) and predicted ( $P$ ) pairs compared;  $M_m$  and  $P_m$  are the means of  $M$  and  $P$ , respectively;  $S_M^2$  and  $S_P^2$  are the variances of  $M$  and  $P$ , respectively;  $b$  is the slope of the line of  $P$  regressed on  $M$ ; and  $R^2$  is the coefficient of determination of the line (Rook et al., 1990).

A high mean bias indicates that the predicted values are consistently higher or lower than the measured values. A low line component indicates that the difference of the slope of the regression is low and a good prediction slope was found. The root mean square prediction error (**RMSPE**) is the root of the MSPE. The RPE is calculated as (Fuentes-Pila et al., 1996):

$$RPE = \left( \frac{RMSPE}{M_m} \right) \times 100. \quad [11]$$

The CCC (Lin, 1989; Nickerson, 1997) evaluates the correlation between 2 data sets but also the deviation from the 45° line. The CCC is composed of 2 components:

$$CCC = \rho \times C_b, \quad [12]$$

where  $\rho$  is the Pearson correlation coefficient and  $C_b$  is the bias correction factor:

$$C_b = \frac{2 \times \sigma_A \times \sigma_P}{\sigma_A^2 + \sigma_P^2 + (\mu_A - \mu_P)^2}, \quad [13]$$

and  $\sigma_A$ ,  $\sigma_P$ ,  $\mu_A$ , and  $\mu_P$  are the standard deviation ( $\sigma$ ) and average ( $\mu$ ) of the actual and predicted data, respectively.

The Pearson correlation coefficient ( $\rho$ ) evaluates how far each observation deviated from the best-fit line, whereas the  $C_b$  evaluates the deviation from the 45° line ( $0 < C_b \leq 1$ ;  $C_b = 1$  if no deviation). The strength of agreement is considered poor if the CCC is  $< 0.65$ , moderate if between 0.65 and 0.80, substantial if between 0.80 and 0.90, and almost perfect if  $> 0.90$  (McBride, 2005).

The different analyses were completed at the individual animal level and at the herd level. At the herd level, all of the milk production parameters were compared across the 38 scenarios.

## RESULTS AND DISCUSSION

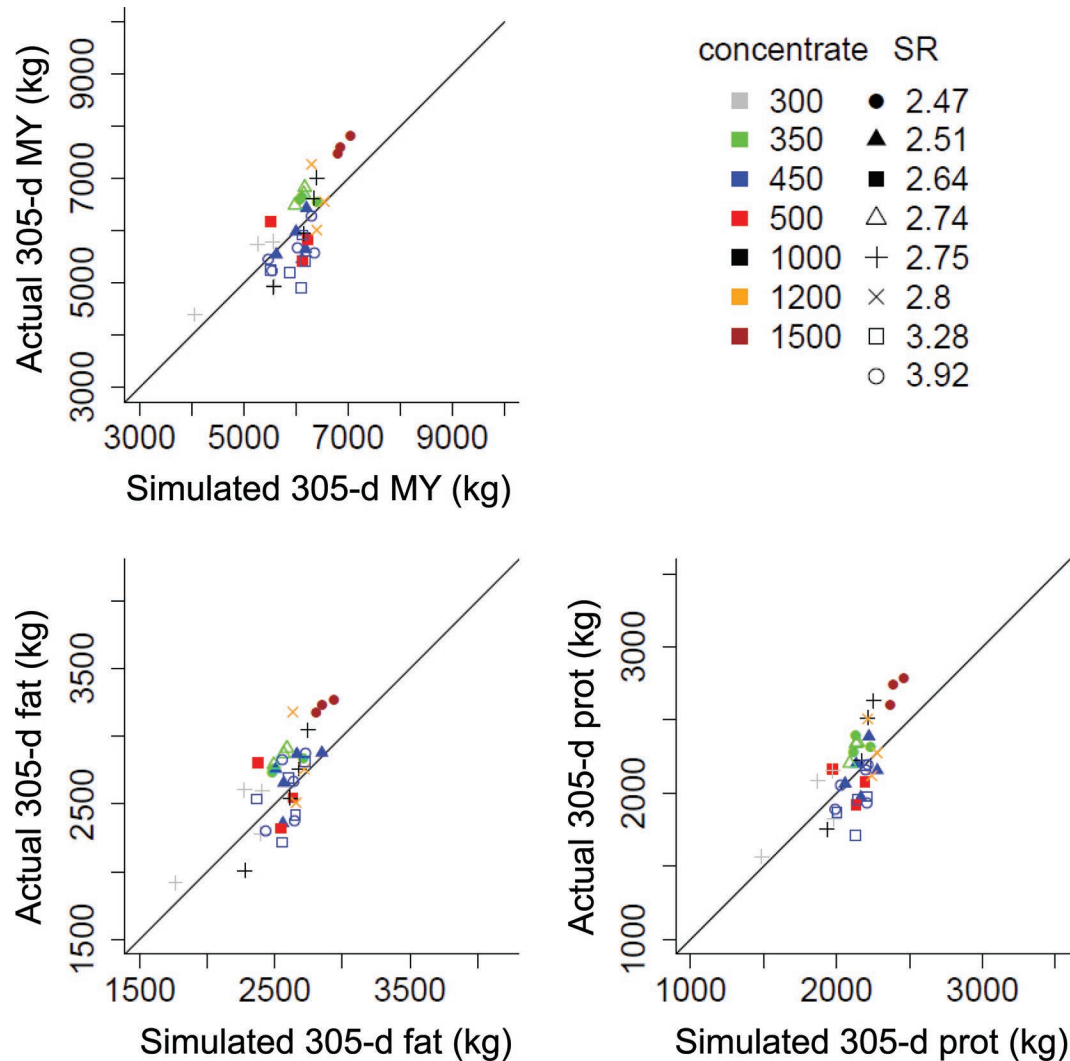
### Creating Links Between PTA and Inputs of the Model

The result of the regression between the 305-d milk, fat, and protein yields and their respective PTA cor-

rected by feeding level (trough SR and concentrate supplementation) led to coefficients of determination ( $R^2$ ) of 0.83, 0.80, and 0.82, respectively (Table 3). The PTA consistently showed a statistically significant effect on the milk traits in their respective equations, demonstrating a direct link between the PTA values and the actual production of the animals. The regression coefficients for the prediction of the milk yield for SR,  $SR^2$ , and concentrate were  $-1,673$  kg/cow per unit of SR,  $237$  kg/cow per unit of  $SR^2$ , and  $0.82$  kg/cow per kg of concentrate. The effect of a change in SR from 2 to 3 cow/ha (when  $PTA_{MY} = 83$ ) on predicted milk resulted in an decrease of 5.6 to 8.7% in milk yield per cow, which is similar to the 8.7 and 7.4% decrease (for different base SR), respectively, found by McCarthy et al. (2011) in a meta-analysis on a database of 109 experiments. The predicted increase of 0.82 kg of milk per kg of concentrate was also in the range of previously published research (Bargo et al. (2002; 0.76–1.1), Delaby et al. (2003; 0.6–1.57), Horan et al. (2005; 0.5–1.18), and Roche et al. (2006; 0.23–1.06).

### Evaluation of the Relationships

Results of using the newly developed equations as inputs to the model on 25% of the remaining data set and running the simulations in terms of comparing cumulative simulated and actual yearly yield are presented in Figure 1 and Table 4 at the herd level and in Figure 2 and Table 4 at the animal level. Overall, for all variables analyzed (milk, protein, and fat), RPE was  $< 10\%$  at the herd level and  $< 13\%$  at the individual animal level, showing satisfactory prediction according to the scale of Fuentes-Pila et al. (1996; Table 4). Accuracy was always higher at the herd level than at the animal level, with accuracy being close for each variable. For every simulation, a very high proportion of the MSPE ( $> 80\%$ ) was attributed to random variation, which shows the quality of the prediction (Table 4). The simulation showed average RPE of 9.1 and 12.7% and CCC of 0.62 and 0.68 at the herd and animal level, respectively. The accuracy of these simulations suggests that the equations developed in this paper could be used to link EBI and inputs of the model without the need for extra information (e.g., actual SR or concentrate supplementation). The precision of the model at the herd level is a major asset for on-farm use because the fat, protein, and milk yields of the herd are more useful than being able to accurately simulate individual cows within a pasture-based system, where all animals are managed as one group. However, the fact that the model is still accurate (all RPE  $< 13\%$ ) at the animal level allows users to study the impact of different individual animal management regimens and could allow



**Figure 1.** Comparison between actual and simulated 305-d yields of milk (MY), fat, and protein (prot) averaged by scenario [concentrate from 300 to 1,500; stocking rate (SR) from 2.47 to 3.92].

significant future research simulations of individual animals under different management regimens. For example, what would be the consequence of feeding extra

concentrate to the higher or lower genetic potential individual animals (e.g., feed to yield), while integrating the economic and fertility components associated with

**Table 4.** Comparison between the actual and predicted 305-d milk, fat, and protein yields (kg/lactation) averaged for the different simulations (herd level) and at the animal level<sup>1</sup>

Item	Yield	Measured	Predicted	RMSPE	Proportion of the MSPE					
					Mean	Line	Random	RPE	CCC	Cbias
Herd level	Milk 305	6,034	6,032	543	0.0	1.7	98.3	9.0	0.67	0.91
	Fat 305	2,677	2,572	246	18.2	0.3	81.5	9.2	0.59	0.84
	Prot 305	2,172	2,136	200	3.2	2.5	94.3	9.2	0.62	0.87
Animal level	Milk 305	5,933	5,944	743	0.0	2.2	97.8	12.5	0.72	0.93
	Fat 305	2,636	2,536	340	8.5	2.9	88.6	12.9	0.66	0.88
	Prot 305	2,137	2,106	269	1.3	4.9	93.8	12.6	0.67	0.89

<sup>1</sup>RMSPE = root mean square prediction error; MSPE = mean square prediction error; RPE = relative predicted error; CCC = concordance correlation coefficient; Cbias = bias of the concordance correlation coefficient.

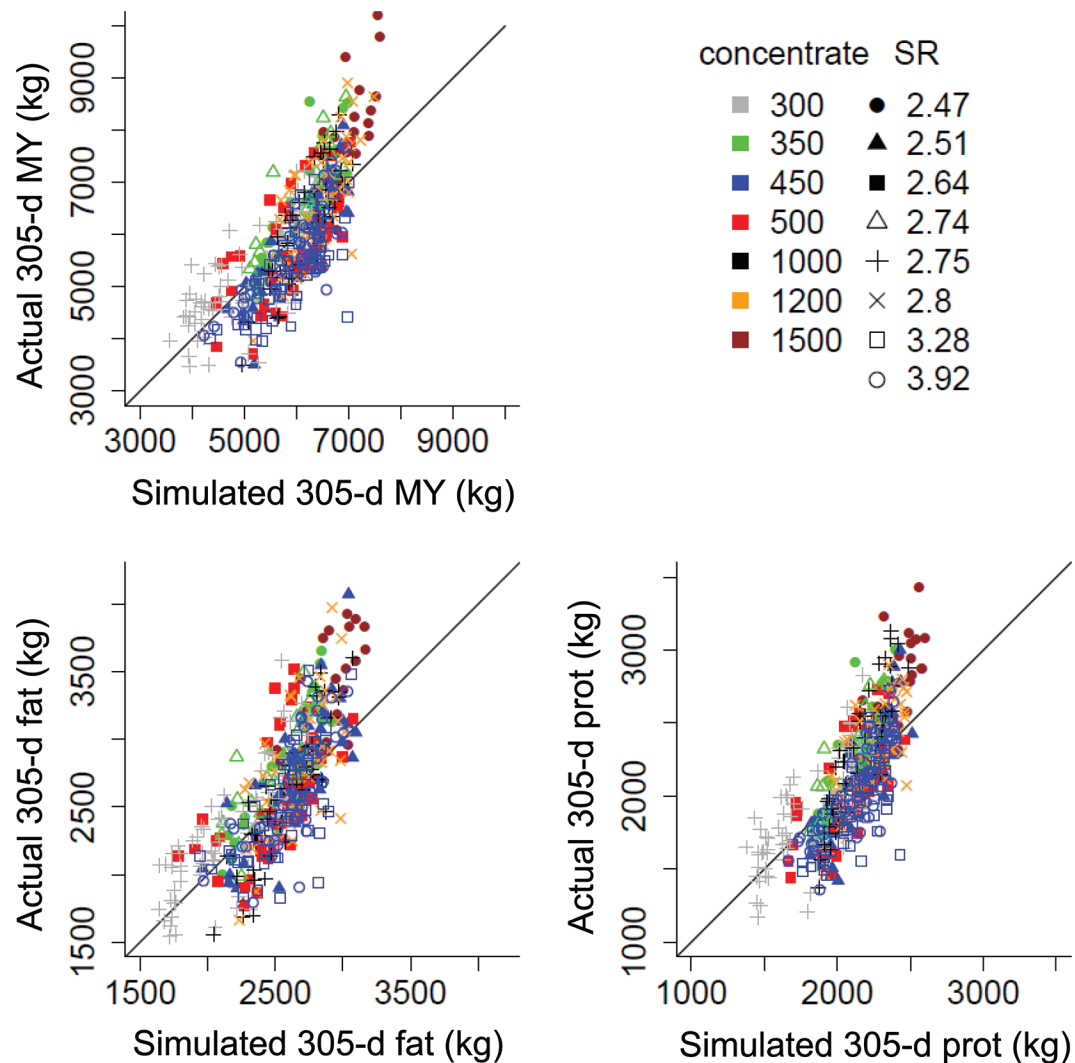


the feed levels at individual animal levels. Although the prediction was very accurate for the lower-producing animals, the model had a tendency to underestimate the higher levels of milk production. Those levels of production were linked to high levels of concentrate supplementation (>1,000 kg per lactation). This underestimation could be due to the simplification assumptions made in term of inputs of the model around the BW of the animal, BCS at calving, concentrate quality, and repartition of the concentrate supplementation. Another hypothesis is that the PBHDM is underestimating the milk production of the animal at high levels of concentrate. However, the model's response to concentrate has been tested previously, and the model is capable of reproducing cow production of >10,000 kg of

milk (Ruelle et al., 2016), which makes this hypothesis unlikely. Another evaluation of the model will have to be conducted to test this hypothesis.

### Limits of this Approach

The equations can be used for an indoor system because no inputs around SR are needed; however, the accuracy of the equations would have to be tested in that scenario, as none of the data used in this study came from indoor systems. A data set composed of data from indoor herds would have to be created and an equation linking inputs to model the genetic index and concentrate supplementation would have to be developed and tested if the same relationships were to be



**Figure 2.** Comparison between the actual and simulated 305-d yields of milk (MY), fat, and protein (prot) at the animal level [concentrate from 300 to 1,500; stocking rate (SR) from 2.47 to 3.92].

used in that scenario. However, from a pasture-based systems perspective, this limitation would have a minor effect on the utilization of the model.

### Future Use of this Approach

The PasturebaseIreland (**PBI**) grassland decision-support system (Hanrahan et al., 2017) is a farm management tool to improve the decision-making process at the farm level, and it includes a back-end database. A long-term objective of PBI is to link the grass grown on farm to animal intakes and subsequently to milk, fat, and protein yields through the PBHDM model. The PBI website would be linked directly with the Irish Cattle Breeding Federation (Bandon, Co. Cork, Ireland; <https://www.icbf.com/wp/>), which would result in direct availability of genetic information of each animal for the model without requiring input by the farmer. Before each calving season, for each farm, an individual specific algorithm would be generated automatically following this approach, depending on the previous year's records of milk production and management of the herd, including the SR and concentrate supplementation of the previous year (as well as the genetic information), if available. Similarly, each year, the overall equation would be automatically regenerated using records from up to the 10 previous years over all farms within PBI. This will allow the PBI platform to provide information about the nutrition of the animal and to help grazing management, thus increasing confidence and management at the farm level. One of the uses could be to examine the effect of the addition of concentrate feeds to the diet at specific times during the lactation. The effect would vary depending on the genetics of the animals. The linkage of these models with an information and communications technology-based solution such as PBI has the potential to revolutionize pasture-based farming in the future.

### CONCLUSIONS

Equations were developed to permit the utilization of PTA for milk, fat, and protein as inputs into the PBHDM model, which increase its usefulness and usability. This will allow use of the model at the farm level through a link to the Irish Cattle Breeding Federation national database and will facilitate the development of individual animal- and herd-level milk, fat, and protein yield predictions.

### ACKNOWLEDGMENTS

The authors acknowledge the funding from the Research Stimulus Fund 2011 administered by the De-

partment of Agriculture, Fisheries and Food (Dublin, Ireland; Project 11/S/132); Frank Buckley (Teagasc, Moorepark, Ireland) for providing the data of the Next Generation Herd experiment; Brendan Horan (Teagasc, Moorepark, Ireland) for providing the data for the Curtins herd; and Anne Geoghegan (Teagasc, Moorepark, Ireland) for the extraction of the datasets.

### REFERENCES

- Bargo, F., L. D. Muller, J. E. Delahoy, and T. W. Cassidy. 2002. Milk response to concentrate supplementation of high-producing dairy cows grazing at two pasture allowances. *J. Dairy Sci.* 85:1777–1792. [https://doi.org/10.3168/jds.S0022-0302\(02\)74252-5](https://doi.org/10.3168/jds.S0022-0302(02)74252-5).
- Baudracco, J., N. Lopez-Villalobos, C. W. Holmes, E. A. Comeron, K. A. Macdonald, and T. N. Barry. 2013. e-Dairy: A dynamic and stochastic whole-farm model that predicts biophysical and economic performance of grazing dairy systems. *Animal* 7:870–878. <https://doi.org/10.1017/S1751731112002376>.
- Baudracco, J., N. Lopez-Villalobos, L. A. Romero, D. Scandolo, M. Maciel, E. A. Comeron, C. W. Holmes, and T. N. Barry. 2011. Effects of stocking rate on pasture production, milk production and reproduction of supplemented crossbred Holstein–Jersey dairy cows grazing lucerne pasture. *Anim. Feed Sci. Technol.* 168:131–143. <https://doi.org/10.1016/j.anifeedsci.2011.03.017>.
- Berry, D., L. Shalloo, A. Cromie, R. Veerkamp, P. Dillon, P. Amer, J. Kearney, R. Evans, and B. Wickham. 2007. Pages 29–34 in *The Economic Breeding Index: A generation on*. Irish Cattle Breeding Federation, Bandon, Co. Cork, Ireland.
- Bibby, J., and H. Toutenburg. 1977. *Prediction and Improved Estimation in Linear Models*. Wiley, New York, NY.
- Bonaiti, B., D. Boichard, E. Verrier, V. Ducrocq, A. Barbat, and M. Briend. 1990. La méthode française d'évaluation génétique des reproducteurs laitiers. *Prod. Anim.* 3:83–92.
- Chardon, X., C. Rigolot, C. Baratte, S. Espagnol, C. Raison, R. Martin-Clouaire, J. Rellier, A. Le Gall, J. Dourmad, and B. Piquemal. 2012. MELODIE: A whole-farm model to study the dynamics of nutrients in dairy and pig farms with crops. *Animal* 6:1711–1721.
- Coleman, J., D. P. Berry, K. M. Pierce, A. Brennan, and B. Horan. 2010. Dry matter intake and feed efficiency profiles of 3 genotypes of Holstein–Friesian within pasture-based systems of milk production. *J. Dairy Sci.* 93:4318–4331. <https://doi.org/10.3168/jds.2009-2686>.
- Delaby, L., S. Leurent, Y. Gallard, and T. Schmitt. 2010. Effet de la race, de la parité, du potentiel laitier et de l'état au vêlage sur l'évolution de l'état corporel des vaches laitières au cours de la lactation in *Proc. 17èmes Rencontres Recherches Ruminants*. Institut de l'Élevage, INRA, Paris, France.
- Delaby, L., J.-L. Peyraud, N. Foucher, and G. Michel. 2003. The effect of two contrasting grazing managements and level of concentrate supplementation on the performance of grazing dairy cows. *Anim. Res.* 52:437–460.
- Faverdin, P., C. Baratte, R. Delagarde, and J. L. Peyraud. 2011. GrazeIn: A model of herbage intake and milk production for grazing dairy cows. 1. Prediction of intake capacity, voluntary intake and milk production during lactation. *Grass Forage Sci.* 66:29–44. <https://doi.org/10.1111/j.1365-2494.2010.00776.x>.
- Fuentes-Pila, J., M. A. DeLorenzo, D. K. Beede, C. R. Staples, and J. B. Holter. 1996. Evaluation of equations based on animal factors to predict intake of lactating Holstein cows. *J. Dairy Sci.* 79:1562–1571. [https://doi.org/10.3168/jds.S0022-0302\(96\)76518-9](https://doi.org/10.3168/jds.S0022-0302(96)76518-9).
- Hanrahan, L., A. Geoghegan, M. O'Donovan, V. Griffith, E. Ruelle, M. Wallace, and L. Shalloo. 2017. PastureBase Ireland: A grassland decision support system and national database. *Comput. Electron. Agric.* 136:193–201. <https://doi.org/10.1016/j.compag.2017.01.029>.
- Horan, B., P. Dillon, P. Faverdin, L. Delaby, F. Buckley, and M. Rath. 2005. The interaction of strain of Holstein–Friesian cows and pas-

- ture-based feed systems on milk yield, body weight, and body condition score. *J. Dairy Sci.* 88:1231–1243. [https://doi.org/10.3168/jds.S0022-0302\(05\)72790-9](https://doi.org/10.3168/jds.S0022-0302(05)72790-9).
- ICBF. 2017. What is EBI? Accessed Jan. 12, 2018. <https://www.icbf.com/wp/?p=5772>.
- ICBF. 2018. Update to Economic Values in the EBI. Accessed Jan. 12, 2018. <https://www.icbf.com/wp/?p=9695>.
- IDELE. 1999. Méthode d'indexation laitière: Les principes. [http://idele.fr/?eID=cmis\\_download&oID=workspace://SpacesStore/62a9c85f-a125-4d42-b245-9a6599a60a4c](http://idele.fr/?eID=cmis_download&oID=workspace://SpacesStore/62a9c85f-a125-4d42-b245-9a6599a60a4c).
- Lin, L. I. K. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45:255–268.
- McBride, G. B. 2005. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. NIWA Client Report: HAM2005–062. National Institute of Water and Atmospheric Research Ltd., Hamilton, New Zealand.
- McCarthy, B., L. Delaby, K. Pierce, F. Journot, and B. Horan. 2011. Meta-analysis of the impact of stocking rate on the productivity of pasture-based milk production systems. *Animal* 5:784–794.
- McCarthy, B., L. Delaby, K. M. Pierce, A. Brennan, and B. Horan. 2013. The effect of stocking rate and calving date on milk production of Holstein–Friesian dairy cows. *Livest. Sci.* 153:123–134. <https://doi.org/10.1016/j.livsci.2013.01.013>.
- McCarthy, J., B. McCarthy, B. Horan, K. Pierce, N. Galvin, A. Brennan, and L. Delaby. 2014. Effect of stocking rate and calving date on dry matter intake, milk production, body weight, and body condition score in spring-calving, grass-fed dairy cows. *J. Dairy Sci.* 97:1693–1706.
- McCarthy, S., D. P. Berry, P. Dillon, M. Rath, and B. Horan. 2007. Influence of Holstein-Friesian strain and feed system on body weight and body condition score lactation profiles. *J. Dairy Sci.* 90:1859–1869. <https://doi.org/10.3168/jds.2006-501>.
- Nickerson, C. A. E. 1997. A note on “A concordance correlation coefficient to evaluate reproducibility”. *Biometrics* 53:1503–1507.
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Roche, J. R., D. P. Berry, and E. S. Kolver. 2006. Holstein-Friesian strain and feed effects on milk production, body weight, and body condition score profiles in grazing dairy cows. *J. Dairy Sci.* 89:3532–3543. [https://doi.org/10.3168/jds.S0022-0302\(06\)72393-1](https://doi.org/10.3168/jds.S0022-0302(06)72393-1).
- Rook, A. J., M. S. Dhanoa, and M. Gill. 1990. Prediction of the voluntary intake of grass silages by beef cattle: 3. Precision of alternative prediction models. *Anim. Sci.* 50:455–466. <https://doi.org/10.1017/S0003356100004931>.
- Rotz, C. A., D. R. Buckmaster, D. R. Mertens, and J. R. Black. 1989. DAFOSYM: A dairy forage system model for evaluating alternatives in forage conservation. *J. Dairy Sci.* 72:3050–3063.
- Rotz, C. A., D. R. Mertens, D. R. Buckmaster, M. S. Allen, and J. H. Harrison. 1999. A dairy herd model for use in whole farm simulations. *J. Dairy Sci.* 82:2826–2840. [https://doi.org/10.3168/jds.S0022-0302\(99\)75541-4](https://doi.org/10.3168/jds.S0022-0302(99)75541-4).
- Ruelle, E., L. Delaby, M. Wallace, and L. Shalloo. 2016. Development and evaluation of the herd dynamic milk model with focus on the individual cow component. *Animal* 10:1986–1997.
- Ruelle, E., D. Hennessy, and L. Delaby. 2018. Development of the Moorepark St Gilles grass growth model (MoSt GG model): A predictive model for grass growth for pasture based systems. *Eur. J. Agron.* 99:80–91.
- Ruelle, E., L. Shalloo, M. Wallace, and L. Delaby. 2015. Development and evaluation of the pasture-based herd dynamic milk (PBHDM) model for dairy systems. *Eur. J. Agron.* 71:106–114. <https://doi.org/10.1016/j.eja.2015.09.003>.
- Schils, R. L. M., M. H. A. de Haan, J. G. A. Hemmer, A. van den Polvan Dasselaar, J. A. de Boer, A. G. Evers, G. Holshof, J. C. van Middelkoop, and R. L. G. Zom. 2007. DairyWise, a whole-farm dairy model. *J. Dairy Sci.* 90:5334–5346. <https://doi.org/10.3168/jds.2006-842>.
- Shalloo, L., P. Dillon, M. Rath, and M. Wallace. 2004. Description and validation of the Moorepark Dairy System Model. *J. Dairy Sci.* 87:1945–1959. [https://doi.org/10.3168/jds.S0022-0302\(04\)73353-6](https://doi.org/10.3168/jds.S0022-0302(04)73353-6).
- Veerkamp, R. F., P. Dillon, E. Kelly, A. R. Cromie, and A. F. Groen. 2002. Dairy cattle breeding objectives combining yield, survival and calving interval for pasture-based systems in Ireland under different milk quota scenarios. *Livest. Prod. Sci.* 76:137–151. [https://doi.org/10.1016/S0301-6226\(02\)00006-4](https://doi.org/10.1016/S0301-6226(02)00006-4).