Technical note: Variation in daily milk yield calculations for dairy cows milked in an automatic milking system

P. P. Nielsen,^{*1} G. Pettersson,* K. M. Svennersten-Sjaunja,* and L. Norell†

^{*}Department of Animal Nutrition and Management, Swedish University of Agricultural Sciences, Kungsängen Research Centre, SE-753 23 Uppsala, Sweden

[†]Unit of Applied Statistics and Mathematics, Swedish University of Agricultural Sciences, PO Box 7013, SE-750 07 Uppsala, Sweden

ABSTRACT

An accurate estimation of the daily milk yield of dairy cows milked in an automatic milking system is not obvious because of variations in milking intervals and frequencies. Daily harvested milk varies substantially, and developing a method to be used for estimating daily milk production is of great importance. Three calculation methods (simple, semiadvanced, and advanced) were used. The simple method calculated rough daily milk production by summing up the yield per day. The semiadvanced used yield in combination with time since last milking to calculate the milk production per hour between milking; an average of the milk production per hour over the day was calculated and multiplied by 24. The advanced method calculated the milk production from midnight to midnight by using information about yield and time since last milking to calculate the exact milk production. The results show a clear preference for the advanced calculation method because the variation [variation for the advanced method = $\ln(1.79)$ for first lactation and $\ln(2.28)$ for later lactations] between days was reduced significantly (3 to 4 times lower compared with the simple method). Variation in daily harvested milk can be used as a management tool.

Key words: automatic milking system, yield estimate, calculation method

Milk recording for the estimation of lactation production is used for breeding purposes, for calculation of nutrient requirements, and for indications of illness. It has been questioned whether monthly or weekly recordings are sufficient for management decisions (Svennersten-Sjaunja et al., 1997). Furthermore, new managements systems such as automatic milking (AM) provide opportunities for more frequent recording without extra labor cost. Bach et al. (2007) found that the daily harvested milk and number of milking appear-

ances decreased with an increased lameness score in cows housed in a barn equipped with an AM system and that, in combination with feeding behavior, the variation in daily harvested milk could act as a tool for early detection of sick cows (Gonzalez et al., 2008). Yet, variation in milk yield between days and within cows needs to be reduced to a minimum to reduce the incidence of diagnosing healthy cows as sick (Lark et al., 1999) and vice versa.

Traditionally, estimation of milk production is performed when cows are milked at regular intervals, such as 2 or 3 times daily, in conventional systems. The prerequisites are different in systems with AM because milking intervals can vary within and between days, normally restricted by a fixed minimum interval between consecutive milking events. As a consequence of the varying intervals between milking and the relationship between milk yield and time since last milking being highly correlated (Friggens and Rasmussen, 2001), daily milk yield can vary substantially and does not reflect true daily milk production. It is of great importance to take this into account when calculating the daily milk yield of cows milked in an AM system.

Deviations in the shape of the lactation curve are a powerful tool used in the management of dairy cows (Grzesiak et al., 2006; Dematawewa et al., 2007). When lactation curves are generated from harvested milk of dairy cows, several models can be used with varying results regarding the fit of the model (Sherchand et al., 1995; Olori et al., 1999; Dematawewa et al., 2007). A few models, including a regression model by Ali and Schaeffer (1987) and the product of Mitscherlich and an exponential function by Rook et al. (1993), have been suggested because of a high level of predictability. These models do not take into account that the shape of lactation curves differ between the first and later lactations. The models proposed by Wilmlink (1987) generating different descriptions of the curves depending on first or later lactations are more useful.

Automatic milking provides data regarding harvested milk, milking intervals, and milking time for each milking. However, no previous study has evaluated how

Received May 26, 2009.

Accepted November 12, 2009.

¹Corresponding author: per.peetz.nielsen@huv.slu.se

Table 1. Number of cows arranged by the lactation number of first and later lactations¹

	Lactation no.							No. of cows
	1	2	3	4	5	6	7	
×	—	—	—	—	—	—	—	30
×	×	—	—	—	—	—	—	11
×	×	×	—	—	—	—	—	5
×	×	×	×	×	—	×	—	1
×	—	×	×	—	—	—	—	1
—	×	—	—	—	—	—	—	23
—	×	×	—	—	—	—	—	6
—	×	×	×	×	—	—	—	4
—	×	×	—	—	×	—	—	1
—	×	—	×	×	—	—	—	1
—	—	×	—	—	—	—	—	5
—	—	×	×	×	—	—	—	4
—	—	×	×	×	×	—	—	1
—	—	×	—	—	×	—	—	2
—	—	—	—	×	—	—	—	5
—	—	—	—	×	×	—	—	2
—	—	—	—	—	×	—	—	4
—	—	—	—	—	—	—	×	1
No. of lactations	48	52	30	18	10	1	1	

¹Total number of cows = 107; total number of lactations = 160.

daily variation in harvested milk affects the calculation and estimation of daily milk production. A calculation method should be found to estimate the most correct yield at any given time to detect a possible irregularity in the milk production, thereby providing information (e.g., indication of health or nutritional problems) that could be used as management tools. Thus, the objective of this study was to validate the variability between different calculation methods for daily milk production.

This study was conducted at the Swedish University of Agricultural Sciences (Kungsängen Research Centre, Uppsala, Sweden). The harvested milk from every milking occasion during 160 lactation periods (48 primiparous and 112 multiparous lactation periods) of 107 cows (Swedish Red) were used. The cows were housed in a loose-housing stall equipped with AM (VMS, DeLaval, Tumba, Sweden). The average yearly milk production in the herd was approximately 9,000 kg/cow.

The cows received grass silage roughage ad libitum and concentrate according to milk yield in concentrate feeding stations in the feeding area. In the AM station, the cows were fed a fixed amount of 0.7 kg of concentrate at every milking. Water was available in the feeding and the resting areas.

The cow traffic system was forced traffic with preselection gates as described in Melin et al. (2006). The resting and feeding areas were separated by 1-way gates, and the cows had to pass through a separation gate at the AM station to gain access to the feeding area. The minimum interval between 2 consecutive milkings was 6 h. All sick cows were removed from the barn for treatment and did not return until they had fully recovered.

Data were collected from May 2005 through December 2007. The lactation data from cows had to have a minimum of 150 d, starting within the first 15 d after parturition, and have fewer than 15 continuous days of missing data because of presumed health problems during the lactation. In total, there were 107 cows and 160 lactations. The number of primiparous cows was 48. Of these cows, 18 had later lactations, and together with 59 other cows they contributed 112 lactations, varying from 1 to 4 per cow (Table 1). The possible dependence among repeated lactations from the same cow was taken into consideration in model [3] below.

Daily milk yield was estimated using 3 different methods. Method I (simple) was defined as the sum of the milk harvested from midnight to midnight. Method II (semiadvanced) was defined as mean milk harvested per hour over 1 d multiplied by 24 h; the milk harvested per hour was calculated as a mean of the yield from each milking during the day divided by the time since last milking. Method III (advanced) was defined as the sum of milk harvested from midnight to first milking, from first to last milking during a day, and from last milking to midnight; the first and last terms were estimated by linear interpolation using the last milking the day before and the first milking the day after, respectively.

The data were analyzed in 2 steps. In step 1, models for lactation curves were estimated. The residual variance of each lactation curve was studied to compare the uncertainties among the 3 methods. These variances were expected to depend on methods for determining daily milk yield, cows, and separate lactation curves from the same cow, if any. These effects were taken into consideration in step 2, where the obtained variances in

step 1 were used as response variables in mixed linear models. As an alternative, the 2-step analysis could be joined into 1 model with 3 parallel lactation curves. But, because such a model would be complicated and because the primary interest was to compare the variations around the lactation curves and not the curves themselves, the 2-step approach was preferred.

The detailed model used for the lactations as proposed by Wilmink (1987) is as follows:

$$Y_{kt} = a_k + b_k t + c_k e^{-0.05t} + d_k t^2 + e_{kt}, \quad [1]$$

where Y_{kt} is estimated milk yield according to method $k = \text{I, II, III}$; $t = \text{days in lactation}$; a_k , b_k , c_k , and d_k are parameters for the curve; and the error e_{kt} is assumed to be normally distributed with mean 0 and variance τ_k^2 , $k = \text{I, II, III}$. According to Wilmink (1987), the effect $d_k t^2$ was excluded from the model for later lactations. Methods I, II, and III were studied one at a time for each cow and lactation, and the resulting variance estimates were kept for further analyses in models 2 and 3, below.

In step 2, the estimates of τ_k^2 were denoted by $\hat{\tau}_{ik}^2$ for the first lactation and $\hat{\tau}_{ijk}^2$ for later lactations, where i identifies the cow and j the lactation. To avoid large heteroscedastic effects, the logarithms were studied instead of the original residual variances of the lactation curves. There was no loss of generality because the primary interest was to search for the method, I, II, or III, having the least variance.

The mixed model used for comparison of the methods during the first lactation was

$$Y_{ik} = \ln(\hat{\tau}_{ik}^2) = \mu + c_i + \alpha_k + w_{ik}, \quad [2]$$

where μ was the overall level, $c_i \sim N(0, \sigma_c^2)$ was the random effect of cow $i = 1, \dots, 48$; α_k was the effect of calculation method $k = \text{I, II, III}$; and $w_{ik} \sim N(0, \sigma_k^2)$ was the error. The model included possible heteroscedasticity by allowing different values of σ_I^2 , σ_{II}^2 , and σ_{III}^2 . The chi-square test based on the difference of residual log-likelihoods was used for the hypothesis of equality $\sigma_I^2 = \sigma_{II}^2 = \sigma_{III}^2$ among the error variances.

The model used when analyzing the calculation methods during later lactations was

$$Y_{ijk} = \ln(\hat{\tau}_{ijk}^2) = \mu + c_i + \beta_j + (c\beta)_{ij} + \alpha_k + (\beta\alpha)_{jk} + w_{ijk}, \quad [3]$$

where μ was the overall level, $c_i \sim N(0, \sigma_c^2)$ was the random effect of cow $i = 1, \dots, 77$; β_j was the effect of lactation number $j = 2, \dots, 7$; $(c\beta)_{ij} \sim N(0, \sigma_{c\beta}^2)$ was the interaction between cow and lactation number; α_k was the effect of calculation method $k = \text{I, II, III}$; $(\beta\alpha)_{jk}$ was the interaction between lactation number and calculation method; and $w_{ijk} \sim N(0, \sigma_k^2)$ was the error. The numerical calculations were performed by procedures in SAS (SAS Institute, 2004) for regression and mixed model analysis.

When analyzing the 3 calculation methods during the first lactation, there was not a significant result for heteroscedasticity ($\chi_{2df}^2 = 3.7$; $P = 0.16$). The joint estimate of the error variance, assuming $\sigma_I^2 = \sigma_{II}^2 = \sigma_{III}^2$, was 0.0367, and cow variance σ_c^2 was estimated at 0.0959. The method effects, corresponding to the logarithms of the estimates of the regression error variances τ_I^2 , τ_{II}^2 , and τ_{III}^2 from model [1], were significantly different when compared by model [2] ($F_{2,94} = 762.8$; $P < 0.001$; Table 2). The pairwise comparisons of the logarithms of the least squares estimates of the variances τ_I^2 , τ_{II}^2 , and τ_{III}^2 for the first lactation showed clear differences. The differences between method I and methods II and III were significant because the t -values corresponded to $P < 0.001$, and the differences between method II and method III were significant because the t -values corresponded to $P < 0.01$. In particular, retransforming by anti-logs showed that calculation method I had a variance approximately 4 times larger than the other 2 methods. These results show that using calculation method III when estimating daily milk production during the first lactation from an AM system resulted in the least variation between the calculated and expected production.

When analyzing the later lactations, the hypothesis $\sigma_I^2 = \sigma_{II}^2 = \sigma_{III}^2$ was rejected ($\chi_{2df}^2 = 14.1$; $P = 0.0009$) and the estimates of the variances were 0.0837, 0.0211, and 0.0241, respectively. The estimate of the cow variance σ_c^2 was 0.0604 and the interaction variance $\sigma_{c\beta}^2$ (i.e., among lactation occasions) was estimated at 0.0748. The fixed effects of lactation number and the interaction with methods were not significant ($F_{5,31} = 2.30$, $P = 0.069$ and $F_{10,212} = 0.56$, $P = 0.84$, respectively). In the final version of model [3], by removing β_j and $(\beta\alpha)_{jk}$, the method effects were significantly different ($F_{2,222} = 928.3$; $P < 0.001$; Table 3). The pairwise comparisons of the logarithms of the least squares esti-

Table 2. Means using model 2 for the logarithm of estimated error variances (τ^2) in regression model 1 applied to calculation methods I, II, and III^{1,2,3}

Item	Calculation method		
	I	II	III
Model 2 estimate of $\ln(\tau^2)$	3.174	1.924	1.790
1		0.039	0.039
2	32.0		0.039
3	35.4	3.4	

¹See text for models 1 and 2.

²Method I defined as the sum of the milk harvested from midnight to midnight. Method II defined as mean milk harvested per hour over 1 d multiplied by 24 h; the milk harvested per hour was calculated as a mean of the yield from each milking during the day divided by the time since last milking. Method III defined as the sum of milk harvested from midnight to first milking, from first to last milking during a day, and from last milking to midnight; the first and last terms were estimated by linear interpolation using the last milking the day before and the first milking the day after, respectively.

³For the pairwise comparisons, the standard error of the difference and the *t*-value are given above and below the diagonal, respectively.

mates of the variances τ^2_I , τ^2_{II} , and τ^2_{III} for the later lactations showed clear differences. All differences were significant because the *t*-values corresponded to $P < 0.001$. In particular, retransforming by anti-logs showed that calculation method I had a variance approximately 3 to 4 times larger than the others. These results showed that using calculation method III when estimating daily milk production during later lactations from an AM system resulted in the least variation between the calculated and expected production.

The daily harvested milk varied substantially when dairy cows were milked in an AM system because of variation in milking intervals. If the variations in daily harvested milk were used to monitor cow health, it would be of importance to examine which estimation methods produced the least variation between days. Our results suggest that to reduce the daily variation in estimated daily yield to a minimum, the most advanced method, method III, was justified. If the deviations in harvested milk were used to monitor health status, or any management routines, the variation in estimated yield between days would have to be reduced to an absolute minimum to reduce the incidence of diagnosing healthy cows as sick and vice versa. The estimate of daily milk yield determined from daily harvested milk by calculation method I resulted in a variation between days that was too large. Even though calculation method III demanded more programming and data manipulation compared with methods I and II, the results presented advocate that this was justified if the harvested milk is used in the future to monitor health and management of cows milked in an AM system.

Table 3. Means using simplified version of model 3 for the logarithm of estimated error variances (τ^2) in regression model 1 applied to calculation methods I, II, and III^{1,2,3}

Item	Calculation method		
	I	II	III
Model 3 estimate of $\ln(\tau^2)$	3.565	2.404	2.283
1		0.030	0.030
2	38.3		0.020
3	42.3	6.0	

¹See text for models 1 and 3.

²Method I defined as the sum of the milk harvested from midnight to midnight. Method II defined as mean milk harvested per hour over 1 d multiplied by 24 h; the milk harvested per hour was calculated as a mean of the yield from each milking during the day divided by the time since last milking. Method III defined as the sum of milk harvested from midnight to first milking, from first to last milking during a day, and from last milking to midnight; the first and last terms were estimated by linear interpolation using the last milking the day before and the first milking the day after, respectively.

³For the pairwise comparisons, the standard error of the difference and the *t*-value are given above and below the diagonal, respectively.

We have shown that by using the model proposed by Wilmink (1987) for describing lactation curves for multiparous cows and a more advanced method (III) for calculating the estimated daily yield, the variation between expected yield (model) and estimated yield (calculation method III) does not vary between lactations. Furthermore, this information allowed prediction of the variation during later lactations from the variation during the present and earlier lactations, something that might be useful when variations in milk production during the early stage of the lactation are used as indicators of health problems.

ACKNOWLEDGMENTS

This study was funded by the Swedish Farmers' Foundation for Agricultural Research, Stockholm, Sweden (project no. V0830426).

REFERENCES

Ali, T. E., and L. R. Schaeffer. 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67:637–644.

Bach, A., M. Dinares, M. Devant, and X. Carre. 2007. Associations between lameness and production, feeding and milking attendance of Holstein cows milked with an automatic milking system. *J. Dairy Res.* 74:40–46.

Dematawewa, C. M. B., R. E. Pearson, and P. M. VanRaden. 2007. Modeling extended lactations of Holsteins. *J. Dairy Sci.* 90:3924–3936.

Friggens, N. C., and M. D. Rasmussen. 2001. Milk quality assessment in automatic milking systems: Accounting for the effects of variable intervals between milkings on milk composition. *Livest. Prod. Sci.* 73:45–54.

Gonzalez, L. A., B. J. Tolcamp, M. P. Coffey, A. Ferret, and I. Kyriazakis. 2008. Changes in feeding behavior as possible indicators

- for the automatic monitoring of health disorders in dairy cows. *J. Dairy Sci.* 91:1017–1028.
- Grzesiak, W., P. Blaszczyk, and R. Lacroix. 2006. Methods of predicting milk yield in dairy cows—Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs). *Comput. Electron. Agric.* 54:69–83.
- Lark, R. M., B. L. Nielsen, and T. T. Mottram. 1999. A time series model of daily milk yields and its possible use for detection of a disease (ketosis). *Anim. Sci.* 69:573–582.
- Melin, M., G. G. N. Hermans, G. Pettersson, and H. Wiktorsson. 2006. Cow traffic in relation to social rank and motivation of cows in an automatic milking system with control gates and an open waiting area. *Appl. Anim. Behav. Sci.* 96:201–214.
- Olori, V. E., S. Brotherstone, W. G. Hill, and B. J. McGuirk. 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. *Livest. Prod. Sci.* 58:55–63.
- Rook, A. J., J. France, and M. S. Dhanoa. 1993. On the mathematical description of lactation curves. *J. Agric. Sci.* 121:97–102.
- SAS Institute. 2004. SAS System for Windows. Version 9.1. SAS Institute Inc., Cary, NC.
- Sherchand, L., R. W. McNew, D. W. Kellogg, and Z. B. Johnson. 1995. Selection of a mathematical model to generate lactation curves using daily milk yields of Holstein cows. *J. Dairy Sci.* 78:2507–2513.
- Svennersten-Sjaunja, K., L. O. Sjaunja, J. Bertilsson, and H. Wiktorsson. 1997. Use of regular milking records versus daily records for nutrition and other kinds of management. *Livest. Prod. Sci.* 48:167–174.
- Wilmink, J. B. M. 1987. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. *Livest. Prod. Sci.* 16:335–348.