

Analysis of water intake and dry matter intake using different lactation curve models

E. Kramer,^{*1} E. Stamer,[†] J. Spilke,[‡] G. Thaller,^{*} and J. Krieter^{*}

^{*}Institute of Animal Breeding and Husbandry, Christian-Albrechts-University, D-24118 Kiel, Germany

[†]TiDa Tier und Daten GmbH, D-24259 Westensee/Brux, Germany

[‡]Biometrics and Informatics in Agriculture Group, Martin-Luther-University, D-06108 Halle/Saale, Germany

ABSTRACT

The objective was to evaluate 6 different lactation curve models for daily water and dry matter intake. Data originated from the Futterkamp dairy research farm of the Chamber of Agriculture of Schleswig-Holstein in Germany. A data set of about 23,000 observations from 193 Holstein cows was used. Average daily water and dry matter intake were 82.3 and 19.8 kg, respectively. The basic linear mixed model included the fixed effects of parity and test-day within feeding group. Additionally, 6 different functions were tested for the fixed effect of lactation curve and the individual (random) effect of cow-lactation curve. Furthermore, the autocorrelation between repeated measures was modeled with the spatial (power) covariance structure. Model fit was evaluated by the likelihood ratio test, Akaike's and Bayesian information criteria, and the analysis of mean residual at different days in milk. The Ali and Schaeffer function was best suited for modeling the fixed lactation curve for both traits. A Legendre polynomial of order 4 delivered the best model fit for the random effect of cow-lactation. Applying the error covariance structure led to a significantly better model fit and indicated that repeated measures were autocorrelated. Generally, the best information criteria values were yielded by the most complex model using the Ali and Schaeffer function and Legendre polynomial of order 4 to model the average lactation and cow-specific lactation curves, respectively, with inclusion of the spatial (power) error covariance structure. This model is recommended for the analysis of water and dry matter intake including missing observations to obtain estimation of correct statistical inference and valid variance components.

Key words: dairy cow, lactation curve, water intake, dry matter intake

INTRODUCTION

Water and feed intake are measured automatically and completely in many test stations or dairy research farms, (e.g., Coffey et al., 2002; Kramer et al., 2008a; Hüttmann et al., 2009). In the future, water and feed intake will become more important for dairy management because of their effect on cow health (González et al., 2008; Lukas et al., 2008) as well as for dairy breeding because of their relationship with cow's energy status and the corresponding liability to diseases in the first part of lactation (Veerkamp and Thompson, 1999; Coffey et al., 2002; Hüttmann et al., 2009). Furthermore, the inclusion of feed intake in future breeding programs is becoming conceivable as cost for feed concentrates increases. Unfortunately, recording daily feed intake is difficult and costly, and hence only feasible in test stations or research herds (Hüttmann, 2007). In contrast, individual recording of water intake is less expensive (Kramer et al., 2008a). Because the correlation between these traits is high ($r = 0.73$; Kramer et al., 2008a), water intake may be suitable as an information trait for feed intake and could be included in dairy programs instead of feed intake. Although necessary for valid statistical inference and correct variance component estimation, literature on model analysis of water and feed intake data are scarce (van der Werf, 2001; Sawalha et al., 2005; Kramer et al., 2008b).

Today, random regression (**RR**) models are increasingly used for the estimation of breeding values and commonly recommended for statistical analysis in this field (e.g., Koenen and Veerkamp, 1998; Veerkamp and Thompson, 1999; Coffey et al., 2002). Such models allow the estimation of cow-specific lactation curves with additional random regression coefficients (Schaeffer and Dekkers, 1994; Schaeffer, 2004). In contrast, conventional fixed regression (**FR**) models only include the fixed effect of lactation curve estimated with universal average regression coefficients, which are the same for all cows (van der Werf, 2001). In a previous study, Kramer et al. (2008a) analyzed the relationship of water and feed intake in the course of lactation. They used the estimated parameters of the Ali and Schaeffer (1987)

Received December 9, 2008.

Accepted May 1, 2009.

¹Corresponding author: ekramer@tierzucht.uni-kiel.de

Table 1. Numbers of cows, observations (Obs.), missing observations, means, standard deviations, and range (minimum, maximum) of the 2 analyzed traits

Trait	Cows (n)	Obs. (n)	Missing obs. (n)	Mean	SD	Minimum	Maximum
Water intake (kg/d)	193	22,660	468	82.3	19.0	10.7	155.6
DMI (kg/d)	193	22,624	496	19.8	4.0	3.8	34.8

function for both the fixed and random regression coefficients to model average and cow-specific lactation curves. In the German national genetic evaluation of milk yield, the Wilmlink function (Wilmlink, 1987) is currently used to model fixed regression coefficients, and a Legendre polynomial of order 2 is used to model random regression coefficients on day of lactation (VIT, 2008). Modeling lactation curves for milk yield has frequently been discussed in the literature, whereas modulation of water or feed intake over the lactation is rare (e.g., Wood, 1967; Guo and Swalve, 1995; van der Werf, 2001; Silvestre et al., 2006).

The objective of the present study was to analyze daily water and DMI with different lactation curve models. In doing so, 1) the best function for the average lactation curve was evaluated and chosen as the basis for 2) the evaluation of the best (co)variance function to model cow-specific lactation curves.

MATERIALS AND METHODS

Data

Data were recorded on the Futterkamp dairy research farm of the Chamber of Agriculture of Schleswig-Holstein, Germany, between March 2005 and April 2006. The dairy herd was subdivided into a research herd and a production herd. The research herd, comprising nearly 70 cows at a time, was divided into 2 feeding groups (group A and group B). Different experiments with feed additives (e.g., propylene glycol for group A and glycerin for group B) were being performed within the research herd. For such feeding experiments, it is important to have cows in early and mid lactation with high milk yield. Thus, for every feeding experiment, cows in early and mid lactation moved from the production into the research herd and vice versa. Complete lactation length could therefore not be recorded because most of the cows in late lactation had already left the feeding groups. During data collection, 3 feeding experiments were being performed. Dry matter intake was not recorded during the maximal interval of 3 wk between the feeding experiments. Hence, nearly 23,000 cow-days were accumulated from a total of 193 Holstein cows in parities 1 to 9. Lactation days were between 6 and 230. Twenty-three cows had observations in 2

lactations. Cows were milked twice daily and fed an ad libitum TMR. The feeding and water troughs (Insentec, Marknesse, the Netherlands) were equipped with an individual cow identification system; hence, cows were only able to access the troughs one at a time. Each visit to the water and feeding troughs was routinely recorded and the amount of collected feed and water were accumulated to daily yields. Extreme values that deviated more than 4 SD from the mean were removed from the data set. Thus, observations ranged between 10.7 and 155.6 kg of water intake and between 3.8 and 34.8 kg for DMI (Table 1), and the average DM content of the TMR was about 45% during the data collection period. In addition, only lactations with at least 15 test-days were considered in the analysis. Furthermore, the first and last days of each feeding experiment were excluded and data from 3 more days were discarded because of general technical problems. A total of 800 records (3.4% of all records) were omitted from data analysis.

Data Analysis

Preliminary investigations were performed using SAS software (SAS Institute, 2005) for analysis of fixed effects. Only the significant fixed effects were included in the following FR model:

$$y_{ijkm}(\text{DIM}) = \mu + P_i + \text{GTD}_j + f_{iw}(\text{DIM}) + c_k + e_{ijkm},$$

where y_{ijkm} is the observation vector of water intake or DMI; μ is the overall mean; P_i is the fixed effect of the i th parity class ($i = 1, \dots, 3$ for first, second, and third or greater); GTD_j is the fixed effect of the j th test-day within feeding group ($j = 1, \dots, 664$); $f_{iw}(\text{DIM})$ describes 6 functions ($w = 1, \dots, 6$) to model the i th lactation curve as described below, where DIM is days in milk, c_k is the random effect of the k th cow ($k = 1, \dots, 193$), and e_{ijkm} is the random error.

In a first step, the fixed effect of the average lactation curve was modeled by the following 6 functions of DIM. These functions were chosen because they are commonly accepted and well established in the literature for modeling lactation curves of different traits:

- 1) Guo and Swalve (1995; **GS**):

$$f_1(\text{DIM}) = a_{1i} + a_{2i} \times (\text{DIM})^{0.5} + a_{3i} \times \ln(\text{DIM});$$

2) Wilmink (1987; **Wi**):

$$f_2(\text{DIM}) = a_{1i} + a_{2i} \times \text{DIM} + a_{3i} \times e^{(-0.05 \times \text{DIM})};$$

3) Legendre polynomial of second order (**Lg2**; Broth-
erstone et al., 2000):

$$f_3(\text{DIM}) = a_{1i} + a_{2i} \times z + a_{3i} \times 0.5 \times (3z^2 - 1);$$

4) Legendre polynomial of third order (**Lg3**; Broth-
erstone et al., 2000):

$$f_4(\text{DIM}) = a_{1i} + a_{2i} \times z + a_{3i} \times 0.5 \times (3z^2 - 1) \\ + a_{4i} \times 0.5 (5z^3 - 3z);$$

5) Legendre polynomial of fourth order (**Lg4**; Broth-
erstone et al., 2000):

$$f_5(\text{DIM}) = a_{1i} + a_{2i} \times z + a_{3i} \times 0.5 \times (3z^2 - 1) \\ + a_{4i} \times 0.5 \times (5z^3 - 3z) \\ + a_{5i} \times 0.125 \times (35z^4 - 30z^2 + 3);$$

6) Ali and Schaeffer (1987; **AS**):

$$f_6(\text{DIM}) = a_{1i} + a_{2i} \times (\text{DIM}/305) + a_{3i} \\ \times (\text{DIM}/305)^2 + a_{4i} \times \ln(305/\text{DIM}) \\ + a_{5i} \times [\ln(305/\text{DIM})]^2;$$

with $z = -1 + 2 [(\text{DIM} - 5)/(230 - 5)]$.

An additional criterion to evaluate model fitting ability is the average residual at different lactation days (Silvestre et al., 2006). The mean error should be small and residuals should randomly oscillate between negative and positive values. This indicates that model fit is consistently appropriate over the whole lactation. To check homogeneity of residual variance along the lactation period, the mean residual obtained with each FR model function was plotted against the lactation day.

The function that delivered the best model fit was chosen as the basis for the RR model (i.e., this function was chosen to model the average lactation curve with the RR model). In the second step, modeling the cow-specific lactation curves was carried out again by applying functions 1 to 6. Hence, the RR model differed from FR model in that the 6 functions were used to describe the random effect of the k th cow ($k = 193$),

which encompasses both the animal genetic and permanent environmental effects. In further contrast to FR model, heterogeneous residual variance was assumed across the 3 parity classes, and the spatial (power) covariance structure [**SP(POW)**] for residuals was applied in RR model. This was because data sets with repeated daily measures within cow are assumed to contain interdependent and thus autocorrelated repeated measures (Littell et al., 1998, 2006). Dependencies between the residuals of repeated yields can be modeled with covariance structures (Sawalha et al., 2005; Mielenz et al., 2006; Kramer et al., 2008b). In a previous study, Kramer et al. (2008b) reported autocorrelated residuals of daily water and feed intake and emphasized the use of a spatial covariance structure for data sets with missing values. This structure assumes constant residual variance at the different stages of lactation. In a previous study, residual variance of water and dry matter intakes was almost constant during the lactation (Kramer et al., 2008b). For the SP(POW) structure, the correlations (r_e) decline as a function of time. The function is defined as $g(d) = r_e^d$, where d is the temporal distance between 2 measurements at times t_1 and t_2 , $d = |t_1 - t_2|$. The SP(POW) structure models the covariance between t_1 and t_2 as $\text{Cov}[Y_{t_1}, Y_{t_2}] = \sigma_e^2 \times r_e^d$. The SP(POW) type can be used for unequally spaced data with characteristically different distances between the measures. The better the RR models fit the data, however, the lower the correlation between residuals is expected to be. Hence as a final step, the model with the best fit was compared with the same model excluding the SP(POW) error covariance structure to confirm the estimated correlations between the residuals and thus verify the necessity of the error covariance structure.

Criteria for Selection of the Models

Comparison of different FR models was based on the maximum likelihood principle. The MIXED procedure (SAS Institute, 2005) provides different model selection criteria by default. The maximum likelihood principle was used for testing the FR models, because the information criteria based on the REML principle would not have been comparable because 2 models including different fixed effects were compared. Random regression model comparison was based on REML. Models, where one could be reduced to a special case of the other, were compared by applying the likelihood ratio test (**LRT**), which is a statistical test of the quality of fit of 2 hierarchically nested models (Littell et al., 2006). Those models are identical in their design matrices of the fixed effects. The LRT is calculated as the log-likelihood dif-

ference [$\Delta(-2\log L)$] of the 2 comparable models and approximates a chi-square distribution with Δq degrees of freedom, where q is the number of estimated covariance components of each model.

Models with different functions for the fixed lactation effect or the random cow effect are not hierarchically nested. For the comparison of such models, the information criteria of Hurvich and Tsai (1989; Akaike's information criteria corrected, **AICC**) and Schwarz (1978; Bayesian information criteria, **BIC**) were used. These values take the number of estimated parameters into account and favor less complex model variants. The model with the smallest values for AICC and BIC is selected without making a statement about underlying significance. In contrast, LRT yields a significance test under the null hypothesis that the reduced model is correct.

RESULTS

Lactation Curves

For primiparous cows, lactation curves for the 4 functions (GS, Wi, AS, and Lg4) fitted in the FR model are shown in Figure 1. The curves of functions Lg2 and Lg3 were very similar to Lg4 and were therefore omitted from Figure 1.

Lactation curves for primiparous cows were only marginally affected by the underlying function. For water intake, functions Wi, GS, and AS showed nearly the same trajectory. The Lg4 function differed slightly from the other 3. A moderate increase could be observed for the AS and Lg4 functions at lactation d 200. For DMI, all functions were very similar. Lactation curves of both traits for the multiparous cows characteristically differed from those of primiparous cows, indicating a higher increase at the beginning and a greater decrease at the end of lactation (results not shown). Additionally, only a marginal influence of the underlying function on the curves could be observed.

Comparison of Different FR Model Functions of Average Lactation Curves

For both traits, residual variance, number of fixed effects, log-likelihood values, and information criteria of different FR models are given in Table 2. For each trait, only the Lg2, Lg3, and Lg4 functions were hierarchically nested. Differences between these nested models, based on LRT, were significant ($P < 0.01$). Otherwise, LRT, AICC, and BIC results did not lead to different conclusions for model selection, although differences between models became smaller with the AICC and even more so with the BIC criterion because both criteria tend to penalize more complex models.

For both traits, the AS and Lg4 functions were the most suitable for modeling the average lactation curve. Of these 2 functions, AS was associated with somewhat lower (more desirable) information criteria values. For water intake, BIC differences between models were small, with the exception of Lg2.

Different models were also compared with respect to the plot of the mean residual water (Figure 2) and dry matter (Figure 3) intake against DIM. This was done to check whether the functions provide a constant model fit and homogeneity of variance of average residuals along the whole lactation. The AS and Lg4 functions showed the best fit, with average residuals distributed uniformly around the value zero. Residuals of the other functions, especially Lg2, showed heterogeneity of variance while observing considerable deviations from a uniform distribution around the value zero. It was obvious that with Lg2, the predicted values at the beginning (end) of the lactation are overestimated (underestimated) leading to negative (positive) residuals. In contrast, residuals obtained with the Wi and GS functions were mainly positive at the beginning and negative at the end of the lactation. This was most pronounced for DMI.

Comparison of Different RR Model Functions of Cow-Specific Lactation Curves

Because of the slightly better model fit compared with Lg4 and no observable differences between the distribution of the separate average residuals for both traits, the AS function was chosen to model the fixed effect of lactation curve while testing RR models. The same 6 functions were tested for modeling cow-specific lactation curves. The number of (co)variance components, restricted log-likelihood values, and information criteria of the different RR models are given in Table 3.

Comparing all model functions, the best model fit was achieved with RR (Lg4) for both traits. The Wi function showed the highest (least desirable) information criteria.

Finally, the model with the best fit [RR (Lg4)] was applied with and without an SP(POW) error covariance structure to confirm the estimated correlations between residuals. This was done to verify the necessity of the assumption of the SP(POW) error covariance structure. For both traits, the average estimated correlation between repeated measures and the estimated residual variance (obtained from the model including the error covariance structure) is shown in Table 4. Results of the LRT comparison to the same model excluding the SP(POW) error covariance structure are also included (Table 4).

For both traits, the LRT indicated a significantly better model fit for the RR (Lg4) model including an error covariance structure. Repeated measures of both traits were autocorrelated with r_e between -0.04 and -0.14 , depending on the parity class. Estimated residual variance increased with parity classes 1, 2, and 3 from 67.34 to 126.77 and from 2.63 to 4.75 for water intake and DMI, respectively.

DISCUSSION

Lactation Curves

Lactation curves for DMI were similar to those presented by Veerkamp and Thompson (1999) and Hüttmann (2007). For water intake, lactation curves were akin to those observed by Murphy et al. (1983)

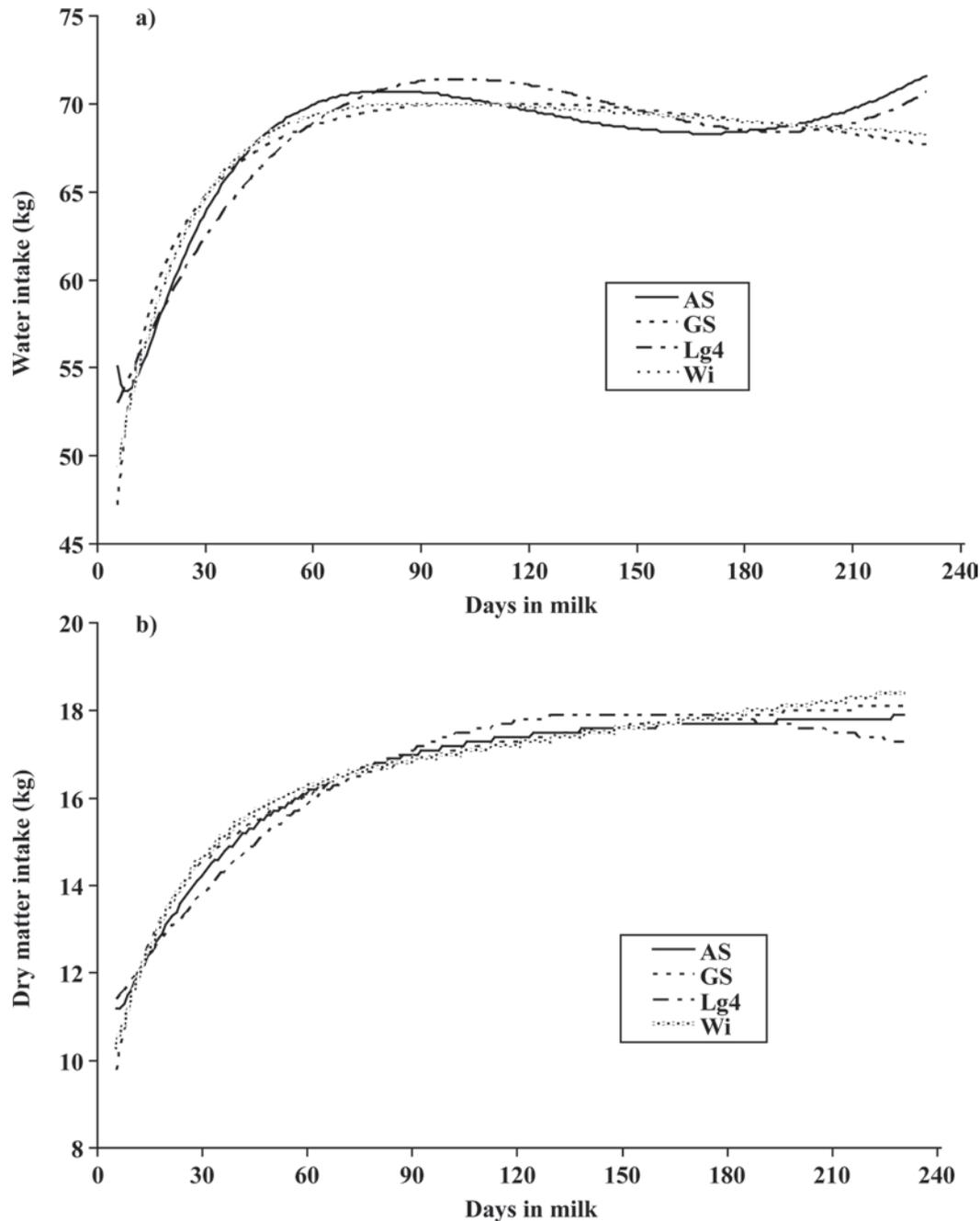


Figure 1. Lactation curves of a) water and b) DMI for the functions of Wilmlink (Wi), Guo and Swalve (GS), Ali and Schaeffer (AS), and the Legendre polynomial of order 4 (Lg4) in primiparous cows.

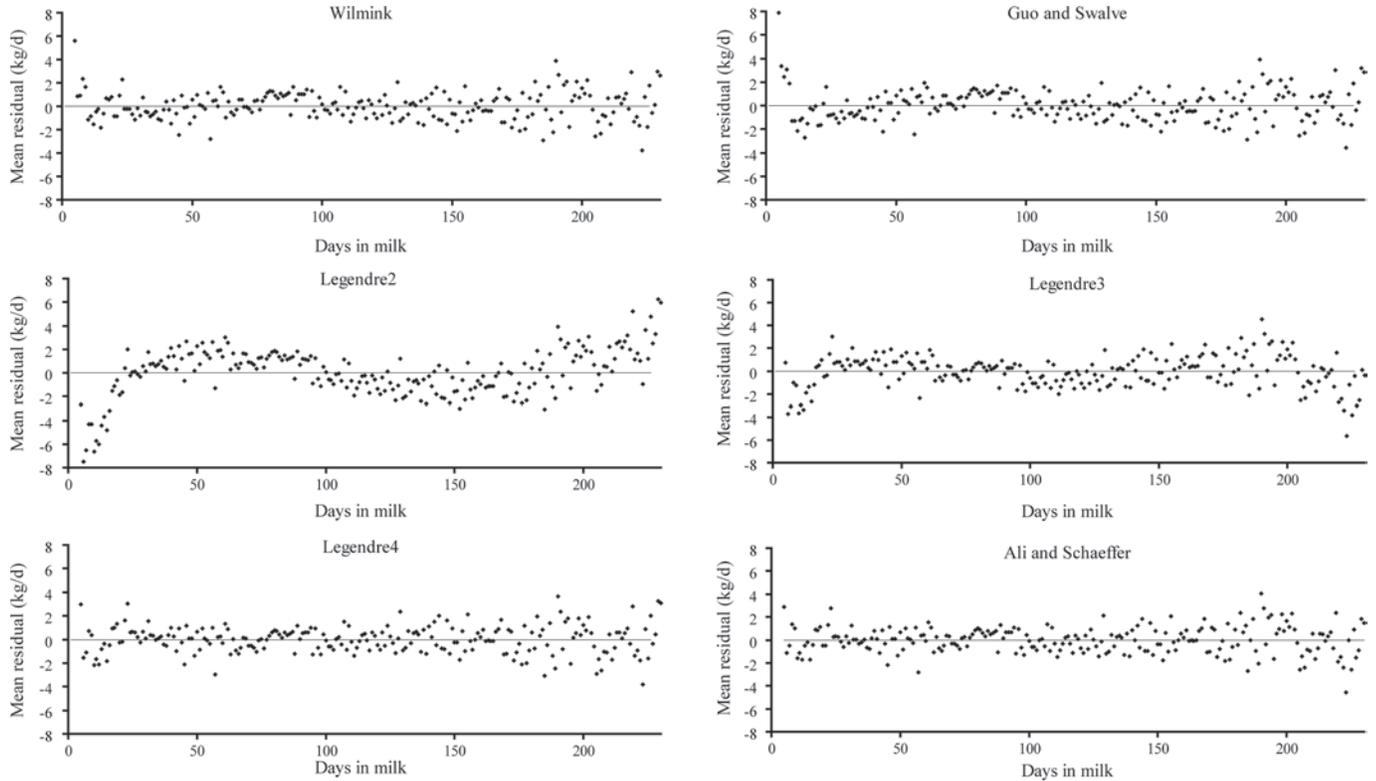


Figure 2. Mean residuals of water intake (kg/d) against DIM for the 6 functions.

except for an increase of water intake at the beginning of late lactation observed with the Lg4 and AS functions. Apart from this increase, curves for water intake showed a trajectory similar to the well-known curves for milk yield. However, the increase of water intake at lactation d 200 could be a good indication

that only functions with 5 parameters (AS and Lg4) are flexible enough to model water intake. Hence, it came as no surprise that the AS and Lg4 functions were found more suitable for modeling lactation curves of water intake compared with the Wi and GS functions. At this point, it should be noted again that only par-

Table 2. Estimated residual variance ($\hat{\sigma}_e^2$) and information criteria of the different functions under the fixed regression (FR) model for water intake and DMI¹

Model (function)	$\hat{\sigma}_e^2$	p	Information criteria ²		
			$\Delta(-2\log L)$	$\Delta(\text{AICC})$	$\Delta(\text{BIC})$
Water intake					
FR (GS)	108.94	669	107	94	76
FR (Wi)	108.63	669	44	32	13
FR (Lg2)	111.22	669	572	560	541
FR (Lg3)	108.96	670	114	118	99
FR (Lg4)	108.47	671	12	11	12
FR (AS)	108.42	671			
DMI					
FR (GS)	4.32	669	218	205	188
FR (Wi)	4.32	669	229	216	197
FR (Lg2)	4.48	669	1,057	1,043	1,024
FR (Lg3)	4.32	670	243	236	227
FR (Lg4)	4.28	671	34	33	33
FR (AS)	4.27	671			

¹GS = function of Guo and Swalve; Wi = function of Wilmink; Lg2–Lg4 = Legendre polynomial of order 2, 3, and 4; AS = function of Ali and Schaeffer; p = number of fixed effect levels included in the FR model.

²AICC = Akaike’s information criteria corrected; BIC = Bayesian information criteria, calculated as difference to the FR (AS) model.

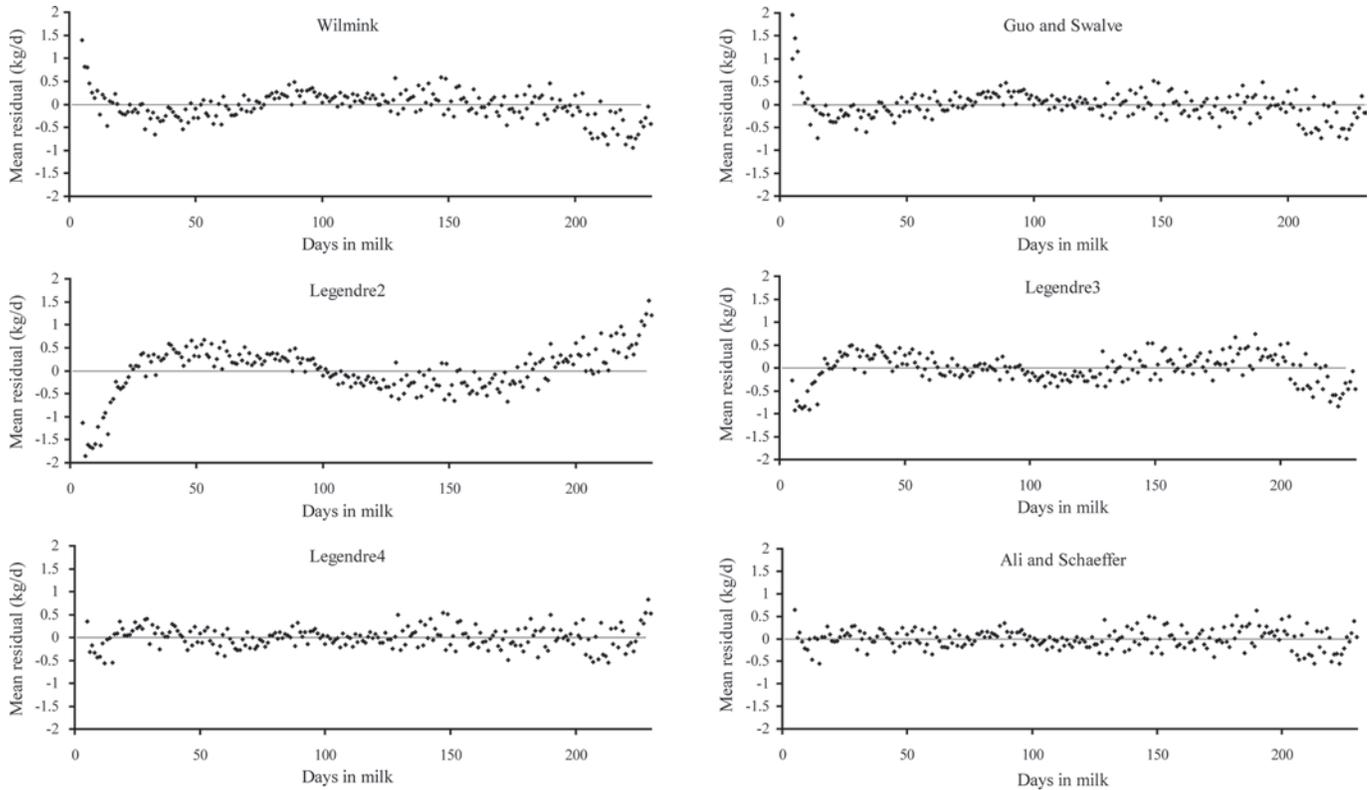


Figure 3. Mean residuals of DMI (kg/d) against DIM for the 6 functions.

tial lactations were modeled because records were only available until lactation d 230. Furthermore, only few observations from lactation d 200 to 230 entered the analysis, and thus the increase of water intake at the beginning of late lactation might not be representative for the whole population.

Comparison of Different FR Model Functions of Average Lactation Curves

The AS function was better than the other functions at modeling fixed average lactation curves. Only the Lg4 function had a fit close to that of AS. Hence, the assumption after visualization of the lactation curves that the functions with 5 parameters might be more flexible and suitable was indeed confirmed. The AS function has been frequently used to model lactation curves of milk yield (López-Romero and Carabaño, 2003). In other studies, this function was also applied to analyze the relationship between water and DMI during the course of a lactation (Kramer et al., 2008a) and to investigate the autocorrelation patterns of residuals of water and DMI (Kramer et al., 2008b). This function was also used for feed intake and energy balance analysis by Woodford et al. (1984), de Vries et

al. (1999), and Collard et al. (2000). Furthermore, the AS function has been used to model the lactation curve of SCC (Reents et al., 1995).

For both traits the average residuals plotted against days of lactation confirmed the model fit statistics and our assumption after visual inspection of the particular lactation curves. It was obvious that functions with 5 parameters (AS and Lg4) were best for modeling, whereas the other functions left at least some degree of heterogeneity of variance unaccounted for. The worst modeling was observed with Lg2, in accordance with results of the study by Silvestre et al. (2006). Generally, modeling the beginning and end of lactation seemed to be more difficult than modeling the middle. This was assumed to be the case especially for functions with inferior model fit because such functions provide systematic over- and underestimation of predicted values and thus negative/positive residuals at the beginning or end of lactation (Silvestre et al., 2006). Additionally, model performance is strongly affected by the number of underlying observations per lactation (Silvestre et al., 2006). Furthermore, the availability of test-day records before peak yield is crucial for the correct estimation of the lactation curve shape (Macciotta et al., 2005). This might also be an explanation for the clear differences

Table 3. Number of (co)variance components (q) and information criteria of the different functions under the random regression (RR) model

Model (function) ¹	q	Information criteria ²		
		$\Delta(-2R\log L)$	$\Delta(\text{AICC})$	$\Delta(\text{BIC})$
Water intake				
RR (GS)	12	353	335	305
RR (Wi)	12	390	372	343
RR (Lg2)	12	311	293	263
RR (Lg3)	16	132	122	105
RR (AS)	21	62	63	65
RR (Lg4)	21			
DMI				
RR (GS)	12	491	472	443
RR (Wi)	12	526	507	479
RR (Lg2)	12	508	489	460
RR (Lg3)	16	308	297	281
RR (AS)	21	106	107	110
RR (Lg4)	21			

¹GS = function of Guo and Swalve; Wi = function of Wilmlink; Lg2–Lg4 = Legendre polynomial of order 2, 3 and 4; AS = function of Ali and Schaeffer.

²AICC = Akaike's information criteria corrected; BIC = Bayesian information criteria, calculated as difference to the RR (Lg4) model.

between performance of separate functions considered in this study, as the research herd is characterized by a frequently changing cow stock. There were several cows with only few observations per lactation (at least 15 by definition) and cows with a considerable interval between calving and first test-day because they remained in the production herd for the first part of lactation. Generally, there may be more appropriate data sets to analyze lactation curve models. However, the average number of 109 records per lactation indicates that data are suitable, albeit not optimal, for this type of analysis.

Comparison of Different RR Model Functions of Cow-Specific Lactation Curves

Dynamics of traits such as water intake or DMI are modeled separately for each cow and lactation day using RR models, leading to a more effective consideration of the underlying biology and, therefore, more accurate results (van der Werf, 2001; Mielenz et al., 2006).

Comparing different model functions, the RR (Lg4) function provided the best fit, indicating its suitability

for mapping individual water and feed intake of each cow during lactation. Legendre polynomials have been used quite often to model lactation curves of milk yield (Liu et al., 2006) and for the estimation of feed intake and energy balance parameters (Coffey et al., 2002). Silvestre et al. (2006) hypothesized that the Lg4 function, which is a polynomial of the fourth degree, may fit daily data better than functions with fewer than 5 parameters such as Wi or Legendre polynomials of fewer than 4 degrees. This was in agreement with results of the present daily-database study. For both fixed and random lactation curves the functions with 5 parameters (AS and Lg4) delivered the best information criteria compared with other functions with fewer parameters (Wi, GS, Lg2, and Lg3).

In the present study, 2 different functions were found to provide the best model fit for fixed and random lactation curves for both traits. In a similar fashion, the routine German national genetic evaluation uses 2 different functions (Wi, Lg2) to model the fixed and random lactation curves of milk yield (VIT, 2008). In the present study, it was suggested that for the random lactation curves a more flexible function is better suited

Table 4. Estimated residual variance ($\hat{\sigma}_e^2$)¹ and correlation between repeated measures (r_e)¹ by parity class (1, 2, and 3), and likelihood ratio test (LRT) results for random regression (Legendre polynomial order 4) model with and without a covariance structure

Trait	$\hat{\sigma}_e^2$			r_e			$\Delta(-2R\log L)$	LRT
	1	2	3	1	2	3		
Water intake	67.34	98.68	126.77	−0.07	−0.04	−0.12	142	<0.01
DMI	2.63	3.80	4.75	−0.05	−0.05	−0.14	166	<0.01

¹For model including the error covariance structure.

for modeling individual deviations from the average lactation curve. The Lg4 function has an advantage compared with the AS function when only a few observations per cow are available, especially at the beginning of lactation (Silvestre et al., 2006). The data set of the present study contained several cows that remained in the production herd for an extended period and, hence, these cows had their first test-day record later in lactation. This might explain the advantage of the Lg4 function in modeling cow-specific lactation curves. However, the effect on modeling the average lactation curve may be negligible, as the AS function actually performed slightly better here. Nevertheless, an RR model using the Lg4 function for both the average and random lactation curve might be an alternative because of the simplicity in programming.

In addition, the comparison of RR (Lg4) models including and excluding the SP(POW) error covariance structure showed that a significantly better fit was achieved when the error covariance structure was present. This is in accordance with results of Hüttmann (2007) who also observed significant differences between RR models with and without an error covariance structure for daily feed intake and energy balance data. The assumption of Littell et al. (2006) that repeated daily measures such as those collected in the present study are not independent and thus autocorrelated was hence confirmed. The choice of model is crucial for statistical inference (van der Werf, 2001; Sawalha et al., 2005). In contrast to models without an error covariance structure, the significance of single fixed effects can be influenced by the model when an error covariance structure is included (Kramer et al., 2008b). However, the estimated correlations between adjacent repeated measures were estimated at marginal values (-0.04 to -0.14 for both traits). One might have expected positive instead of negative correlations. It could be argued that excessive eating and drinking on a certain day may result in cows consuming less food and water on the following day, although this is speculative.

The highest residual variance was estimated with observations from the third parity class (lactations 3–9) indicating that the model could not explain as much of the variation as for first- and second-parity cows. The (absolute value of the) correlation between the residuals of adjacent repeated measures in the third parity class were the highest, confirming that leftover covariance between residuals may increase with an inferior model fit.

Based on results of this study, the model with the best fit—RR (Lg4) including the error covariance structure SP(POW)—is recommended for the analysis of unequally spaced data sets of water and DMI to obtain valid statistical inference and correct variance

components (van der Werf, 2001; Sawalha et al., 2005; Rosário et al., 2007).

CONCLUSIONS

Mean residuals of water and DMI plotted against DIM showed clear differences in model performance between different FR functions. This was confirmed by the information criteria, which indicated that the best model fit for the fixed effect of DIM was attained using the AS function, and the best model fit for the random effect was achieved using the Lg4 function. Furthermore, application of a covariance structure was necessary. In conclusion, the model with the best fit RR (Lg4) is recommended for analyses of unequally spaced data sets of water intake and DMI including missing observations.

ACKNOWLEDGMENTS

This manuscript has benefited from the critical and helpful comments of 2 anonymous reviewers.

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.
- Brotherstone, S., I. M. S. White, and K. Meyer. 2000. Genetic modelling of daily milk yield using orthogonal polynomials and parametric curves. *Anim. Sci.* 70:407–415.
- Coffey, M. P., G. Simm, and S. Brotherstone. 2002. Energy balance profiles for the first three lactations of dairy cows estimated using random regression. *J. Dairy Sci.* 85:2669–2678.
- Collard, B. L., P. J. Boettcher, J. C. M. Dekkers, D. Petitclerc, and L. R. Schaeffer. 2000. Relationships between energy balance and health traits of dairy cattle in early lactation. *J. Dairy Sci.* 83:2683–2690.
- de Vries, M. J., S. van der Beek, L. M. T. E. Kaal-Lansbergen, W. Ouweltjes, and J. B. M. Wilmink. 1999. Modeling of energy balance in early lactation and the effect of energy deficits in early lactation on first detected estrus postpartum in dairy cows. *J. Dairy Sci.* 82:1927–1934.
- González, L. A., B. J. Tolkamp, M. P. Coffey, A. Ferret, and I. Kyriazakis. 2008. Changes in feeding behaviour as possible indicators for the automatic monitoring of health disorders in dairy cows. *J. Dairy Sci.* 91:1017–1028.
- Guo, Z., and H. H. Swalve. 1995. Modelling of the lactation curve as a sub-model in the evaluation of test day records. *Proc. Interbull Mtg. Prague, Czechoslovakia. International Bull Evaluation Service, Uppsala, Sweden. Interbull Bull.* 4–7.
- Hurvich, C. M., and C. L. Tsai. 1989. Regression and time series model selection in small samples. *Biometrics* 76:297–397.
- Hüttmann, H. 2007. Analyse der Futteraufnahme und der Energiebilanzmerkmale bei hochleistenden, erstlaktierenden Milchkühen. Schriftenreihe des Instituts für Tierzucht und Tierhaltung der Christian-Albrechts-Universität zu Kiel. Band 164.
- Hüttmann, H., E. Stamer, W. Junge, G. Thaller, and E. Kalm. 2009. Analysis of feed intake and energy balance of high yielding first lactating Holsteins with fixed and random regression models. *Animal* 3:181–188.
- Koenen, E. P. C., and R. F. Veerkamp. 1998. Genetic covariance functions for live weight, condition score, and dry-matter intake

- measured at different lactation stages of Holstein Friesian heifers. *Livest. Prod. Sci.* 57:67–77.
- Kramer, E., E. Stamer, K. Mahlkow, W. Lüppling, and J. Krieter. 2008a. Relationship between water intake, dry matter intake and daily milk yield on a German research farm. *Livest. Sci.* 115:99–104.
- Kramer, E., E. Stamer, J. Spilke, and J. Krieter. 2008b. Analysis of water intake, dry matter intake and daily milk yield using different error covariance structures. *Animal* 2:1585–1594.
- Littell, R. C., P. R. Henry, and C. B. Ammerman. 1998. Statistical analysis of repeated measures data using SAS procedures. *J. Anim. Sci.* 76:1216–1231.
- Littell, R. C., G. A. Milliken, W. W. Stroup, R. D. Wolfinger, and O. Schabenberger. 2006. Pages 174–203 in *User's Manual: SAS for Mixed Models*. SAS Institute Inc., Cary, NC.
- Liu, Y. X., J. Zhang, L. R. Schaeffer, R. Q. Yang, and W. L. Zhang. 2006. Optimal random regression models for milk production in dairy cattle. *J. Dairy Sci.* 89:2233–2235.
- López-Romero, P., and M. J. Carabaño. 2003. Comparing alternative random regression models to analyse first lactation daily milk yield data in Holstein-Friesian cattle. *Livest. Prod. Sci.* 82:81–96.
- Lukas, J. M., J. K. Reneau, and J. G. Linn. 2008. Water intake and dry matter intake changes as a feeding management tool and indicator of health and estrus status in dairy cows. *J. Dairy Sci.* 91:3385–3394.
- Macciotta, N. P. P., D. Vicario, and A. Cappio-Borlino. 2005. Detection of different shapes of lactation curve for milk yield in dairy cattle by empirical mathematical models. *J. Dairy Sci.* 88:1178–1191.
- Mielenz, N., J. Spilke, H. Krejcova, and L. Schüler. 2006. Statistical analysis of test-day milk yields using random regression models for the comparison of feeding groups during the lactation period. *Arch. Tierernähr.* 60:341–357.
- Murphy, M. R., C. L. Davis, and G. C. McCoy. 1983. Factors affecting water consumption by Holstein cows in early lactation. *J. Dairy Sci.* 66:35–38.
- Reents, R., J. Jamrozik, L. R. Schaeffer, and J. C. M. Dekkers. 1995. Estimation of genetic parameters for test day records of somatic cell score. *J. Dairy Sci.* 78:2847–2857.
- Rosário, M. F., M. A. N. Silva, A. A. D. Coelho, and V. J. M. Savino. 2007. Estimating and predicting feed conversion in broiler chickens by modeling covariance structure. *Poult. Sci.* 86:508–514.
- SAS Institute. 2005. *SAS/STAT User's Guide Version 9.0*. SAS Inc., Cary, NC, USA.
- Sawalha, R. M., J. F. Keown, S. D. Kachman, and L. D. van Vleck. 2005. Evaluation of autoregressive covariance structures for test-day records of Holstein cows: Estimates of parameters. *J. Dairy Sci.* 88:2632–2642.
- Schaeffer, L. R. 2004. Application of random regression models in animal breeding. *Livest. Prod. Sci.* 86:35–45.
- Schaeffer, L. R., and J. C. M. Dekkers. 1994. Random regressions in animal models for test-day production in dairy cattle. Pages 443–446 in *Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Canada. Communication No. 18*.
- Schwarz, G. 1978. Estimating the dimension of a model. *Ann. Stat.* 6:461–464.
- Silvestre, A. M., F. Petim-Batista, and J. Colaco. 2006. The accuracy of seven mathematical functions in modeling dairy cattle lactation curves based on test-day records from varying sample schemes. *J. Dairy Sci.* 89:1813–1821.
- van der Werf, J. 2001. Random regression in animal breeding. In *Course notes. Jaboticabal, SP Brazil, November 2001*. <http://www-personal.une.edu.au/~jvanderw/CFcoursenotes.pdf> Accessed Nov. 28, 2008.
- Veerkamp, R. F., and R. Thompson. 1999. A covariance function for feed intake, live weight, and milk yield estimated using a random regression model. *J. Dairy Sci.* 82:1565–1573.
- VIT. 2008. Zuchtwertschätzungsbeschreibung0808. http://www.vit.de/fileadmin/user_upload/vit-fuers-Rind/zuchtwertschaetzung/milchrinder-zws-online/Zws_Bes0808_eng.pdf Accessed Oct. 17, 2008.
- 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.
- Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164–165.
- Woodford, S. T., M. R. Murphy, and C. L. Davis. 1984. Water dynamics of dairy cattle as affected by initiation of lactation and feed intake. *J. Dairy Sci.* 67:2336–2343.