French experience of using ICAR approved method for predicting 24-hour fat percentage and yield from one sampled milking in Automatic Milking Systems

X. Bourrigan[1], R. Vallée[1] and G. Augier[2]

[1]Institut de l'Elevage, 149, rue de Bercy, 75595 Paris, France
[2]Eliance, 149, rue de Bercy, 75595 Paris, France
Corresponding Author: xavier.bourrigan@idele.fr

Abstract

The use of Automatic Milking Systems is increasing in France, from about 1,550 herds with official records in 2010 to about 3,350 in 2022. The Dairy Cattle Milk Recording Guidelines allows 2 types of robot protocols approved by ICAR, one with at least two milkings per recording test day sampled for components and, to meet the simplification and cost-saving needs of farmers, another with only one milking sampled per recording test day. Regarding the latter, which is used in France since 2017 (Minéry *et al.,* 2018), 24-hour fat percentage and yield are estimated with the ICAR Peeters and Galesloot method (Peeters and Galesloot, 2002).

The aim of this study was to try to improve the prediction of the 24-hour fat percentage and yield by using more complex models described by Peeters and Galesloot. Therefore, we compared the accuracy (r^2, prediction error and standard deviation of prediction error) of the predictions at a recording test day level for the multiple regression model currently used and for 6 other models also considering the effect of class variables, such as milking interval and fat to protein ratio.

The estimation of regression coefficients and the validation studies were performed on independent updated data sets (with at least two milkings sampled by cow), using a total of 620,272 milkings for 125,905 cows spread over 1,277 French farms from 2017 to 2019.

The results confirmed the relevance of the model currently used but highlighted a possible improvement. Indeed, adding the effect of class variables to the prediction model slightly improved the correlation between the 24-hour reference and the 24-hour prediction for fat percentage and yield, from 0.776 to 0.786 and from 0.910 to 0.913 for fat percentage and fat yield respectively, for the model giving the best results. There was no effect on the prediction error (0.0003% for fat percentage and 0 kg for fat yield) while the standard deviation of the prediction error was slightly reduced, from 0.308 to 0.301% and from 98 to 96g for fat percentage and fat yield respectively.

Keywords: milk recording, automatic milking systems, 24-hour fat percentage, prediction, accuracy.

Introduction

The number of dairy farms in France using Automatic Milking Systems (AMS) in Official Milk Recording increased significantly from 1550 in 2010 to 3550 in 2022 and this number of is relatively stable since 2019. Today the percentage of AMS farms in Official Milk Recording represents 12% of the total of farms in Official Milk Recording and 14% of cows in Official Milk Recording. The average number of cows by farm is equal to 85 (more than 15 cows compared with traditional farms) and 55% of farms are fitted with one AMS box.

This growth creates difficulties for the Milk Recording Organizations (MRO): cost of milk recording, use of sampling equipment…

On the published literature, several studies have been made to answer some of these issues and challenges for Milk Recording with (Bouloc, 2001; Peeters and Galesloot, 2002; Hand *et al.,* 2006; Leclerc *et al.,* 2012, Bourrigan *et al.,* 2013).

To answer needs and expectations of MRO's and AMS farmers, the French Milk Recording Guidelines proposes 2 types of AMS schemes approved by ICAR:

- at least two sampled milkings per recording test day by MRO's technician=AR scheme or by farmer BR scheme (this is the Gold Standard for genetic evaluation),
- one-sampled milking per recording test day, used since 2017 with a specific identification: AR* scheme or BR* scheme by using Peeters and Galesloot's method (defined in Section 2 of current ICAR Guidelines, 2022) for predicting 24-hour fat% and yield. Regarding this AMS milk recording scheme, weighting factors have been defined for genetic evaluation (from the determination coefficient r^2 and the repetability level of fat% and fat yield traits) and applied since 2020 (Vallee and al, 2021).

For helping technicians, farmers during Milk Recording test day (set up Automatic Milking Samplers, parameters, data transfer,…), today 25 different AMS Standard Operating Procedures (SOP) have been described in the French Milk Recording Guidelines in collaboration with all AMS Manufacturers.

The aims of this study carried out in 2022, consisted to:

- to check and verify the accuracy level of the Peeters and Galesloot's first regression coefficients (defined in 2017) from a new relevant dataset,
- to improve the actual 24-hour performances predicted, by using more complex Peeters and Galesloot's models (7 different models tested) and described in Journal of Dairy Science article (2002),
- to calculate accuracy results on recording test day / 24-hour reference,
  - to evaluate the accuracy of the method on recording test day,
  - to propose potentially changes of the French Milk Recording Guidelines, according to the results achieved.

Material and methods

Presentation of the Peeters and Galesloot models tested

The Peeters and Galesloot method is a multiple linear regression declined in different models. The "simple" model allows to estimate 24-hour fat percentage and yield from one-sampled milking by taking into account fat and protein percentage, milk weight and milking interval of the sampled milking and milk weight and milking interval of the previous milking (Table 1). Six more "complex" models (models Ca to Cf), similar to the simple one, also include different classifications of variables such as the time of day of the sampled milking, the parity number, the stage of lactation, the interval preceding the sampled milking and/or the fat% to protein% ratio (Table 2). (Peeters and Galesloot, 2002; ICAR Guidelines., 2022)

Table 1. Peeters and Galesloot's simple model.

24-hour Fat% = $b_0 + b_1$* Fat%(n) + b_2* Prot%(n) + b_3* MI(n) + b_4* MI(n-1) + b_5* Milk(n) + b_6* Milk(n-1) + e
b_0 = intercept, b_1 to b_6 = regression coefficients
MI = milking interval, Milk = milk weight, e = residual effect
(n) = milking sampled, (n-1) = previous milking

Table 2. Peeters and Galesloot's complex models.

24-hour Fat%i = b_0i + b_1i * Fat%(n) + b_2i * Prot%(n) + b_3i * MI(n) + b_4i * MI(n-1) + b_5i * Milk(n) + b_6i * Milk(n-1) + ei
b_0i = intercept, b_1i to b_6i = regression coefficients
i = subclass of classification for class variables Cx for x = a, b, c, d, e, f
Ca = day time of sampled milking (h) 0-5.59, 6.00-11.59, 12.00-17.59, 18.00-23.59
Cb = interval preceding the sampled milking n (min) 0-360, 361-510, 511-700, 701-1440
Cc = fat to protein% ratio of the sampled milking 0-1.10, 1.10-1.25, 1.25-1.40, >1.40
Cd = parity 1, 2, $\geq$ 3
Ce = lactation stage 1-99, 100-199, $\geq$200
Cf = interval preceding the sampled milking n (min) 0-360, 361-510, 511-700, 701-1440 and fat to protein % ratio of the sampled milking 0-1.10, 1.10-1.25, 1.25-1.40, >1.40

Description of the datasets used in the study

Milkings collected by Milk Recording Organizations over the years 2017 to 2019 in herds where at least two milkings per cow were sampled were used to establish a 24-hour reference population. Data from breeds other than Holstein (71% of the milkings), Montbeliarde (24%), Normande, Simmental and Brown Swiss were deleted due to insufficient numbers, as well as milkings with an milking interval lower than 4 hours, a milk yield lower than 1 kg or higher than 30 kg, a sampling period lower than 12 hours and outlier fat (less than 1.5% and more than 9%) and protein percentage (less than 1% and more than 7%).

Thus a total of 620,272 milkings (described in Table 3) were taken into account for 125,905 cows spread over 1,277 French farms.

The reference dataset was then split into two independent sets, a training data set of 414,394 milkings, used to estimate regression coefficients for the 7 different Peeters and Galesloot models and a validation data set of 205,878 milkings, for which 24-hour fat percentage and yield were predicted. The study consisted in comparing the predicted performances to the 24-hour reference ones by analyzing the accuracy of the predicted values (r^2, prediction error/bias and standard deviation of prediction error/bias).

Results

Compared to the performances of the simple model, the analysis of the accuracy of the different complex models shows (Table 4):

- a reduction of standard deviation of bias from 0.001% (model Cd) to 0.007% (model Cf) for fat percentage and from 0.42g (model Cd) to 1.83g (model Cf) for fat yield,

Table 3. Description of the reference data set.

	Sampled milking				Preceding milking		24-hour reference		
	MY (kg)	Fat% (%)	Prot% (%)	MI (minutes)	MY (kg)	MI (minutes)	MY (kg)	Fat% (%)	Prot% (%)
Mean	11.8	3.99	3.32	587.2	11.4	558.6	30.1	4.02	3.31
Std	3.9	0.78	0.37	169.8	3.7	163.2	8.8	0.65	0.36
Min	1.0	1.50	1.01	60	1.0	60	2.4	1.51	1.64
Max	30.0	8.99	6.83	1440	30.0	1440	73.1	8.85	6.80

Table 4. Correlations (r^2), bias and standard deviation of bias between predicted 24-hour fat% and yield and reference 24-hour fat% and yield.

	Fat% (%)			Fat yield (g)		
Prediction model	Bias	Std bias	r^2	Bias	Std bias	r^2
Uncorrected	-0.034	0.422	0.706	-11	135.96	0.849
Simple model	0.0003	0.308	0.776	-0.11	97.97	0.910
Complex model Ca	0.0002	0.306	0.779	-0.16	97.36	0.911
Complex model Cb	0.0003	0.304	0.781	-0.31	97.03	0.912
Complex model Cc	0.0003	0.305	0.781	-0.07	97.20	0.911
Complex model Cd	0.0003	0.307	0.778	-0.05	97.55	0.911
Complex model Ce	0.0002	0.306	0.778	-0.11	97.48	0.911
Complex model Cf	0.0002	0.301	0.786	-0.30	96.14	0.913

- an improvement of correlations (r^2) from +0.2 (models Cd and Ce) to +1.0 (model Cf) point for fat percentage and from +0.01 (models Ca, Cc, Cd and Ce) to +0.3 (model Cf) point for fat yield.

Regardless of the prediction model, a tendency to overestimate fat percentage and underestimate fat yield can be observed, but with a negligible bias overall (less than 0.0003% for fat percentage and less than 0.31g for fat yield).

Discussion and conclusion

Regarding the accuracy of the Peeters and Galesloot's regression coefficients used in France since 2017 with the simple model, there no difference between the regression coefficients tested during this study in 2022 (overall same level of r^2, or fat% and fay yield). But it's necessary to check regularly the accuracy level of the regression coefficients (every 4 or 5 years from a new dataset, according to changes of performance, herd management,…).

Regarding the accuracy of 6 Peeters and Galesloot's complex models tested during this study on milk recording test day, the gain of accuracy r^2 (in comparison with the current simple model) is equal to 1.0 point for fat% and 0.3 point for fat yield with a complex model (Cf) which combines milking interval and fat/protein% ratio.

A tendency to overestimate fat% and underestimate fat yield can be observed with Cf complex model but overall with a negligible bias and a reduction of standard deviation of bias to 0.007% for fat% and to 1.83g for fat yield.

After six years of using an ICAR approved method for predicting 24-hour fat% and yield from one-sampled milking in Automatic Milking Systems, the Peeters and Galesloot method's is widely use today by French Milk Recording Organizations. The goal is to answer the expectations of the farmers to simplify schemes and to reduce the cost of AMS milk recording test day while maintaining a sufficient accuracy for genetic evaluation and cow management purpose.

This new French study about the possibility of improving the accuracy level of the Peeters and Galesloot's method shows that from a new regression formula, adding of milking interval and fat/protein ratio class variable, a gain of accuracy is observed especially for fat%.

Another study (Roelofs *et al.,* 2006) showed that the Peeters and Galesloot's method regression formula was improved to estimate the 24-hour fat% based on one-sampled milking, especially by adding other variable such as month of sampling (pasture effect),..

References

Bouloc, N., 2001. Milk recording and automatic milking systems, characteristics and possibilities for simplification of milk recording. Institut de l'Elevage, France, 14 p.

Bourrigan, X., 2013. Study about the possibility of changes robots protocols in milk recording - Report n°001371045 - Institut de l'Elevage, France, 30 p.

Hand, K.J., Lazenby, D., Miglior, F., Kelton, D.F., 2006. Comparison of Protocols to Estimate 24-hour Fat and Protein Percentages for Herds with a Robotic Milking System. Journal of Dairy Science. 89:1723-1726.

ICAR Guidelines., 2022. Procedure 1 of Section 2 - Cattle Milk Recording - Computing 24-hour yields - Estimation of fat and protein yield (Galesloot and Peeters, 2000).

Leclerc, H., Huquet, B., Minery, S., Bourrigan, X., Thomas, G., Saunier, D., 2012. Analysis of the accuracy of protocols in robotic milking herds for estimation of 24-hour fat and protein (percentage and yields) - Proceedings of the 38th Biennal Session of ICAR, Cork, Ireland.

Minery, S., Bourrigan, X., Saunier, D., Augier, G., 2018. Analysis of the accuracy of a method for estimating 24-hour fat (percentage and yields) with robots protocols and one single sample.

Roelofs, R.M.G., G. de Jong and A.P.W. de Roos. 2006. Renewed estimation method for 24-hour fat % in AM/PM milk recording scheme. NRS, P.O. Box 454, 6800 AL Arnhem, The Netherlands. Proceedings of the 35th ICAR conference in Kuopio, Finland.

Vallee, R., Bourrigan, X., Augier, G., 2021. Analysis of the accuracy of lactation qualification methods and use of weighting factors for genetic evaluation.