Application of lactation models on dairy cow farms

Miroslav Janković[1]**, Ana Leko**[1] **and Nenad Šuvak**[2,*]

[1]*Farmeron Ltd., Stjepana Radića 16, 31000 Osijek, Croatia*
E-mail: ⟨*miroslav.jankovic, ana.leko@farmeron.com*⟩

[2]*Department of Mathematics, J. J. Strossmayer University of Osijek, Trg Ljudevita Gaja 6, 31000 Osijek, Croatia*
E-mail: ⟨*nsuvak@mathos.hr*⟩

Abstract. An important branch of the modern global economy is food production. Our focus is milk production on dairy farms, the period in which cows produce milk (lactation) and the quantity of milk produced. Special attention will be given to the time-dependent function that describes the quantity of the milk produced over a lactation period. The respective graph is known as the lactation curve and is one of the most important indicators in dairy farm management.
This paper will present time-dependent parametric models for daily average milk production on a particular dairy farm. Besides the well-known Wood's model, the data will also be fitted to some less known models such as the MilkBot model and also to a combination of these two models. Model parameters will be interpreted relative to milk production. Finally, we will compare all of the observed models.

Keywords: lactation, milk yield, Wood's model, MilkBot model, Levenberg-Marquart algorithm

1. Introduction

Lactation is the period in which a cow produces milk. It starts with calving and usually lasts for 305 days. Lactation is followed by a period in which the cow does not produce milk and during which it rests and prepares for the next calving and lactation. This period usually lasts for 60 days, and it is called a dry period. Under the milk yield we consider the amount of milk produced during lactation. We will distinguish between the daily and the overall milk yield. The daily milk yield is the milk yield obtained on a certain day of lactation, while the overall milk yield refers to the total amount of milk obtained for a certain lactation period. The daily milk yield over time is usually described by a lactation curve, the graph of a time-dependent parametric function. This function defines the so-called lactation model.

*Corresponding author.

Lactation models provide insight into the milk yield of each cow, enabling farmers to compare how the specific cow behaves in relation to a group of cows and how the performance of the specific cow meets their expectations. Additionally, for cows living in similar conditions, lactation models are used to describe their average milk yield.

The Wood's model is a widely-used lactation model in many countries (reflecting the differences in food, living conditions, animal health care, etc.) and is applicable to different dairy animals (see e.g. [4, 6] and [5]). The popularity of this model has led to its modification and improvement over time (see e.g. [6]). Besides the Wood's model, there are many other lactation models - here we will point out the Milkbot model (see e.g. [7]) and Wilmink model (see e.g. [5]), also utilizing daily average milk yield data in [8] and [9]. A comparison of different lactation models for a variety of milk yield data is available in Abas et al. [1], Aziz et.al. [2], Babnik et al. [3] and Bilgin et al. [4].

In this paper, three lactation models will be described: Wood's (Section 3), MilkBot (Section 4) and the new lactation model defined in terms of the previous two (Section 5). Based on daily lactation data from more than 400 cows located on a farm within a portfolio of the Croatian company Farmeron Ltd., we obtained daily average milk yields and focused on yields in the third or higher lactation only. Based on these data, we estimated the parameters of each model and compared them to the maximum milk yield, the time in which it is reached and the overall milk yield during lactation (Subsection 5.2). The objective criterion for comparing these three models relied on using the adjusted determination coefficient, denoted by R^2_{adj} (Subsection 5.2).

2. Lactation models

In this paper, three lactation models will be used to fit the daily average milk yields. Since the observed dairy cows are bred on the same farm, we can presume that they have similar living and health care conditions, which makes our sample homogeneous and justifies the modeling of the average milk yield rather that the milk yield of each cow. The final sample of the average milk yields consists of 305 data, which is also the expected duration of one lactation cycle. In building the models, i.e. estimating their parameters, 275 data will be used, while the remaining 30 data are left for the inspection of the model quality.

First, two classical lactation models will be described and used for data fitting - the Wood's model and MilkBot model. After observing the corresponding lactation curves and some objective indicators for model quality, e.g. adjusted determination coefficient (R^2_{adj}) (see e.g. [8]), the new model will be defined as a combination of the previous two.

Each of these models, besides modeling of the daily milk yield, enables computation of the theoretical maximum milk yield (Y_{max}), the moment in which it is reached (t_{max}) and the theoretical overall milk yield during the selected part of the lactation cycle (e.g. M_{275} and M_{305} for the first 275 and 305 days, respectively). In order to compare some of these characteristics calculated according to a specific model, in Table 1 we outlined their real values for the first, the second and the third or higher lactation, although in this paper we focus on the third or higher lactation only. All

characteristics regarding the quantity of milk are given in pounds (*lb*).

Lactation	t_{max}	Y_{max}	M_{275}	M_{305}
First	97	86.28	21525.86	23685.5
Second	58	109.33	25409.94	27452.89
Third or higher	55	122.27	28185.62	30343.94

Table 1: *Real values of characteristics t_{max}, Y_{max}, M_{275} and M_{305} for the first, the second and the third or higher lactation.*

Table 1 shows that the third or higher lactation is characterized by the highest overall yield and the highest maximal yield which is also reached much earlier than in the first and the second lactation.

3. Wood's lactation model

In this section, we introduce Wood's lactation model and use it to fit the previously described daily average milk yield data for cows in the third or higher lactation.

3.1. Model function and interpretation of parameters

In Wood's model, the milk yield in time $t > 0$ is described by the time-dependent function

$$Y_t = at^b e^{-ct}, \tag{1}$$

depending on parameters a, b and c. These parameters are interpreted as follows: the value of parameter $a > 0$ reflects the milk yield at the beginning of the observed lactation, parameter $b > 0$ controls the increasing part and parameter $c > 0$ controls the decreasing part of the lactation function (1).

In this model, the theoretical maximum milk yield Y_{max} and the time t_{max} in which it is reached as well as the theoretical overall milk yield M_u during u consecutive days of the lactation cycle are defined in Table 2.

t_{max}	Y_{max}	M_u
b/c	$a\left(b/c\right)^b e^{-b}$	$a\int_1^u t^b e^{-ct}\, dt$

Table 2: *Wood's model: maximum milk yield Y_{max}, the time t_{max} in which it is reached and the theoretical overall milk yield M_u during u consecutive days of the lactation cycle.*

To obtain the values of these characteristics, based on the daily average milk yield data dealt with, we need to estimate parameters a, b and c.

3.2. Methodology for parameter estimation

The same methodology is used for parameter estimation in all models used for fitting the daily average milk yields in this paper. In this subsection, we provide a few informative sentences about the parameter estimation method and the corresponding algorithm that will be used.

For parameter estimation we use the least-squares approach. Computationally, this method is usually performed by the Levenberg-Marquardt algorithm, the classical tool in solving the non-linear least-squares problems. This algorithm performs iterative improvements of the parameter values in order to minimize the sum of squares of residuals, i.e. the differences between the observed (real) and the theoretical (predicted) values. This algorithm is the "combination" of two methods - the method of the steepest descent and the famous Gauss-Newton method. For further details on these methods we refer to Madsen et al. [10].

3.3. Parameter estimation for Wood's lactation model

Parameters $a > 0$, $b > 0$ and $c > 0$ are estimated by the method described in Subsection 3.2, based on the daily average milk yield data in the third or higher lactation. The estimates, together with the standard error of the estimate and the p-value of the t-test for testing the null-hypothesis on the zero value of the parameter, are given in Table 3.

Parameter	Estimate	Standard error	p-value (t-test)
a	60.24992	0.78639	$< 2 \cdot 10^{-16}$
b	0.22001	0.00391	$< 2 \cdot 10^{-16}$
c	0.0036	0.000046	$< 2 \cdot 10^{-16}$

Table 3: *Parameter estimates for the Wood's model*

The p-value in Table 3, in case of all three parameters, implies that on the significance level 0.05, we reject the null-hypothesis and state that all three parameters are significantly larger than zero, i.e. all three parameters are significant for the shape of the Wood's lactation curve. The corresponding lactation curve together with the actual (observed) daily average milk yields is shown in Figure 1.
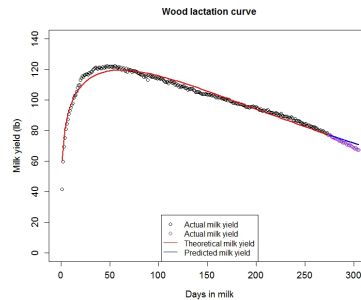

Figure 1: *Wood's lactation curve and the actual daily average milk yields*

The most obvious deviation of the estimated lactation curve from the actual data is during the period of actual and theoretical (model based) maximum milk yields - the model overestimates the time in which the maximum milk yield is reached. For clarification purposes, Table 3 shows the theoretical maximum milk yield Y_{max} and the time t_{max} in which it is reached, as well as the theoretical total milk yield

for both 275 and 305 days (M_{275} and M_{305}, respectively) for the third or higher lactation, all based on the Wood's model (1) with the estimated parameters given in Table 3.

	t_{max}	Y_{max}	M_{275}	M_{305}
Actual values	55	122.27	28185.62	30343.94
Wood's model values	61.1	119.5	28119.06	30334.1

Table 4: *Wood's model: Actual and theoretical maximum milk yield, the time in which it is reached and the theoretical total milk yields for both 275 and 305 days.*

4. MilkBot lactation model

4.1. Model function and interpretation of parameters

The time-dependent MilkBot model function is defined by the following expression:

$$Y_t = a \cdot \left(1 - \frac{1}{2}e^{\frac{c-t}{b}}\right) \cdot e^{-dt}. \tag{2}$$

The function (2) depends on four parameters and therefore this model is expected to be more sensitive than Wood's model. Parameter a is the scaling parameter and its role in the model is associated with the daily milk yield (usually, a is interpreted in lb/day). The other three parameters of the model are shape parameters. Parameter $b > 0$ controls the increasing part of the lactation curve and it is usually interpreted in terms of time between calving and reaching the theoretical maximum milk yield. Parameter c controls the time between calving and the beginning of lactation and it can be positive, negative or zero, with the latter reflecting the fact that lactation begins at the exact moment of calving. Parameter $d > 0$ controls the decreasing part of the lactation curve and therefore it is associated with lactation persistency, i.e. the ability of the cow to maintain production at a high level after the maximal yield.

The theoretical maximum milk yield Y_{max}, the time t_{max} in which it is reached and the theoretical overall milk yield M_u during u days of the lactation cycle in the framework of this model are given in Table 5 (where $k = 2bd/(1 + bd)$).

t_{max}	Y_{max}	M_u
$c - b \cdot ln\left(\frac{2bd}{1+bd}\right)$	$ak^{bd}\left(1 - \frac{1}{2}k\right)e^{-dc}$	$a \int_1^u \left(\left(1 - \frac{exp(\frac{c-t}{b})}{2}\right) \cdot e^{-dt}\right)dt$

Table 5: *MilkBot model: maximum milk yield Y_{max} and the time t_{max} in which it is reached and the theoretical overall milk yield M_u during u days of the lactation cycle.*

To obtain the values of these characteristics based on daily average milk yield data that we dealt with, we need to estimate model parameters a, b, c and d.

4.2. Parameter estimation for the MilkBot lactation model

Estimation of all parameters is based on the 275 daily average milk yields from the third or higher lactation, and it is performed using the method from the Subsection

3.2. Parameter estimates, together with the standard errors and the p-value of the t-test for testing the null-hypothesis on the zero value of the parameter, are given in Table 6.

Parameter	Estimate	Standard error	p-value (t-test)
a	137.25193	0.38086	$< 2 \cdot 10^{-16}$
b	14.07275	0.28626	$< 2 \cdot 10^{-16}$
c	3.26698	0.12284	$< 2 \cdot 10^{-16}$
d	0.00191	0.000017	$< 2 \cdot 10^{-16}$

Table 6: *Parameter estimates for the MilkBot lactation model.*

According to p-values in Table 6, all parameters are significant for the shape of the MilkBot lactation curve which is, together with the actual daily average milk yields, shown in Figure 2.
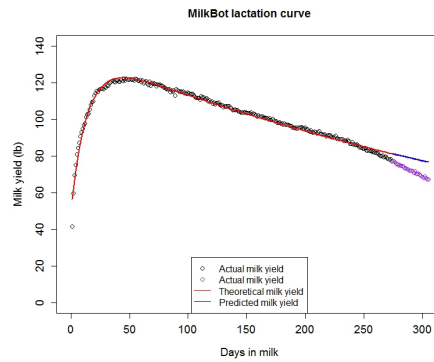

Figure 2: *MilkBot lactation curve and the observed daily average milk yields*

The lactation curve in Figure 2 shows the significant deviation from the observed data only at the end of the lactation cycle. Table 7 shows the theoretical maximum milk yield Y_{max}, the time t_{max} in which it is reached and the theoretical total milk yield for both 275 and 305 days (M_{275} and M_{305}, respectively) for the third or higher lactation for the MilkBot model (2) with estimated parameters from the Table 6.

	t_{max}	Y_{max}	M_{275}	M_{305}
Actual values	55	122.27	28185.62	30343.94
MilkBot model values	44.77	122.7	28118.8	30485

Table 7: *MilkBot model: Actual and theoretical maximum milk yield, the time in which it is reached and the theoretical total milk yield for* 275 *and* 305 *days.*

Here, unlike the Wood's model, the theoretical maximum milk yield is slightly over-estimated. Furthermore, the increase in the theoretical total milk yield (with respect to the actual value) if the additional 30 data are included in this analysis can be explained by the overestimation of the actual milk yield at the end of the lactation.

Remark 1. *Parameter estimates of the Wood's and MilkBot model for the first and*

the second lactation as well as the corresponding maximum milk yield, the time in which it is reached and the theoretical total milk yield for 275 *and* 305 *days can be found in [8].*

5. The new lactation model based on the Wood's and the MilkBot model

A comparison of the Wood's and MilkBot lactation curves for the third or higher lactation shows that the MilkBot model fits the data better at the beginning of the lactation cycle, whereas the Wood's model performs better at the end of the lactation period. This was the motivation for building the new model by "combining" the MilkBot model at the beginning and Wood's model at the end of the lactation cycle. Clearly, the purpose of this new model is to reduce underestimation of the actual average milk production at the beginning of lactation in case of Wood's model and overestimation of milk production at the end of the lactation period in case of the MilkBot model.

The idea behind the new model is based on finding the intersections of lactation curves of these two models and choosing the appropriate interval at the beginning of the lactation cycle where the milk yield is described by the MilkBot model and the interval at the end of the lactation where Wood's model is used. The criteria for choosing these time-intervals will be explained in the next section. The parameters of the new model will be estimated independently as the MilkBot parameters and Wood's model parameters, based on the daily average milk yields in the intervals where the new model behaves like the MilkBot and Wood's model, respectively. This approach allowed us to avoid the issue of nondifferentiability in the new model function in relation to the used estimation method.

Note that in Wood's model (1) parameter b controls the increasing part of the lactation function. Since in this setting Wood's model will be used to describe the milk production at the end of the lactation cycle, where the lactation function decays, we will introduce a logical simplification - the value of parameter b in the new model will be fixed to zero. Therefore, for the case of Wood's model, only parameters a and c are estimated. For the case of the MilkBot model, all four parameters are estimated.

5.1. Model function and interpretation of parameters

The new model describes the milk yield by the following function:

$$Y_t = a_1 \left(1 - \frac{1}{2} e^{\frac{c_1 - t}{b_1}} \right) \cdot e^{-d_1 t} \cdot I_{[1,T]} + a_2 e^{-c_2 t} \cdot I_{<T,305]}, \tag{3}$$

where the time T, as observed at the beginning of Section 5, is chosen as the maximum of all times in which the MilkBot model and Wood's model take the same value. In other words, T is the x-coordinate of the last intersection of MilkBot and Wood's lactation curves. The choice of time T in model (3) in an optimal sense is subject to further research.

Interpretation of parameters a_1, b_1, c_1 and d_1 from the MilkBot part and parameters a_2 and c_2 from Wood's part of the new model remain the same as in classical models. Furthermore, since time in which the actual maximum milk yield is reached belongs to the time-interval over which the milk production is described by the MilkBot model, both the theoretical maximum milk yield and the time in which it is reached are calculated according to the expressions given in Table 5.

5.2. Parameter estimation for new lactation model

The MilkBot and Wood's lactation curves for the third or higher lactation shown in Figures 2 and 1, respectively, have three intersections and the times in which they occur are 13.89, 72.29 and 207.7. The MilkBot model fits the daily average milk yield data better up to the time of the last intersection, after which Wood's model is more appropriate for fitting the observed data. Therefore, according to the previously described criteria, we choose $T = 207.7$. This rather arbitrary choice is supported by observing the sum of squares of residuals for both models at intervals $[0, 207.7]$ and $(207.7, 275]$ and the sum of squares of prediction errors over the interval $(275, 305]$ (see Table 8).

Model	Sum of sq. of res. [0, 207.7]	Sum of sq. of res. (207.7, 275]	Sum of sq. of pred. err. (275, 305]
Wood's	1553.19	197.73	126.17
MilkBot	510.14	189.94	1476.82

Table 8: *Sum of squares of residuals and sum of squares of prediction errors on appropriate intervals for the MilkBot and Wood's model.*

Table 8 shows that the MilkBot model better fits the actual daily average milk yields over the time-interval $[0, 275]$, which contains all the real data used to estimate model parameters. However, since Wood's model fits the milk yields from the interval $(275, 305]$ better than the MilkBot model, and since the sum of squares of residuals for both models in $(207.7, 275]$ are close, we follow the described criteria and use the MilkBot model over the interval $[0, 207.7]$ and Wood's model afterwards. Parameter estimates and standard errors for this new model are given in Table 9.

Model	Parameter	Estimate	Standard error
MilkBot	a_1	134.91403	0.42278
[0, 207.7]	b_1	12.99814	0.26627
	c_1	3.13863	0.10877
	d_1	0.00176	0.000024
Wood's	a_2	167.88602	1.97005
(207.7, 275]	c_2	0.00277	0.000049

Table 9: *Parameter estimates and corresponding standard errors for the lactation model* (3).

The lactation curve for the model (3) as well as the MilkBot and Wood's lactation curves estimated in subsections 3.3 and 4.2 are shown in Figure 3.
By comparing the new lactation curve with the MilkBot and Wood's lactation curves we can assume that the new model is an improvement. Significant deviations are
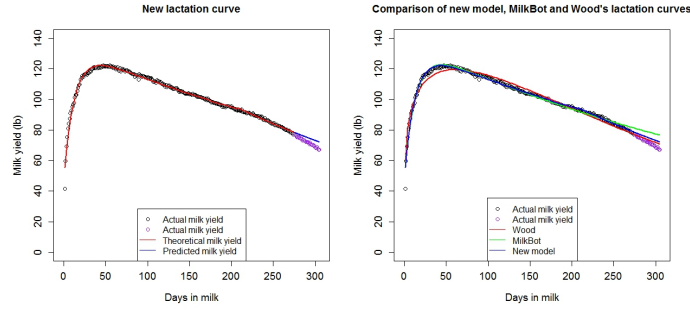
Figure 3: *Comparison of the new lactation curve with the MilkBot and Wood's lactation curves.*

present only in case of predictions (referring to the time period $(275, 305]$ which was not included in parameter estimation). The values predicted by the new model show greater deviations from actual values than in Wood's model, but significantly smaller deviations that in the MilkBot model. A more exact comparison of the new model with the classical ones is given in Table 10, by observing the theoretical maximum milk yield Y_{max} and the time t_{max} in which it is reached for both 275 and 305 days of the lactation cycle.

Model	t_{max}	Y_{max}	M_{275}	M_{305}
Actual values	55	122.27	28185.62	30343.94
MilkBot model values	44.77	122.7	28118.8	30485
Wood's model values	61.1	119.5	28119.06	30334.1
New model values	43.52	122.17	28120	30374.26

Table 10: *Theoretical maximum milk yield, the time at which it is reached and the theoretical overall yield for both 275 and 305 days of the lactation cycle for the MilkBot, Wood's and new lactation model.*

Unlike the MilkBot model that slightly overestimates the maximum milk yield, the new model slightly underestimates it. However, in comparison to both classical models, this underestimation (0.1 lb) is the smallest deviation from the actual maximum. The underestimation of the overall milk yield during the 275 lactation days is also the smallest for the case of the new model. For 305 lactation days, the new model overestimates the overall milk yield, but slightly less than the MilkBot model. However, the actual overall milk yield is closest to the value estimated by the Wood's model, hence for the purpose of estimating this characteristics, the new model is not the best choice.

In order to compare the new model with the classical ones, Table 11 shows the adjusted determination coefficients (R^2_{adj}) and the corresponding residual standard errors (RSD) for all three models, as the exact measures of the model quality.

As is evident in Table 11, when comparing with the MilkBot and Wood's model, the new model has the highest value of the adjusted determination coefficient R^2_{adj} and the smallest residual standard error. This means that for the observed set of daily average milk yields for cows in the third or higher lactation, the new model performs better than the classical ones. However, we want to emphasize that this

Model	R^2_{adj}	RSD
Wood's	0.9673	2.537
MilkBot	0.9869	1.607
New	0.9913	1.311

Table 11: *Adjusted R^2_{adj} and the residual standard errors for the MilkBot, Wood's and new model.*

conclusion is relevant for daily average milk yield data at the observed farm only. The appropriateness of the model for fitting the milk yields based on other farms should be further tested.

6. Conclusion

Wood's and the MilkBot lactation models are a usual choice for modeling daily average milk production data. However, as is evident in this paper, both models exhibit the time-intervals of a poor fit, i.e. underestimation or overestimation of real production data, depending on the specific period of lactation. In the new model, we tried to choose the appropriate interval at the beginning of the lactation cycle for the milk yields described by the MilkBot model and the interval at the end of lactation cycle when the Wood's model fits the milk yields better. By choosing time T as described in Subsection 5.1, the new model shows the highest value of the adjusted determination coefficient and the smallest residual standard error, i.e. the new model presents an improvement over the observed classical models. However, the appropriateness of the new model for fitting the daily milk yields based on other data sets should be taken with care, since milk production depends on animal health, milking hygiene, the right nutrition, animal welfare, environment, management and many other factors that vary from farm to farm. Also, care should be taken when using the predefined models on different data, in terms of possible overfitting of the model. Specifically, the model should not only fit the current sample, but also new samples, too. In that case, the new model provides good predictions, which is important for farmers to make suitable, statistically-based decisions at the appropriate time.

Acknowledgement

References

[1] Abas, Z., Giannakopoulos, C., Karagneli, M., Koutroumanidis, T. and Malesios, C. (2011). Comparison of models for describing the lactation curves of

Chios sheep using daily records obtained from an automatic milking system, DOI: 10.13140/2.1.2703.4885, Conference HAICTA, Skiathos Island, Greece

[2] Aziz, M.A., El-Shafie, O.M., Mahdy, A.T., Nishida, A. and Shalaby, N.A. (2006). Comparison between the shapes of lactation curve of Egyptian buffalo milk yield estimated by the incomplete gamma function and a new model, Livestock Research for Rural Development, 18(5)

[3] Babnik, D., Jeretina J. and Škorjanc, D. (2013). Modeling lactation curve standards for test-day milk yield in Holstein, Brown Swiss and Simmental cows, The Journal of Animal and Plant Sciences, 23(3), 754-762

[4] Bilgin, O.C., Esenbuga, N. and Davis, M.E. (2010) Comparison of models for describing the lactation curves of Awassi, Morkaraman and Tushin sheep, Archiv Tierzucht, 53(4), 447-456

[5] Cankaya, S., Unalan, A. and Soydan, E. (2011). Selection of mathematical model to describe the lactation curves of Jersey cattle, Archiv Tierzucht, 54 (1), 27-35

[6] Cappio-Borlino, Pulina, G. and Rossi, G. (1995) A non-linear modification of Wood's equation fitted to lactation curves of Sardinian dairy ewes, Small Ruminant Research, Volume 18(1), 75-79

[7] Ehrlich, J.L. (2013). Quantifying inter-group variability in lactation curve shape and magnitude with the MilkBot lactation model, PeerJ 1:e54 https://doi.org/10.7717/peerj.54

[8] Janković, M. (2016). New lactation model based on the Wood's and the MilkBot models, Master thesis (in Croatian), J.J. Strossmayer University of Osijek, Department of Mathematics

[9] Janković, M. and Paradžik, F. (2016). Wood's and MilkBot lactation models, Student seminar paper (in Croatian), J.J. Strossmayer University of Osijek, Department of Mathematics

[10] Madsen, K., Nielsen, H.B. and Tingleff, O. (2004). Methods for non-linear least squares problems, Informatics and Mathematical Modelling, Technical University of Denmark, Kongens Lyngby, Denmark