Studies on test-day and lactation milk, fat and protein yield of dairy cows

Promotor: dr.ir. R.D. Politiek, hoogleraar in de veeteeltwetenschappen
Co-promotor: dr.ir. J. Dommerholt, directeur informatieverwerking
Nederlands Rundvee Syndicaat

J.B.M. Wilmink

STUDIES ON TEST-DAY AND LACTATION MILK, FAT AND PROTEIN YIELD OF DAIRY COWS

Proefschrift
ter verkrijging van de graad van
doctor in de landbouwwetenschappen,
op gezag van de rector magnificus,
dr. C.C. Oosterlee,
in het openbaar te verdedigen
op vrijdag 20 maart 1987
des namiddags te vier uur in de aula
van de Landbouwuniversiteit te Wageningen.

STELLINGEN

1. Voor een nauwkeurige voorspelling van de melkproduktie is het gewenst te beschikken over de geschatte gemiddelde produktie van koeien die lacteren op hetzelfde bedrijf en die op gelijke leeftijd en in dezelfde maand afkalven.

Dit proefschrift.

2. Selectie op het lactatiedeel tussen 90 en 180 dagen na afkalven van vaarzen leidt tot een grotere genetische vooruitgang per jaar in 305-dagenproduktie dan directe selectie.

Dit proefschrift.

3. Bij het extrapoleren van deellactaties in de fokwaardeschatting voor stieren hoeft geen onderscheid gemaakt te worden tussen lopende en afgebroken lactaties.

Dit proefschrift.

4. Doordat geëxtrapoleerde deellactaties meegenomen worden bij het schatten van fokwaarden voor stieren, is het publiceren van dochtergemiddelden van 100-dagenprodukties verwarrend en dient achterwege te blijven.

Dit proefschrift.

5. Dat K.I.-verenigingen in het kader van kostenbesparingen stoppen met het verzamelen van gegevens over voederconversie en groei van centraal opgefokte stieren voor selectiedoeleinden, kost de boer geld.

6. Het toepassen van embryo-transplantatie dreigt een dans om het gouden kalf te worden.

7. De extra opbrengsten door het verhogen van de produktie per koe nemen af met het toenemen van de gemiddelde produktie per bedrijf.

8. De grote belangstelling voor het gebruik van somatotropine om de melkgift te stimuleren staat in geen verhouding tot de beperkte bedrijfseconomische betekenis ervan.

9. De gepubliceerde fokwaarden voor exterieurkenmerken nemen toe in waarde wanneer met behulp van selectie-indextheorie de fokwaarden voor lineair gescoorde onderdelen worden samengevat in de bovenbalk.

10. Gezien de toenemende behoefte aan het beoefenen van duursporten zijn douches in kantoorgebouwen geen luxe.

11. Het beleid van de KNAU om de recreatiesport te stimuleren dient ook afgestemd te zijn op investeringen in de kweek van topatleten.

12. De verpleegkundige is het oog en oor van de arts.

Proefschrift van J.B.M. Wilmink
Studies on test-day and lactation milk,
fat and protein yield of dairy cows.
Wageningen, 20 maart 1987.

Aan mijn moeder en wijlen vader
Aan Hermine en Anne

Wilmink, J.B.M., 1987. Studies on test-day and lactation milk, fat and protein yield of dairy cows (Onderzoek naar melk-, vet- en eiwitprodukties per dag en tijdens de lactatie bij melkkoeien). Doctoral thesis, Royal Dutch Cattle Syndicate, Arnhem.

VOORWOORD

Dit proefschrift bestaat uit een bundeling van 5 artikelen. Het onderzoek werd uitgevoerd bij het Instituut voor Veeteeltkundig Onderzoek te Zeist en bij het Nederlands Rundvee Syndicaat te Arnhem.

Mijn promotor prof. dr. ir. R.D. Politiek dank ik voor zijn voortdurende belangstelling tijdens het onderzoek. De stimulans die van hem uitging en zijn vertrouwen in het onderzoek heb ik erg gewaardeerd. Mijn copromotor dr. ir. Jan Dommerholt dank ik voor zijn intensieve betrokkenheid. Met name zijn frisse kijk op een aantal vraagstukken heeft indruk op mij gemaakt.

Mede dank zij de inbreng van dr. ir. Johan van Arendonk en dr. ir. Jan Jansen in de leescommissie is dit onderzoek tot een goede afronding gekomen. De discussiepunten die jullie aandroegen hebben zeker geleid tot een aanscherping van het onderzoek en de artikelen. Onze besprekingen heb ik steeds als zeer prettig ervaren.

De directie van het I.V.O. ben ik zeer erkentelijk voor de gastvrijheid die zij mij geboden heeft. Het bestuur en de directie van het N.R.S. dank ik voor de ruimte en de financiële ondersteuning die zij mij gegeven hebben om tot een goede afronding van dit proefschrift te komen.

Een woord van dank gaat uit naar ing. Hans Kooper voor zijn hulp bij het verzamelen van de gegevens en naar Annet Pop voor de snelle en zorgvuldige wijze waarop zij het typewerk verzorgde. Haar geduld is herhaaldelijk op de proef gesteld.

Tenslotte wil ik iedereen bedanken die op een of andere wijze bij het tot stand komen van dit proefschrift waren betrokken, maar in dit voorwoord niet genoemd zijn.

CONTENTS

INTRODUCTION

In The Netherlands, about 1.8 million or about 77% of total population of dairy cows are recorded (NRS Jaarverslag, 1985). Individual milk samples are taken at 3 or 4 week intervals and results on milk, fat and protein yield are sent to dairy farmers. On completion of a cow's lactation, farmers receive a summary with
1) all test-day yields in her last lactation,
2) cumulative 305-day and total lactation yields for all her completed lactations and 3) her predicted breeding value.

Because almost 80% of farmer's income originates from milk production, data from milk recording forms the basis to control herd management and genetic improvement of cows. With respect to management, a lactation index is given to the farmer in The Netherlands that is based on predicted 305-day milk, fat and protein yield, standardized for age and season of calving. The lactation index ranks cows within herd phenotypically on future production capacity and is, among others, used to decide whether to inseminate a cow. In addition a herd index is computed for the whole herd by averaging test-day milk yields, standardized for age, season of calving and stage of lactation. The herd index controls feeding and management of the dairy herd.

Van Arendonk (1985) demonstrated the need to account for age at calving and time in lactation when using the lactation index to decide whether a cow should be inseminated. New guides were developed to decide about whether a cow should be inseminated and about optimal replacement of cows. These guides represent expected profit in case of conception and future profitability respectively and are based on predicted milk production of a cow. Adjustment factors for age and season of calving are involved in calculating these guides.

With regard to genetic improvement, breeding values of bulls and cows are predicted for 305-day milk, fat and protein yield in first lactation. Breeding values should be predicted as early as possible to keep generation interval short and to reduce prediction bias by selection (Danell, 1982).

The basis for calculating management guides and predicting breeding values as early as possible is correct adjustment of milk, fat and protein yield for age and season of calving and prediction of future yields. Because level of production has increased dramatically in the last decade, new adjustment factors may be required. In addition, it was necessary to have more precise adjustments and a procedure by which these adjustment factors can be updated easily if production level changed. These issues have not been investigated thoroughly although a number of studies have been carried out with regard to adjustment of milk and fat yield. Moreover, results with regard to adjustment of protein yield are very scarce in literature.

Prediction of 305-day yield can be improved by using correct averages within classes of herd by age by season of calving (Van Arendonk and Fimland, 1983). Methods of estimating these averages, have received no attention.

Extended or part lactation yields have been used to predict breeding values for early selection of bulls (Danell, 1982). However, a number of other information sources in first lactation are also available to allow early selection. The aspect of which and how information sources in first lactation should be used to predict breeding values for 305-day milk, fat and protein yield as early as possible with high accuracy and low probability of bias by selection of heifers has received little attention in the literature.

This study was to improve procedures for correcting milk, fat and protein yield records, to improve prediction of 305-day yield and to develop a procedure to predict breeding values for 305-day milk, fat and protein yield in first lactation as early as possible.

In chapter 1, the adjustment of 305-day yields for age at calving is treated. The use of mathematical functions and the relation of age correction factors to level of production will be discussed. Chapter 2 deals with use of functions for adjustment of test-day yields for age at calving and stage of lactation. Prediction of 305-day yield and estimation of means will be studied in chapter 3. In chapter 4, genetic parameters for cumulative yield in different intervals in first lactation and predicted 305-day yield, and their relationship with 305-day yield in first lactation will be presented. Annual responses in 305-day yield, when selection is on alternatives for 305-day yield, will be discussed. In chapter 5, the effect of including incomplete first lactation records on genetic parameters and on extension of part lactation records will be studied. A general discussion and a comparison of predicted breeding values for the different alternatives will be presented in chapter 6.

REFERENCES

Danell, B. (1982). Studies on lactation yield and individual test-day yields of Swedish dairy cows. 4. Extension of part-lactation records for use in sire evaluation. Acta Agric. Scand. 32:103-114.

NRS (1985). Jaarverslag 1985 van het Koninklijk Nederlands Rundvee Syndicaat, Arnhem.

Van Arendonk, J.A.M. (1985). Studies on the replacement policies in dairy cattle. Ph.D. Thesis, Agric. Univ. Wageningen.

Van Arendonk, J.A.M. and E. Fimland (1983). Comparison of two methods to extend partial milk records. Z. Tierz. u. Züchtungsbiol. 100:33-38.

Chapter 1

ADJUSTMENT OF LACTATION YIELD FOR AGE AT CALVING IN RELATION TO LEVEL OF PRODUCTION

J.B.M. Wilmink

Royal Dutch Cattle Syndicate
P.O. Box 454, 6800 AL Arnhem

Accepted for publication in Livestock Production Science

ABSTRACT

The influence of age at calving on 305-day milk, fat, and protein yield in relation to herd and population level of production was investigated and a procedure for age adjustment was derived.
Age differences were estimated by a repeatability model for 49669 complete 305-day lactation records (kg milk, kg fat and kg protein) from 27965 Dutch Friesian cows calving from 1-6-1979 until 31-5-1982. Several mathematical functions were applied to the estimated age differences. The best fit was by a second degree polynomial for lactations 1 and 2 separately ($R^2 \leq 0.995$) and a linear function including the log of age at calving for third and later lactations ($R^2 = 0.967$).

One set of adjustment factors was found to apply to all herd levels of production. However, age adjustment factors were different from those found by Dommerholt (1975) where the population level was about 1000 kg milk lower. A more accurate age adjustment would be obtained by updating the mean production of heifers at 24 months of age and using mathematical functions for computation of age factors.

INTRODUCTION

Apart from genetic influences, a cow's 305-day milk yield is influenced within herds by age at calving. Age explains from 20 to 40 % of the total sums of squares as reported by Auran (1973), Dommerholt (1975) and Cooper and Hargrove (1982). After age and herds, 305-day milk yield is affected by month of calving (about 1 %), number of days open (about 3 %) (Danell, 1982), preceding dry period (about 1 %) (Witt et al., 1969; Hoque and Hodges, 1980) and interactions like age * season and year * season. Therefore correct age adjustment factors for milk, fat and protein yield should be available if 305-day records are to be preadjusted for use in the genetic evaluation of bulls and cows or comparison of cows within a herd.

In The Netherlands multiplicative adjustment factors for age and season at calving, developed by Dommerholt et al. (1977) are still in use. With regard to age factors, Dommerholt (1975) showed a relationship with herd production level, which resulted in one set of factors for different herd levels. No interaction was found between season of calving and herd level or age at calving.
In the years 1975 to 1984, the average production of the Dutch Black-and-White population increased by 724 kg (jaarverslag NRS, 1985). New multiplicative age factors and additive season factors were shown to be required (Wilmink, 1985). The need for continual updating of correction factors was stressed by Keown and Everett (1985). Multiplicative factors, however, are dependent on level of production.

The objective of this study was to investigate the influence of age at calving on the 305-day milk, fat and protein yield in relation to both herd and population level and to present a procedure for age adjustment.

MATERIAL

After screening of milk recording files, 49,669 complete lactations from 27,965 Dutch Friesian cows, calving from 1-6-1979 until 31-5-1982 in 732 herds drawn at random, were extracted (Data set I). Screening included at least 280 days in lactation, at least 22 months old at calving and at least 4 paternal half sisters per herd level in order to avoid disconnectedness between age and sires. Furthermore cows had to remain in one herd, to allow for absorption of cows and herds in the mixed model equations and were not required to have a first lactation. All lactations longer than 304 days had a known 305-day milk, fat and protein yield while lactations shorter than 305 days were extrapolated to a 305-day equivalent using the formula proposed by Poutous and Mocquot (1975). Herds were split into low (< 5250 kg), medium (5250-5750 kg) and high (> 5750 kg) level groups according to the average 305-day milk yield per herd in the above calving period. In

Table 1. Overall means and standard deviations for 305-day yield

	dataset I		dataset II	
	mean	SD	mean	SD
kg milk	5,813	1,015	5,935	1,098
kg fat	244	43	251	47
kg protein	197	34	200	36

calculating this average the records were adjusted for age at calving with factors determined in a previous study (Wilmink, 1985). Grouping was such that 50% of herds were in the medium level, resulting in 10544, 26060 and 13055 records for the low, medium and high level groups, respectively. Groupings for milk yield were also used for fat and protein yields. Data set I was used for the study of the relationship of age at calving with herd and population levels of production.

To study the validity of the adjustment procedure, data set II was formed consisting of 305-day yields from lactations, 280 to 310 days in milk, for 5108 Dutch Friesian cows in 227 herds, randomly chosen from the above 732 herds, and calving between 1-6-1982 and 31-5-1983. No restrictions were made with respect to age at calving and sire identification. Again herds were split into 3 levels (< 5400 kg; 5400 - 6200 kg (50%); > 6200 kg) based on the average age corrected milk yield of all cows in the herd. Numbers of records per herd level were 909, 2887 and 1312 respectively. Overall means and standard deviations for 305-day kg milk, kg fat and kg protein in both data sets are in table 1.

METHODS

Estimation of age and season effects

Age and season effects were estimated in data set I within herd level and trait with a repeatability model (Henderson ,1975a) allowing a simultaneous estimation of fixed and sire and cow effects:

$$y_{ijklmnop} = \mu_i + hy_{ij} + a_{ik} + m_{il} + l_{im} + s_{in} + c_{ino} + e_{ijklmnop} \quad (1)$$

where y = 305 day kg milk, fat or protein;

μ_i = overall mean for herd level i;

hy_{ij} = fixed effect of herd-year j (j = 1, 2196);

a_{ik} = fixed effect of age (k = 1, 17);

m_{ik} = fixed effect of month of calving (l = 1, 12);

l_{im} = fixed effect of lactation length (m = 1, 10);

s_{in} = random effect of sire n with expectation zero and variance equal to σ^2_s;

c_{ino} = random effect of cow o with expectation zero and variance equal to σ^2_{ci};

$e_{ijklmnop}$ = residual with expectation zero and variance equal to σ^2_e.

All covariances between sires, cows and residuals were zero. Further covariances among cows and among residuals were zero. Sires were related by their sire and maternal grandsire. Years were defined from June to May. Age classes 1 to 17 were defined for 22-23, 24-25, 26-27, 28-29, 30-31, 32-34, 35-36, 37-38, 39-40, 41-42, 43-44, 45-50, 51-56, 57-68, 69-72, 93-105 and > 105 months for age at calving respectively.
With regard to lactation number, age classes 1 to 5, 6 to 11 and 12 to 17 were considered as lactation number 1, 2 and greater than 2, respectively. Lactation length was divided into 10 classes as follows: 280-299, 300-319, 320-339, 340-359, 360-379, 380-399, 400-439, 440-479 and more than 479 days. The tenth class consisted of cows removed from the herd after their lactation, and considered to be non-pregnant.

The effects included in equation (1) were a result of a pilot study, (Wilmink, 1985) for which age * month of calving was found to be non-significant.

Age and month effects were estimated by solving the equations:

$$\begin{pmatrix} X_1'X_1 & X_1'X_2 & X_1'Z & X_1'W \\ & X_2'X_2 & X_2'Z & X_2'W \\ \text{Symm.} & & Z'Z + kA^{-1} & Z'W \\ & & & W'W + tI \end{pmatrix} \begin{pmatrix} \hat{hy} \\ \hat{f} \\ \hat{s} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} X_1'Y \\ X_2'Y \\ Z'Y \\ W'Y \end{pmatrix} \quad (2)$$

where X_1, X_2, Z and W are known incidence matrices and $\hat{hy}$, $\hat{f}$, $\hat{s}$, $\hat{c}$ are vectors of solutions for $\mu_i + hy_{ij}$, other fixed effects, sires and cows respectively. A is a matrix of relationships based upon sire and maternal grandsire (Henderson, 1975b) and I is the identity matrix. In calculating k (σ_e^2 / σ_s^2) and t (σ_e^2 / σ_c^2) a heritability and repeatability of 0.25 and 0.53 respectively, was used for all traits in all herd levels.

The restrictions $\sum_{k=1}^{17} \hat{a}_{ik} = \sum_{l=1}^{12} \hat{m}_{il} = \sum_{m=1}^{10} \hat{l}_{im} = 0$ were imposed.

Data were ordered on herd-years, sires within herd years, and cows within herd-years and sires for absorption of the **hy** and **c** equations. Per cow, $S_w = I - W(W'W + tI)^{-1}W'$ was first computed followed by $X_1'S_wX_1$, $X_1'S_wX_2$, $X_1'S_wZ$, $X_1'S_wY$, $X_2'S_wX_2$, $X_2'S_wZ$, $X_2'S_wY$, $Z'S_wZ$ and $Z'S_wY$.

The matrices $X_1'S_wW_1$ to $Z'S_wY$ were then accumulated over the cows within a herd. After a herd was processed the 3 year equations were absorbed into the **f** and **s** equations resulting in:

$$\begin{pmatrix} X_2'HX_2 & X_2'HZ \\ Z'HX_2 & Z'HZ + kA^{-1} \end{pmatrix} \begin{pmatrix} \hat{f} \\ \hat{s} \end{pmatrix} = \begin{pmatrix} X_2'HY \\ Z'HY \end{pmatrix} \quad (3)$$

where $H = S_w - S_wX_1 (X_1'S_wX_1)^{-1} X_1'S_w'$.

The matrices in (3) were accumulated over herds. Approximate effective numbers for the effects in **f** were on the diagonal of $X_2'HX_2$. After obtaining solutions for **f** and **s**, the **hy** equations were solved, giving an estimate for μ_i by the restriction $\sum_{j=1}^{2196} \hat{hy}_{ij} = 0$

Relationship between age differences and age, herd and population level

For the influence of age at calving on 305-day milk, fat and protein yield, three equations were fit to the estimated age differences $\hat{a}_{ik} - \hat{a}_{i2}$ (k = 1, 17):

$$ac_I = a + bx + cx^2 \qquad (4)$$

$$ac_{II} = a + bx + c\ln x \qquad (5)$$

$$ac_{III} = a + bx + cx^2 + dx^3 \qquad (6)$$

where ac_I, ac_{II}, ac_{III} = age differences for 305-day yield with respect to 24 months of age,
x = age in months, calculated as the weighted average age within an age class,
a, b, c, d = parameters
ln = natural logarithm

The parameters a to d were estimated by ordinary least squares. Equations (4) and (5) were used for lactation numbers 1, 2 and 3 and higher, and equations (4) to (6) for all parities. Comparison criteria were the coefficient of determination (R^2) and mean square error (MSE). The function giving the best fit to all age differences (ac_f) was then used.

Age adjustment factors were calculated per month of age by $(\mu_i + ac_f)/(\mu_i + ac_{fr})$ where ac_f was fit within herd level and ac_{fr} was the difference at reference age, chosen to be 96 months, the base age in The Netherlands. Adjustment factors for different herd levels were compared. For the relationship with population level which may change with time, the age differences calculated by ac_f per herd level were compared with the differences found by Dommerholt (1975). In Dommerholt's study, age differences were estimated in a fixed model, as given in equation (7).

Validity of adjustment

The validity of age adjustment was tested in data set II by a fixed model, using LSML76 (Harvey, 1977):

$$y^*_{iklm} = \mu + hl_i + a_k + m_l + (a * hl)_{ikn} + e_{klm} \qquad (7)$$

Here y^*_{iklm} represents 305-day kg milk, kg fat and kg protein adjusted for age, and hl is herd level. Criteria for comparison were the F-variable and the least square constants if an effect was significant.

Table 2. The estimated differences ($\hat{a}_{ik} - \hat{a}_{i2}$) between ages at calving for 305-day milk yield (kg) with effective number of records (n_e)

	herd level					
	low		medium		high	
age class	n_e	$\hat{a}_{ik} - \hat{a}_{i2}$	n_e	$\hat{a}_{ik} - \hat{a}_{i2}$	n_e	$\hat{a}_{ik} - \hat{a}_{i2}$
1	314	-115	696	-116	405	-112
2	655	0	1,763	0	880	0
3	324	170	797	156	409	122
4	134	295	326	327	155	312
5	74	434	164	421	67	514
6	78	643	173	714	120	818
7	478	713	1,221	840	674	953
8	455	862	1,146	938	557	1,045
9	225	912	549	1,090	243	1,194
10	113	1,035	295	1,161	122	1,353
11	71	1,086	169	1,216	74	1,429
12	773	1,248	1,939	1,415	1,010	1,550
13	352	1,419	943	1,557	399	1,716
14	766	1,528	2,036	1,653	1,010	1,797
15	676	1,600	1,762	1,707	881	1,925
16	201	1,512	453	1,613	252	1,885
17	128	1,335	267	1,459	148	1,810

RESULTS

Equations fit to the age differences

Age differences ($\hat{a}_{ik}$ - $\hat{a}_{i2}$) with the effective number of records per age class are given for milk yield (table 2). At mature age the differences from 24 months were 1600, 1707 and 1925 kg milk for low, medium and high herd level respectively. Level of production for age class 2 was 4225, 4798 and 5414 kg milk respectively. Equations (4), (5) and (6) were fit to the age differences for the medium herd level. Results are in table 3. Considering the fit of ac_I and ac_{II} within lactation number 1, 2 and 3 and higher, R^2 ranged from 0.899 to 0.996 and MSE from 345.4 to 2160.8 kg^2. Function ac_I gave the best fit to the age differences in parity 1 and 2 and ac_{II} to the age differences in later parities. When ac_I, ac_{II} and ac_{III} were fit to the age differences over all lactations, the best fit was obtained by ac_{III} with 0.998 and 1055.5 kg^2 for R^2

Table 3. The coefficient of determination (R^2) and mean square error (MSE) for fitting ac_I, ac_{II}, ac_{III}[a] to the age differences (305-day milk yield) for medium herd level

lactation number	equation	R^2	MSE
1	ac_I	0.996	429.5
2	ac_I	0.995	345.4
3 and higher	ac_I	0.899	2160.8
1	ac_{II}	0.995	469.5
2	ac_{II}	0.994	387.7
3 and higher	ac_{II}	0.967	708.1
all	ac_I	0.956	18586.6
all	ac_{II}	0.995	2191.5
all	ac_{III}	0.998	1055.5

a) $ac_I = a + bx + cx^2$
$ac_{II} = a + bx + c\ln x$
$ac_{III} = a + bx + cx^2 + dx^3$
where x = age in months
a,b,c,d = parameters
ln = natural logarithm

and MSE respectively. When function ac_{III} for all lactations was compared with a function, called ac_f and consisting of ac_I for lactation 1 and 2 separately and ac_{II} for later parities, the mean square error was 1055.5 and 547.7 kg^2 for ac_{III} and ac_f respectively. The average absolute standard error was 22.0 and 13.8 kg respectively with maximum values of 66 and 26 kg milk. The closest fit of the age differences was therefore by ac_f.

In figure 1 ac_f is plotted against the age differences derived from model 1. A very close fit of ac_f to the age differences is demonstrated for all traits. The function ac_f was chosen for future use.

Relationship with herd level

Parameters for ac_f per herd level are in table 4 for milk yield. The pattern of age differences within lactation 1 and 2 was convex, except for lactation 1 and high herd level. This was mainly caused by the differences $\hat{a}_{31}$ - $\hat{a}_{32}$ and $\hat{a}_{33}$ - $\hat{a}_{32}$ (table 2), which were smaller for high than for medium and low herd levels.

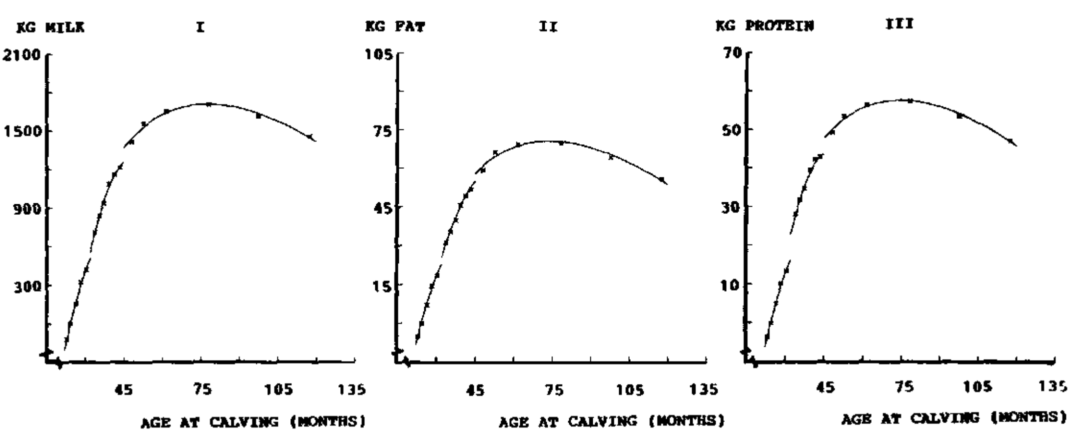

Figure 1 The relationship between age differences for kg milk (I), fat (II) and protein (III) and age in months.

Table 4. The estimated parameters in $ac_I = a + bx + cx^2$ for lactation number 1 and 2 and $ac_{II} = a + bx + c\ln x$ for lactation number ≥3 in fitting age differences for 305 day milk yield (kg)

herd level		lactation 1	lactation 2	lactation 3
low	a	- 2723	- 3133	- 8666
	b	145.3	161.3	- 38.52
	c	- 1.365	- 1.476	3047
medium	a	- 3300	- 5130	- 7335
	b	188.5	265.6	- 34.62
	c	- 2.168	- 2.748	2696
high	a	560	- 2707	- 6292
	b	- 113.2	135.1	- 27.31
	c	3.677	- 0.912	2372

A reason for this may be better rearing of heifers in high level herds. On the other hand, age class 5 could contain some second parity cows. If lactation number was greater than 1, the pattern of ac_f was in all cases convex upward.

Multiplicative age adjustment factors for 305-day milk are presented in table 5. These factors were computed using ac_f and the average production of heifers in age class 2 within herd level. At the medium level, age adjustment factors increased from 0.717 up to 1.000 for cows of about 7.5 years old. Age factors for low and high herd levels differed from the medium level by only 0.017 at the maximum, if cows had not reached mature age. According to Lush and Shrode (1950), one set of multiplicative age factors for all levels would be sufficient. Thus, similar to Dommerholt (1975) age differences were proportional to herd level of production. Parameters for ac_f for 305-day kg fat and kg protein are in table 6. The level at 24 months of age is given so that age factors can be computed.

Table 5. Multiplicative age adjustment factors for 305-day milk yield per herd level

Age class	herd level low	herd level medium	herd level high	maximum difference from medium
1	0.704	0.717	0.724	0.013
2	0.726	0.739	0.738	0.012
3	0.751	0.762	0.757	0.011
4	0.776	0.784	0.781	0.008
5	0.799	0.803	0.809	0.006
6	0.833	0.846	0.850	0.013
7	0.851	0.867	0.867	0.016
8	0.868	0.882	0.885	0.014
9	0.884	0.901	0.902	0.017
10	0.899	0.915	0.919	0.014
11	0.912	0.925	0.936	0.013
12	0.945	0.959	0.954	0.014
13	0.962	0.972	0.968	0.010
14	0.985	0.990	0.984	0.006
15	1.000	1.000	1.000	0.000
16	0.988	0.988	0.999	0.011
17	0.952	0.960	0.984	0.024

Relationship with population level

Age differences for 305-day kg milk with respect to 23.5 month of age at calving were computed with ac_f per herd level and compared with the differences found by Dommerholt (1975). In that study the contrasts were based on 4000 lactation records of Dutch Friesian cows, calving from 1-7-1970 until 30-6-1971. Level of production was about 1000 kg lower (table 7) than in the current study.

For medium level herds, the maximum age difference was 1765 kg in both studies, but mature age was reached earlier in the present study. For cows, varying in age from about 25 - 60 months, the differences were about 150 to 250 kg higher, with an exception for

Table 6. Parameters in the function ac_f[a)] describing the pattern of age differences with regard to 24 month age for 305-day milk, fat and protein yield (all kg) and medium herd level

	lactation	a	b	c
kg milk	1	- 3300	188.53	- 2.168
	2	- 5131	265.63	- 2.748
	≥3	- 7336	- 34.62	2696
kg fat	1	- 147.88	8.509	- 0.099
	2	- 158.33	8.210	- 0.077
	≥3	- 330.36	- 1.645	121.46
kg protein	1	- 97.22	5.435	- 0.059
	2	- 166.01	8.933	- 0.095
	≥3	- 240.35	- 1.202	89.88
level at 24 months		4798	201	162

a) $ac_f = a + bx + cx^2$ if lactation nr = 1,2.
$ac_f = a + bx + c\ln x$ if lactation nr ≥ 3.
x = age in months.
ln = natural logaritm

cows around 31 months old. The differences between both studies for low level herds were as for medium level herds. In high level herds, Dommerholt (1975) found age differences which were about 100 - 200 kg higher than in the present study for cows older than 60 months.

With respect to age adjustment factors, these increased in 10 years with a factor varying from 0.068 to 0.021 (table 7), indicating that age differences were not proportional to population level. When level of production was updated and Dommerholt's differences were used for computing age factors, differences with the age

Table 7. Age differences and adjustment factors for 305-day milk yield (kg) found in this study with ac_f[a] and by Dommerholt (1975)

	Dommerholt (1975) herd level			This study herd level			Adjustment factors[b]		
average age (months)	low	medium	high	low	medium	high	I	II	III
23,5	0	0	0	0	0	0	0.687	0.730	0.736
26	-30	64	198	194	203	172	0.699	0.740	0.767
31	379	568	711	531	528	654	0.788	0.817	0.817
38	706	800	1064	931	1062	1177	0.829	0.853	0.899
44,5	926	1093	1605	1188	1315	1566	0.881	0.897	0.938
52	1216	1437	1742	1434	1584	1730	0.942	0.950	0.979
71	1412	1687	2069	1651	1765	1950	0.986	0.988	1.007
95	1669	1765	2208	1614	1719	1986	1.000	1.000	1.000
mean[c]	3479	3884	4178	4214	4785	5425			

a) see table 6.

b) For medium herd level

I : Computed from Dommerholt's differences and level of 3884 kg milk.

II : Computed from Dommerholt's differences and level of 4785 kg milk.

III: Computed from the difference in the present study and level of 4785 kg milk.

c) at 23.5 months for age at calving

factors computed from data set I varied from 0.000 to 0.027 for heifers (up to 33 months old) from 0.041 to 0.046 for second parity cows (34 - 46 months old) and from 0.029 to 0.019 for cows aged from 47 to 80 months. As a result, Dommerholt's factors were closer to the age factors derived from data set I, when level of production was updated.

Validity of adjustment

Lactation records in dataset II were adjusted for age and month of calving. Age factors per month were calculated with ac_f and level at 24 months as given in table 6. After adjustment, age and herd-level * age were not significant (P < 0.05). Although not significant, least squares constants for age showed a slightly higher production for young cows when adjusted to mature age (30 kg with respect to overall), which could be explained by genetic trend.

DISCUSSION AND CONCLUSIONS

The disagreement between the age differences, found to exist in Dommerholt (1975) and the current study, revealed that, relative to mature cows, heifers produce more now than they did in the early seventies. Reasons for this may be better rearing of heifers and/or selection on first lactation yield. Breeding and rearing programmes have been emphasized in farmers' educational/ advisory programmes in The Netherlands.

Adjustment for selection bias may not be complete because all records for some cows in data set I were not available (Henderson et al., 1959, Thompson, 1979). However, selection bias will be reduced when using a repeatability model instead of the gross and paired comparison method, presented by Lush and Shrode (1950). Gross and paired comparison methods have been applied for the most part in literature. The repeatability model used was a mixture of gross and paired comparison methods. Age factors tend to be biased upward in the gross comparison by r(m'-m) and to be biased downward in the paired comparison by (1-r)(m'-m), where r is the repeatability between lactations and m,m' are means of lactations before and after selection (Henderson et al., 1959). An analysis of 305-day yield for the high level herds in data set I, adjusted for half the sires breeding values, (computed from best linear unbiased predictions in the May 1983 run), with computer programme LSML76 (Harvey 1977) and a model containing fixed effects for herd, year, age and month of calving, did not yield other age differences.

Age differences were proportional to herd level of production. As a

consequence, one set of age factors can be used for different herd levels, which is in agreement with earlier studies (Miller, 1973; Dommerholt, 1975). With regard to population level, no proportionality was found. Cooper and Hargrove (1982) found relatively small differences ($\leq$ 0.05) between their age factors and those reported by Norman et al. (1974).
Difference in population level was about 200 kg. Keown and Everett (1985) reported that age factors of cows freshening during May and June tended to be lower and during the remaining months higher than the factors reported by Miller et al. (1970). An update of average production for heifers, aged 24 months, in calculating age adjustment factors may be worthwhile.

Using the age classes defined by Dommerholt (1975) i.e. 22-24, 25-28, 29-35, 36-41, 42-47, 48-59, 60-71, 72-97, 98-132 months for class 1 - 9 respectively, age adjustment with Dommerholt's factors (table 7) resulted in a highly significant 27.8 F-value for the effect of age in model (7) for data set II.
When the average production of 24 month old heifers was updated, the F-value decreased to 10.6 but age was still significant. Cows in age class 1 - 5 were strongly favoured (+ 250 kg), whereas cows in age classes 4 and 5 were over evaluated (+ 130 kg). This means that an increase in accuracy of age adjustment can be obtained by updating production level of heifers. However, age factors need to be estimated, for example every 10 years, because the age production curve can change due to breeding and/or management/feeding.

For practical purposes age adjustment factors can be computed by $(\mu_i + f(x))/(\mu_i + f(x_r))$ where μ_i is production level of heifers aged 24 months at calving and f(x) is a function of age in months as given in table 6. Base age is represented by x_r. Age factors can be updated by changing μ_i, if production level of heifers has increased.

REFERENCES

Auran, T. (1973). Studies on monthly and cumulative monthly milk yield records. 1. The effect of age, month of calving, herd and length of the first period. Acta Agr. Scand. 24; 339-348.

Cooper, J.B. and G.L. Hargrove (1982). Age and month of calving adjutments of Holstein protein, milk, and fat lactation yields.
J. Dairy Sci. 65: 1673-1678.

Danell, B. (1982). Studies on lactation yield and individual test-day yields of Swedish dairy cows. 1. Environmental influence and development of adjustment factors. Acta Agr. Scand. 32: 65-81.

Dommerholt, J. (1975). Correctie van de melkgift van koeien voor verschillen in leeftijd, seizoen en lactatiestadium. (with English summary). Agriculture research reports 844, Centre for Agricultural Publishing and Documentation, Wageningen.

Dommerholt, J.; S.R. Sijbrandij and W.M.G. Wismans (1977). De individuele standaardkoeproduktie en de lactatiewaarde. Bedrijfsontwikkeling 8: 319-321.

Harvey, W.R. (1977). User's guide for LSML76, a mixed model least-squares and maximum likelihood computer program. Mimeo, Ohio State University.

Henderson, C.R. (1975a). Use of all relatives in intraherd prediction of breeding values and producing abilities. J. Dairy Sci. 58: 1910-1916.

Henderson, C.R. (1975b). Inverse of a matrix of relationships due to sires and maternal grandsires. J. Dairy Sci. 58: 1917-1920.

Henderson, C.R.; O. Kempthorne; S.R. Searle and S.M. von Krosigh (1959). The estimation of environmental and genetic trends from records subject to culling. Biometrics 15: 193-218.

Hoque, M. and J. Hodges (1980). Genetic and phenotypic parameters of lifetime production traits in Holstein cows. J. Dairy Sci. 63: 1900-1910.

Keown, J.F. and R.W. Everett (1985). Age-month adjustment factors for milk, fat, and protein yields in Holstein cows. J. Dairy Sci. 68: 2664-2669.

Lush, J.L. and R.R. Shrode (1950). Changes in milk production with age and milking frequency. J. Dairy Sci. 33: 338-357.

Miller, P.D. (1973). A recent study of age adjustment. J. Dairy Sci. 56: 952-958.

Miller, P.D., W.E. Lentz and C.R. Henderson (1970). Joint influence of month and age of calving on milk yield of Holstein cows in the Northeastern United States. J. Dairy Sci. 53: 351-357.

NRS (1985). Jaarverslag 1985 van het Koninklijk Nederlands Rundvee Syndicaat, Arnhem, 104 pp.

Norman, H.D.; P.D. Miller; B.T. McDaniel; F.N. Dickinson and C.R. Henderson (1974). USDA-DHIA Factors for standardizing 305-day lactation records for age and month of calving. Report of U.S. Dept. Agr. ARS-NE-40, 91 pp.

Poutous and Mocquot (1975). Etudes sur la production latière des bovines. III. Relations entre critères de production duree de lactation et intervalle entre le 1er et le 2ième velage. Ann. Genet. Sel. Anim. 17: 181-189.

Thompson, R. (1979). Sire evaluation. Biometrics 35: 339-353.

Wilmink, J.B.M. (1985). De correctie van melkproduktiegegevens voor leeftijd, seizoen en lactatiestadium (with English summary). Report B-243, IVO "Schoonoord", Zeist.

Witt, M.; D. Flock and U.E. Pfeiderer (1969). Untersuchungen über den Verlauf von Lactationskurven auf der Grundlage täglicher Milch- und Fettkontrollen. Z. Tierz. und Züchtungsbiol. 86: 1-29.

Chapter 2

ADJUSTMENT OF TEST-DAY MILK, FAT AND PROTEIN YIELD FOR AGE, SEASON AND STAGE OF LACTATION

J.B.M. Wilmink

Royal Dutch Cattle Syndicate
P.O. Box 454, 6800 AL Arnhem

Accepted for publication in Livestock Production Science

ABSTRACT

The objective of this study was to develop a method for the adjustment of test-day milk, fat and protein yield for age at calving, stage of lactation, and month of calving. Test-day records of 14275 purebred Dutch Friesians were analysed by generalized least squares, (GLS). Functions of days in lactation and age in months were fit to the age differences for different stages in lactation. For both milk, fat and protein yield a close fit to the GLS means was obtained. From these functions, adjustment factors (CF_{ageI}) were computed depending on the cow's age in months and number of days in lactation. A multiplicative adjustment for age and stage of lactation by CF_{ageI} was compared with an adjustment using factors derived from the constant estimates (CF_{ageII}). Comparison was on the F-variable and constant estimates for age after adjustment. When CF_{ageI} was used instead of CF_{ageII}, lower F-values were obtained and the constants showed a more regular pattern. Additive adjustment factors for month of calving were derived from the GLS constant estimates. After correction, month of calving was only significant for protein yield. The adjustment of test-day yields for age at calving and stage of lactation was improved by using mathematical functions.

INTRODUCTION

A cow's test-day yield is influenced by systematic environmental effects such as age and season of calving, stage of lactation, herd, and number of days open (Danell, 1982a) A standardization of test-day yields for all these effects, except herd, allows a comparison of cows within herd which is useful for management and selection. In the Netherlands, for example, test-day milk yields of cows, recorded prior to 250 days in milk, are standardized for age and season of calving, and stage of lactation. The standardized tests are averaged to give a herd index which is used as a management guide by the farmer. Test-day records of cows longer than 250 days in milk are not considered because the effect of days open

increases substantially in that part of the lactation (Auran, 1974; Danell 1982a).
Test-day records of heifers collected in early lactation and standardized for age and season, offer a proposition for an early genetic evaluation of bulls and cows for 305-day yield. The possible advantages of an early evaluation are described by Danell (1982b).

Adjustment factors for age and season of calving for test-day milk yield have been developed by Auran (1973), Dommerholt (1975), Schaeffer and Burnside (1976) and more recently by Danell (1982a), who also developed factors for test-day fat yield and percentage.
Age and season adjustment factors for test-day protein yield were not found in the literature.
Age and season factors for test-day yields are dependent on stage of lactation. Dommerholt (1975) showed that age at calving accounted for about 46 % of the variation in daily milk yield at the beginning of lactation, which decreased to about 5 % at the end. Season of calving accounted for about 1 % at the beginning which increased to about 8 % at the end of lactation. Traditionally age and season factors were presented for classes of age * stage of lactation. The accuracy of adjustment depends on the size of classes defined for age, season and stage of lactation. Because age at calving has a major effect on test-day yields, it might be worthwile to develop a new method using functions for the adjustment of age at calving. By these functions age factors can be computed for each age at calving (months) by day in lactation, resulting in a more precise adjustment.

The purpose of this paper was to develop functions for age factors and to present season factors, for standardizing test-day milk, fat and protein yield.

MATERIAL

Individual test-day records of 14275 pure bred Dutch Friesian cows calving from 1 June 1981 to 1 June 1982 in 721 randomly chosen herds were extracted from the milk recording files (Data set I). Sires of these cows were known and each cow had at least 7 paternal halfsibs in order to avoid disconnectedness between age and sires. Lactation length was between 280 and 340 days. Because individual tests were taken at variable intervals, new test-day yields were calculated by inter- or extrapolation of the two nearest tests to day 10,..., 290 post partum with 20-day intervals. From these new test-day yields, 305-day yield (Y_{305}) was computed by

$\sum_{i=1}^{15} 20 * x_i + 5 * x_{15}$ where x_i was the i-th test-day yield.

Table 1. Overall means and standard deviations (SD) for test-day (x_t) and 305-day yield (Y_{305}) in data set I and II

	Data set 1		Data set 2	
	mean	SD	mean	SD
kg milk				
x_{70}	24.2	5.2	24.7	5.2
x_{150}	19.7	4.2	20.0	4.1
x_{230}	15.3	3.6	15.4	3.5
Y_{305}	5865	1096	5941	1098
kg fat				
x_{70}	0.960	0.219	0.978	0.218
x_{150}	0.811	0.167	0.826	0.168
x_{230}	0.687	0.152	0.700	0.152
Y_{305}	248	47	252	47
kg protein				
x_{70}	0.763	0.164	0.770	0.159
x_{150}	0.669	0.139	0.671	0.136
x_{230}	0.558	0.126	0.558	0.123
Y_{305}	199	37	200	36
number of records				
	14,275		4,664	

The correlation of Y_{305} with usual 305-day milk yield, as computed in the milk recording files was 1.00 for kg milk, kg fat and kg protein and there was a difference of 2.0 kg in mean 305-day milk yield and no difference in standard deviation.

Data set I was used for the estimation of effects for age and season of calving. Another data set II was formed with 4664 purebred Dutch Friesian cows calving from 1 June 1982 to 1 June 1983 in 227 herds, chosen at random from the above 721 herds. Lactation length was between 280 and 320 days and, as in data set I, new test-day yields were computed by inter- and extrapolation. Data set II was used to check the validity of the adjustment procedure. In order to test the age * herd level interaction, herds were split into low ($<$ 5400 kg), medium (5400-6200 kg) and high ($>$ 6200 kg) level groups according to the average 305-day milk yield, corrected for age, season and breed (Dutch * Holstein Friesian), of all cows in the herd with a completed lactation. The correction factors used were from Wilmink (1985). Grouping was such that 25%, 50%, 25% of the herds were in low, medium and high levels respectively. Overall means and standard deviations for test-day yields at 70, 150 and 230 days in milk and the 305-day yields are shown in table 1.

METHODS

Estimation of age effects

Age and season effects were estimated for each test-day milk, fat or protein yield separately by a mixed model (model 1).

$$Y_{ijkmno} = \mu + h_i + a_j + m_k + l_m + s_n + e_{ijkmno} \qquad (1)$$

where Y = test-day milk, fat or protein yield

μ = overall mean

h_i = effect of herd i (i = 1,721)
a_j = " " age class j (j = 1,10)
m_k = " " season k (k = 1,6)
l_m = " " lactation length class m (m = 1,3)
s_n = " " sire n (n = 1,244); $E(s_n) = 0$; $var(s_n) = \sigma_s^2$
e_{ijkmno} = residual; $E(e) = 0$; $var(e) = \sigma_e^2$.

All effects were assumed to be fixed, except for sires and residuals. Covariances between sires and residuals and among residuals were assumed to be zero; sires were related with their sire and maternal grandsire.

Ten age classes were defined, corresponding to 22-24, 25-27, 28-32, 33-37, 38-44, 45-56, 57-68, 69-92, 93-104 and more than 104 months at calving. Seasons were defined for two month periods, starting with June-July and so on.

Three classes for lactation length were defined as 280-299; 300-319 and 320-340 days in milk respectively.

Estimated age effects were solutions of

$$\begin{pmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 & \mathbf{X}_1'\mathbf{Z} \\ & \mathbf{X}_2'\mathbf{X}_2 & \mathbf{X}_2'\mathbf{Z} \\ \text{symm.} & & \mathbf{Z}'\mathbf{Z} + k\mathbf{A}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{h}} \\ \hat{\mathbf{f}} \\ \hat{\mathbf{s}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{Y} \\ \mathbf{X}_2'\mathbf{Y} \\ \mathbf{Z}'\mathbf{Y} \end{pmatrix} \qquad (2)$$

Here the matrices $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{Z}$ are known incidence matrices, $\hat{\mathbf{h}}_i$ is a vector with solutions for $\mu + h_i$, $\hat{\mathbf{f}}$ a vector with solutions for age, season and lactation length and $\hat{\mathbf{s}}$ a vector with sire solutions. Matrix $\mathbf{A}$ represents the additive genetic relationship matrix between sires and their sires and maternal grandsires. $\mathbf{A}^{-1}$ was computed according to Henderson (1975). Scalar k represented σ_e^2/σ_s^2, which was set equal to 15 for each test-day yield.

In solving (2), the $\hat{\mathbf{h}}$ equations were absorbed into the $\hat{\mathbf{f}}$ and $\hat{\mathbf{s}}$ equations and the restrictions

$\sum_{j=1}^{10} \hat{a}_j = 0$, $\sum_{k=1}^{6} \hat{m}_k = 0$ and $\sum_{m=1}^{3} \hat{l}_m = 0$ were imposed.

Approximate effective numbers for the age solutions were taken from the diagonal of the coefficient matrix for the age * age part after absorbtion of $\hat{\mathbf{h}}$. The μ in (2) was estimated by substituting $\hat{\mathbf{f}}$ and $\hat{\mathbf{s}}$ in (2) by their estimates and imposing $\Sigma \hat{h}_i = 0$.

Fitting functions to the age effects

The generalized least square (GLS) means $\hat{\mu} + \hat{a}_1$ for milk, fat and protein yield, were fit by

$$\text{LEVEL}(x) = a + bx + ce^{-0,05x} + dx^2 \qquad (3)$$

where x= days in lactation and a, b, c and d are parameters.
The $\hat{\mu} + \hat{a}_1$ reflected the mean test-day yield adjusted for the other effects in (1) of heifers in age class 1. Equation (3) represents the level of test-day yields, based on cows in age class 1, necessary for computation of multiplicative adjustment factors. For older cows, the age differences for age class j ($\hat{a}c_j$) were all defined as $\hat{a}_j - \hat{a}_1$. The $\hat{a}c_j$ were fit by

$$\text{DIF}(x,t) = a_t + b_t\, x + c_t\, e^{-0.05x} \qquad (4)$$

In equation (3) and (4) factor a is associated with level of production, b with production decrease after peak yield, c with production increase towards peak yield and the factor -0.05 with the moment of peak yield, i.e. about 50 days post partum. Both equations were a modification of the lactation curve function in Cobby and Le Du (1978)

The parameters (a_t, b_t, c_t) in (4) are a function of age at calving (PARAM(t)).
The relationship between these parameters and age in months (t) was described by

$$\text{PARAM}(t) = p + qt \tag{5}$$

and

$$\text{PARAM}(t) = p + qt + r\ln(t), \tag{6}$$

where ln(t) is the natural logarithm of t.
Equation (5) was for parity 1 and 2 separately, whereas equation (6) was for higher parities. Age classes 1 up to 3, 4 and 5, 6 to 10 were associated with parity 1, 2 and 3 or greater respectively. All constants in equations (3) to (6) were estimated by ordinary least squares.

Adjustment factors

Multiplicative adjustment factors for age at calving (CF_{ageI}) in month t' and day x' in lactation were computed as follows:

$$CF_{ageI} = \frac{\text{LEVEL}(x') + a(t') + b(t')x' + c(t')\ e^{-0.05x'}}{\text{LEVEL}(x') + a(t_r) + b(t_r)x' + c(t_r)\ e^{-0.05x'}} \tag{7}$$

where t_r represented the base age, which was taken to be 96 months. Let a cow be x' days in lactation and t' months old at calving. Then the level of production (yield/day) for the class x' days in milk by t' months for age is estimated by LEVEL(x') + DIF(x',t'). Using (5), when a cow is in the second lactation, DIF(x',t') = $a(t') + b(t')x' + c(t')\ e^{-0.05x'}$, where $a(t') = p_a + q_a t'$, $b(t') = p_b + q_b t'$ and $c(t') = p_c + q_c t'$.
Likewise the level of production is calculated for the reference class x' by t_r, or x_r by t_r when test-day yields are adjusted for both age and stage of lactation. In fact LEVEL (x') + DIF (x',t') represent the lactation curve of all cows which aged t' months at calving. Age factors were multiplicative because age effects were found to be proportional to herd level (Wilmink, 1986). When correcting for stage of lactation, x' in the denominator of (7) was taken to be 50 days.
In deriving adjustment factors for month of calving, month effects were computed from the season effects in model 1 by interpolation and smoothing, using routine ICSSVC (IMSL, 1980). Adjustment was in

an additive manner, because the effect of season of calving was independent of production level, as reported by Wilmink (1985)

Validity of adjustment

Test-day milk, fat and protein records in data set II were multiplicatively adjusted for age by CF_{ageI} and additively by the smoothed constants for month of calving. After correction, test-day yields in data set II were analyzed by a fixed model (model 2):

$$Y_{ijkl} = \mu + hl_i + a_j + m_k + (hl*a)_{ij} + e_{ijkl},$$

where hl_i is the effect of herd level i. Analysis of model 2 was by LSML76 (Harvey, 1977). There was chosen for model 2 instead of model 1 because of saving computer expenses. Besides, an adjustment with multiplicative age factors per age class 1 to 10, derived from the constant estimates in model 1 (CF_{ageII}) was examined. CF_{ageII} was calculated as $(\hat{\mu} + \hat{a}_j)/(\hat{\mu} + \hat{a}_9)$. Both methods of age adjustment were compared. The advantage of CF_{ageI} should be that the adjustment factors are different per day in lactation by age in months, while they are different per lactation stage by age classes for CF_{ageII}.

In order to get a better insight into the validity of age adjustment, in model 2, 17 instead of 10 age classes were defined as follows (in months):

22-23, 24-25, 26-27, 28-29, 30-31, 32-34, 35-36, 37-38, 39-40, 41-42, 43-44, 45-50, 51-56, 57-68, 69-92, 93-104 and 105 and higher.

RESULTS

Age effects

In table 2, the results for fitting LEVEL(x) (equation (3)) to the GLS test-day means in age class 1 are presented. Plots of LEVEL(x) and the GLS means are shown in figure 1. In all cases, the coefficient of determination was close to 1, and the test for correlation

Table 2. Estimates of the parameters, in the function[a] fit to the GLS means of heifers in age class 1

	milk/day	fat/day ($kg*10^{-3}$)	protein/day ($kg*10^{-3}$)
a	20.7	771	595
b	-0.041	-1.000	-0.376
c	-4.00	-21.23	15.05
d	$4.42*10^{-5}$	$1.23*10^{-3}$	$-0.45*10^{-3}$
R^2(%)	99.9	99.9	99.9
DW[b]	2.19	2.14	2.34

[a] $LEVEL(x) = a + bx + ce^{-0.05x} + dx^2$
where x = days in milk

[b] DW = Test for serial correlation of the residuals.
(significant ($P < 0.05$) if $DW < 0.82$ or $DW > 3.18$)

among the residuals, (Durbin and Watson, 1951), was not significant or inconclusive ($P < 0.05$). When heifers commenced their lactation, level of production was 20.7 kg milk, 0.771 kg fat and 0.595 kg protein per day. Milk/day then increased until about 50 days post partum followed by a decrease (figure 1). Fat/day and protein/day decreased over the entire lactation.

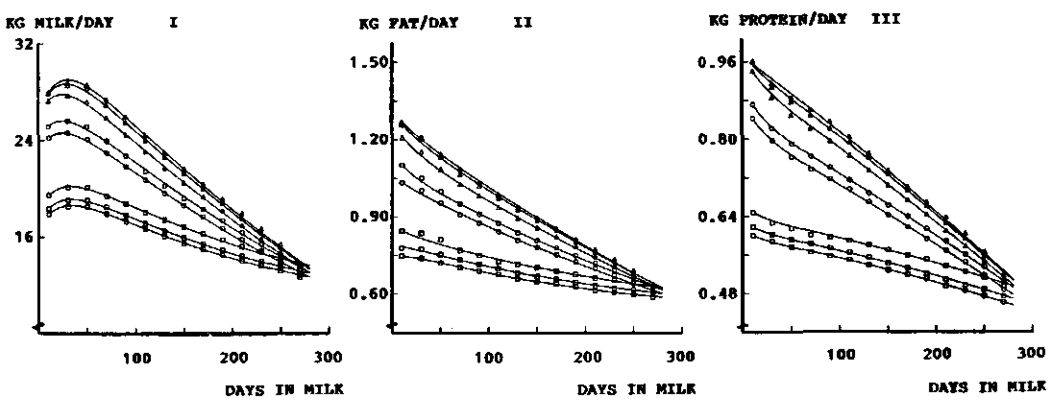

Figure 1. Plots of the equations fitted to the GLS means (□ for first, o for second and Δ for higher parities) for daily milk(I), fat(II) and protein yield (III). At 50 days post partum, the lines from below to upper refer to age class 1 to 8 respectively.

Equation (4) was fit to the age differences expressed as a deviation from age class 1. Estimates for the parameters a, b and c for milk, fat and protein per day are presented in table 3. The pattern of the lactation curves was strongly affected by parity. The b-values, for age differences after 50 days in milk, were rather similar within first, second and later parities with a maximum difference of $0.53 * 10^{-2}$, 0.25 and 0.13 for milk, fat and protein yield per day, respectively. Between parities differences in b-values were $0.93 * 10^{-2}$, 0.40 and 0.32 at the minimum and $0.41 * 10^{-1}$, 1.67 and 1.23 at the maximum for the three yield traits.
The a-values for milk/day increased from 0.0 to 1.7 kg in the first parity, from 6.74 to 7.82 kg in the second and from 10.45 to 12.32 kg in higher parities. For fat/day these increases were 0.078, 0.316 and 0.496 and for protein/day 0.040, 0.268 and 0.410 kg respectively. The R^2 for all second and higher parity cows was higher than 99.5%. In the first parity R^2 ranged between 32.2 and 95.5%. These values were lower than in later parities, due to the

Table 3. Estimates of the parameters in the function[a)] fitting the lactation curves at different age classes relative to the lactation curve at age class 1

Age			milk (kg)				fat ($kg*10^{-3}$)				protein ($kg*10^{-3}$)			
class[b]	n_e[c]	age	a_t	b_t	c_t	R^2	a_t	b_t	c_t	R^2	a_t	b_t	c_t	R^2
1	1110	23.5	0	0	0	0	0	0	0	0	0	0	0	0
2	1108	25.7	0.58	-0.0009	-0.17	87.2	32.3	-0.063	3.7	86.8	5.4	0.033	17.5	32.2
3	325	29.3	1.70	-0.0027	-0.12	95.5	78.1	-0.139	45.3	85.3	40.4	0.008	39.5	53.7
4	1203	35.8	6.74	-0.0241	-0.24	99.9	271.7	-0.919	42.6	99.7	244.9	-0.799	78.6	99.8
5	817	39.7	7.82	-0.0266	-0.40	99.8	315.9	-1.023	84.1	99.8	267.6	-0.793	63.8	99.8
6	1429	49.4	10.45	-0.0359	-1.10	99.9	423.7	-1.416	91.1	99.8	350.5	-1.110	47.9	99.8
7	1090	61.6	11.82	-0.0398	-2.33	99.9	476.1	-1.578	106.7	99.8	392.0	-1.204	15.6	99.7
8	1206	77.8	12.32	-0.0408	-3.10	99.9	496.4	-1.645	55.3	99.8	410.4	-1.235	- 9.3	99.7
9	289	97.6	12.21	-0.0412	-3.82	99.9	484.9	-1.667	30.1	99.6	389.5	-1.218	-12.8	99.5
10	249	117.5	11.27	-0.0383	-3.20	99.7	435.2	-1.537	3.9	99.7	377.3	-1.171	-39.6	99.6

a) $DIF(x,t) = a_t + b_t x + c_t e^{-0.05x}$
where x = number of days in lactation

b) age classes 1 to 3 refer to parity 1, 4 and 5 to parity 2 and more than 5 to parity 3 and higher.

c) effective number

fact that b was almost nonsignificant.

The parameters in table 3 were fit by equations (5) and (6) for parity 1, 2, 3 and later giving the relationship of a, b and c in (4) with age at calving in months (table 4). Thus when parity and age at calving are known, equation (4) can be applied for that age. In figure 1 the total sum of LEVEL(x) + DIF(x,t') is plotted, given the weighted average age in months (t') in age class 1 up to 8 in data set I. Notice that DIF(x,t')=0 for age class 1. The GLS means (k = 1,8) estimated from model 1 were also plotted. Figure 1 shows that LEVEL(x) + DIF(x,t') where the parameters in DIF(x,t') were estimated by equation (5) and (6), fit the age differences well for all days in lactation. Compared with 24 month old heifers, older heifers produced more over the entire lactation. Second and later parity cows had higher test-day yields than heifers up to about 250 days post partum. The differences were greatest at the beginning of lactation and were almost zero or even negative after about 250 days post partum.

Table 4. Parameters of the function a), fitting the a_t, b_t, c_t values of the functions in table 3 by lactation number

parity		milk (kg) a_t	b_t	c_t	fat ($kg*10^{-3}$) a_t	b_t	c_t	protein ($kg*10^{-3}$) a_t	b_t	c_t
1	p	-7.05	0.01134	0.350	- 318	0.56130	-200	-161	0.01727	-43.0
	q	0.299	-0.00048	-0.017	13.6	-0.02403	8.27	6.87	-0.00098	1.77
2	p	-3.05	0.00155	1.24	- 131	0.02805	-355	-34.3	-0.26910	26.6
	q	0.274	-0.00063	-0.041	11.2	-0.02646	10.6	7.13	-0.01242	0.509
≥ 3	p	-45.8	0.11720	45.3	-2033	5.4850	-98.7	-118	2.4960	1004
	q	-0.205	0.00055	0.141	-9.34	0.02456	-2.52	-5.60	0.01234	2.72
	r	17.0	-0.04630	-13.7	748	-2.0800	82.8	459	-1.0590	-283

a) PARAM(t) = p + q t   if number of lactation is 1 or 2
PARAM(t) = p + qt + r ln t   if number of lactation is 3 or higher
where PARAM(t) = a_t, b_t or c_t
t = age at calving in months
ln = natuaral logarithm

Table 5. Generalized least squares constant estimates of daily milk, fat and protein yield for season of calving (model 1)

	Days in lactation														
	10	30	50	70	90	110	130	150	170	190	210	230	250	270	290
Milk/day (kg * 10^{-1})															
J,J	4	- 3	- 5	- 5	- 8	-13	-19	-21	-20	-17	-14	-12	- 9	- 5	1
A,S	- 3	- 7	-11	-15	-16	-15	-12	- 9	- 7	- 5	- 2	3	9	14	17
O,N	- 6	- 5[a)]	- 5	- 3	- 1	0	1	4	8	13	17	19	18	13	8
D,J	- 2	0	2	2	3	7	11	14	16	16	14	10	5	0	- 6
F,M	0	4	7	10	12	14	15	13	9	4	- 1	- 7	-12	-14	-13
A,M	7	11	12[a)]	11	9	6	3	- 1	- 6	-11	-14	-13	-12	- 9	- 6
Fat/day (kg * 10^{-3})															
J,J	-53	-53	-46	-45	-44	-46	-53	-55	-46	-32	-23	-21	-15	- 6	10
A,S	-19	-28	-33	-32	-21	-13	- 5	0	2	3	8	22	33	39	42
O,N	37	31	28	33	33	34	34	41	45	44	42	44	38	29	18
D,J	19	26	28	27	29	36	34	28	22	23	24	18	13	3	-16
F,M	2	18	28	28	18	12	11	8	6	2	- 3	-20	-38	-45	-44
A,M	13	6	- 5	-11	-16	-22	-21	-22	-30	-40	-48	-43	-31	-20	-11
Protein/day (kg * 10^{-3})															
J,J	- 7	- 1	- 3	- 4	- 8	-20	-41	-55	-57	-51	-44	-41	-34	-25	- 7
A,S	16	6	- 6	-23	-33	-33	-27	-26	-25	-22	-17	0	22	39	47
O,N	- 7	-13	-12	- 9	- 5	- 2	- 1	6	21	34	45	51	45	36	26
D,J	- 8	-12	-10	- 7	- 3	12	26	37	36	37	35	30	24	14	- 4
F,M	-15	-10	4	19	31	34	34	30	25	21	15	- 2	-22	-34	-40
A,M	21	30	28	23	18	10	9	9	- 1	-19	-34	-38	-35	-29	-23

a) Entries, underlined as --, are the effects of months October-November on daily yields regardless of date of calving.
Entries, underlined as __, are the effects of months May-June.

Season effects

In table 5, the GLS constant estimates for the season effects in model 1 are presented. With respect to daily milk yield, spring calving cows produced relatively most in early lactation, whereas autumn calving cows produced relatively most in late lactation. When looking at the effect of calendar month, regardless of month of calving, daily milk yield increased by about 1-2 kg in May and June (underlined with __) and decreased by 1-2 kg in October and November (underlined with --). For example, the effect of October and November on daily milk yield of cows, calved in June and July, is about -2 kg. These cows are then 120-180 days in milk.
A similar pattern was observed for daily protein yield. For daily fat yield, cows calving in April to July produced less than average over almost the entire lactation, whereas cows calving in October

to January produced more. With respect to the effect of months, there was a tendency for yield to increase in April, May and to decrease in September and October, regardless of date of calving.

Validity of adjustment

Test-day records in data set II were adjusted multiplicatively for age with CF_{ageI} (factors derived from equation (7)) or CF_{ageII} (factors derived from the constant estimates in (1)). Adjustment for month of calving was additive and for stage of lactation multiplicative towards a base of 50 days post partum. The adjusted records were analyzed by model 2, using LSML76 (Harvey, 1977). Results are in table 6. Age, month and stage of lactation were found to be significant. The least square constants for age ranged from -0.8 to 0.8 kg milk when adjustment was by CF_{ageI}, and from -1.0 to 1.0 kg milk when adjustment was by CF_{ageII}. The corresponding F-values were 2.3 and 2.8 respectively. For daily fat yield only age was significant. Least squares constants varied from -0.041 to 0.041 and from -0.052 to 0.065 kg fat if age adjustment was by CF_{ageI} and CF_{ageII} respectively. The F-values were 3.0 and 3.8 respectively. The constants for age at calving for protein yield/day ranged between -0.027 and 0.030 and between -0.031 and 0.029 kg when adjustment was by CF_{ageI} and CF_{ageII} respectively. Corresponding F-values were 2.7 and 3.1. Month effects for daily protein yield were found to be significant. The interaction herdlevel * age was non significant in all cases.

DISCUSSION AND CONCLUSION

In this study, equations were fit to the GLS means for test-day yields in different age classes. These equations consisted of a function of number of days in milk, describing the level of test-day yields in age class 1, and a function of number of days in milk and age in months fitting the age differences relative to age class 1. Using these equations, age factors for test-day milk, fat and protein yield can be computed (CF_{ageI}) for any combination of age at calving and day in milk. An adjustment for stage of lacta-

Table 6. Least squares constants for age and F-values for age and month of calving, and stage of lactation after adjustment (model 2)

age class	Milk/day (kg * 10^{-1}) I[a]	Milk/day (kg * 10^{-1}) II[b]	fat/day (kg * 10^{-3}) I	fat/day (kg * 10^{-3}) II	protein/day (kg * 10^{-3}) I	protein/day (kg * 10^{-3}) II
1	7	4	41	24	30	20
2	-3	- 4	- 3	- 6	- 9	-11
3	3	5	11	22	12	18
4	8	4	10	- 9	19	8
5	1	5	14	30	20	29
6	3	0	18	3	15	11
7	-4	- 5	- 6	-10	-10	-13
8	-1	- 2	- 5	2	1	- 1
9	2	0	18	13	12	8
10	-6	- 2	-31	-13	-27	-16
11	1	10	23	65	- 4	21
12	1	- 2	- 4	-14	- 5	-11
13	6	9	18	29	9	16
14	-3	- 4	- 9	-13	- 9	-12
15	-7	- 8	-41	-45	-25	-28
16	-1	- 1	-23	-25	- 6	- 8
17	-8	-10	-41	-52	-24	-31
F-values						
Age	2.3[c]	2.8[c]	3.0[c]	3.8[c]	2.7[c]	3.1[c]
Month	2.2	2.2	1.0	1.0	2.9[c]	2.9[c]
Stage of lactation	2.3[c]	2.4[c]	1.1	1.1	2.0	2.0

a) Adjustment for age with CF_{ageI}
b) " " " " CF_{ageII}
c) $P < 0.01$

tion can be carried out, if required, by simply replacing x' in the denominator in (7) by a reference day in lactation. As an example, using the values in table 2 and 4, the factor for standardizing a daily milk yield to mature age (96 months) and 50 days in milk, of a cow which is 150 days in milk and 36 months old at calving, equals (15.54 + (6.814-3.1695-0.00013)) / (18.43 + (12.110-2.0665-0.30330)) = 0.681

In table 6 age at calving was significant for all traits after adjustment. Similar results were found by Dommerholt (1975) and Danell (1982a). Both authors reported that small age and season effects remained after adjustment. A reason for this may be that the aim of age adjustments is to remove age effects. Because sires were not included in model 2 genetic differences between different age classes should still be present. A small negative trend in age differences was observed with increasing age. A confounding of age with season of calving might be another reason. In The Netherlands heifers younger than 27 months generally calve from January to April, whereas heifers older than 29 months generally calve from June to October. Nevertheless, a more regular pattern in age contrasts was found when adjustment was by CF_{ageI} than by CF_{ageII}. The F values decreased when using CF_{ageI}. Adjustment of test-day yields for age at calving was improved by using mathematical functions.

Month of calving was not significant for daily milk and fat yield after adjustment. The reason that month of calving was significant for protein yield might be due to herd-year-season or year-season interactions. An attempt was made to improve the adjustment for month of calving by using functions. Functions of calving date and days in milk with sine and cosine terms were fit to the GLS estimates for season in model 1. In most cases a poor fit was obtained. Functions for season adjustment were not worthwhile.

Stage of lactation was not significant after adjustment for fat and protein yield but was for milk yield. Daily milk yield at 70 days post partum was 0.5 kg higher in data set II than in data set I. An interaction with year of calving may play a role here. Differences of 0.5 kg milk for lactation stage after correction were also reported by Dommerholt (1975).

Adjustment of test-day records for environmental effects is useful for early selection of bulls and cows or for farm-management purposes. In the Netherlands, first sire selection, based on progeny performance for milk, is carried out on an average of the first test-day records of the daughters. A best linear unbiased prediction of sire contributions on these records, preadjusted for age and month of calving and stage of lactation in a model with herd-year-season effects will substantially improve the unbiasedness of selection. For farm management a herd index based on standardized test-day milk yield records at the last test is used. In this index fat and protein yield might also be taken into account, giving a better reflection of, for example, the amount and quality of feed given to the cows. By mutually comparing the herd indices of successive tests, a farmer can observe possible changes of environmental conditions within his herd.

REFERENCES

Auran, T. (1973). Studies on monthly and cummulative monthly milk yield records. 1. The effect of age, month of calving, herd and length of the first period. Acta Agric. Scand. 23 : 189-199.

Auran, T. (1974). Studies on monthly and cummulative monthly milk yield records. II. The effect of calving interval and stage in pregnancy. Acta Agric. Scand. 24 : 339-348.

Cobby, J.M. and Y.L.P. Le Du (1978). On fitting curves to lactation data. Anim. Prod. 26 : 127-133.

Danell, B. (1982a). Studies on lactation yield and individual test-day yields of Swedish dairy cows. 1. Environmental influence and development of adjustment factors. Acta Agric. Scand. 32 : 65-81.

Danell, B. (1982b). Studies on lactation yield and individual test-day yields of Swedish dairy cows. IV. Extension of part-lactation records for use in sire evaluation. Acta Agric. Scand. 32 : 103-114.

Dommerholt, J. (1975). Correctie van de melkgift van koeien voor verschillen in leeftijd, seizoen en lactatiestadium (with English summary). Agricultural Research reports 844, Centre for Agricultural Publishing and Documentation, Wageningen.

Durbin, J. and G.S. Watson, (1951). Testing for serial correlation in least squares regression. Biometrica 38 : 159-177

Harvey, W.R. (1977). User's guide for LSML76, a mixed model least-squares and maximum likelihood computer program. Ohio State University.

Henderson, C.R. (1975). Inverse of a matrix of relationships due to sires and maternal grandsires. J. Dairy Sci. 58 : 1917-1920.

IMSL, (1982). IMSL library, reference manual, 9th edition. Int. mathematical and statistical libraries, Houston.

Schaeffer, L.R. and E.B. Burnside (1976). Estimating the shape of the lactation curve. Can. J. Anim. Sci. 56 : 157-170.

Wilmink, J.B.M. (1985). De correctie van melkproduktiegegevens voor leeftijd, seizoen en lactatiestadium (with english summary). Report-B-243, IVO-"Schoonoord", Zeist.
Wilmink, J.B.M. (1986). Adjustment of lactation yield for age at calving in relation to level of production. Submitted to Livestock Prod. Sci.

Chapter 3

COMPARISON OF DIFFERENT METHODS OF PREDICTING 305-DAY MILK YIELD USING MEANS CALCULATED FROM WITHIN HERD LACTATION CURVES

J.B.M. Wilmink

Royal Dutch Cattle Syndicate
P.O. Box 454, 6800 AL Arnhem

Accepted for publication in Livestock Production Science

ABSTRACT

The prediction of the unknown part of a 305-day lactation is usually based on regression models. For these models the relation between known test-day yields and remaining lactation yield as well as their means within classes of environmental effects have to be known. Due to environmental changes, means may vary severely. In this study a procedure for estimation of the means is presented and single regression (SR), multiple regression (MR), and factor analysis (FA) models for prediction are compared. Both MR and FA use all known test-day yields as information sources. Method SR uses the last known test-day yield. Lactation curves were estimated per herd on all records, adjusted for age and season of calving. By back-adjustment for age and season of calving, lactation curves were obtained for a group of cows pertaining to a certain class of herd * age * month of calving. Means necessary for prediction were calculated from these lactation curves (type II). Correlations with means, which were directly calculated in the data set (type I), ranged from 0.61 to 0.78 for test-days and from 0.51 to 0.69 for remaining lactation milk yield. The average difference between type I and II means was almost zero.

Using type II means, no differences were found between SR, MR and FA. The correlation between predicted and realized 305-day milk yield was 0.87, 0.88 and 0.88 respectively for parity 1, 2 and greater than 2, when the last test-day yield was known at 50 days post partum. The use of type II means resulted in a more precise adjustment.

INTRODUCTION

The prediction of a cow's 305-day milk yield is important for early breeding value prediction and for farm management purposes. Many methods have been applied to the prediction of a 305-day milk yield, e.g. a single regression of the remaining part lactation yield on the last known test-day yield with and without an intercept, a multiple regression of the unknown part on known

test-day yields, and the use of functions describing the lactation curve (Auran 1976, Schaeffer et al. 1977, Congleton and Everett, 1980). In general the last known test-day yield provided the most information about yield in the remaining lactation.
A single regression method using only the last test-day yield was recommended by Dommerholt (1975), Wiggans and Van Vleck (1979) and Danell (1982). In such a single regression method an intercept may be included, which, according to Dommerholt (1975), results in a more accurate prediction. With the intercept, the remaining part lactation yield is estimated by multiplying the deviation of the cow's last test-day yield from the mean test-day yield with a prediction factor and adding it to the mean remaining lactation yield. Such means have to be known within classes of environmental effects such as herd, age and year-season of calving. Prediction factors are based on a ratio of covariances between remaining part lactation and last test-day yields, and variances of the last test-day yields.

Means for last-test day yields are very sensitive to environmental effects, and means for remaining lactation yields are not known when predicting a 305-day milk yield. Ratios of covariances and variances are much less sensitive. While level of 305-day milk yield increased by 1000 kg, the correlations among test-day milk records hardly changed from 1975 to 1985 (Dommerholt, 1975; Wilmink, 1985). A procedure for the estimation of means within classes of environmental effects would be worthwhile, if a regression model including an intercept is to be used for prediction of 305-day milk yields.

In this case a multiple regression model would theoretically result in the most accurate linear prediction, because all available test-day yields are used. If factor analysis is used, the correlation matrix between test-day yields is approximated, the accuracy of which depends on the number of factors used (Harman, 1970). For practical purposes this may be of interest because of a possible further decrease of the sensitivity of the approximated correlation matrix and prediction factors for environmental effects.

The objectives of this paper are a) to present a procedure for estimation of means and b) to compare the multiple regression, factor analysis and single regression methods, using the estimation procedure for means, with regard to unbiasedness and accuracy.

MATERIAL

Test-day records (kg milk/day), taken at variable intervals, of 30185 complete lactations from purebred Dutch Friesian cows in about 700 randomly chosen herds were extracted from the milk recording files. The data were split according to date of calving, into set I (1 June 1981 - 31 May 1982) and set II (1 June 1982 - 31 May 1983). Data set I was used for computation of correlations between test-day records, prediction factors and means, while set II was

Table 1. Overall means and standard deviations per data set for test-day milk yields (kg/day) x_t (x_i) at t days post partum

x_i, i=	x_t, t=	Data set I mean	Data set I SD	Data set II mean	Data set II SD
1	10	24.6	6.0	25.2	6.1
2	30	25.4	5.6	25.9	5.6
3	50	25.2	5.5	25.6	5.5
4	70	24.2	5.3	24.5	5.2
5	90	23.0	5.0	23.2	4.8
6	110	21.9	4.7	22.0	4.5
7	130	20.8	4.4	20.8	4.3
8	150	19.7	4.2	19.8	4.1
9	170	18.7	4.0	18.7	3.9
10	190	17.6	3.8	17.6	3.8
11	210	16.5	3.7	16.5	3.6
12	230	15.4	3.6	15.3	3.6
13	250	14.2	3.5	14.0	3.5
14	270	12.8	3.4	12.5	3.4
15	290	11.4	3.5	10.9	3.5
305-day yield		5885	1107	5903	1096
number of records		17607		12578	

used for comparison of prediction methods. Lactation length was from 280 to 339 days (data set I) and from 280 to 319 days (data set II). Incomplete records and records of heifers younger than 22 months at calving were removed. New test-day yields at day 10, 30, 50...290 post partum (p.p.) were calculated by inter- or extrapolation, providing estimates of 15 test-day yields at equal intervals within the lactation. Overall means and standard deviations are presented in table 1.

In order to test the procedure for estimaton of means, herds were split into low, medium and high level groups as described by Wilmink (1986). For computation of the correlations and prediction factors, classes were defined for age and season at calving. Age at calving was divided into 10 classes for months 22-24, 25-27, 28-32, 33-37, 38-44, 45-56, 57-68, 69-92, 93-104 and higher. Age classes 1 to 3 were defined as lactation number 1, 4 and 5 as lactation number 2, and from 5 onwards as later parities. Seasons were defined as two month periods, starting with June and July.

METHODS

Prediction methods

In order to compare the different methods, the mathematical relationship between multiple regression (MR), single regression (SR) and factor analysis (FA) to be used for prediction of a 305-day milk yield is shown. A predicted 305-day yield ($\hat{Y}_{305}$) can be written as the sum of production in the known part lactation (Y_D) and predicted production in the remainding part ($\hat{Y}_R$). When p test-day yields are known, $\hat{Y}_R$ can be estimated according to

$$\hat{Y}_R = \sum_{j=p+1}^{14} \hat{x}_j * 20 + 25 * \hat{x}_{15} \qquad (1)$$

which is the sum of 15-p predicted test-day yields ($\hat{x}_j$) multiplied by the length of the corresponding intervals. Note that $\hat{x}_{15}$, i.e predicted test-day yield at day 290, is multiplied by an interval of 25 days. The x_j can be predicted by a multiple regression of x_j on the known test-day yields x_i (i=1,p):

$$\hat{x}_j = \bar{x}_j + \sum_{i=1}^{p} b_{ij} \frac{(x_i - \bar{x}_i)}{\sigma_i} * \sigma_j \qquad (2)$$

where the b_{ij} are partial regression coefficients, and $\bar{x}_i$, σ_i are means and standard deviations of the i-th test-day yield respectively. For the situation where the last q test-day yields are regressed on the first p test-day yields (p+q = 15), partial regression coefficients are estimated by:

$$\mathbf{Z}_{nq} = \mathbf{Z}_{np}\, \mathbf{B}_{pq} + \mathbf{E}_{nq} \qquad (3)$$

where $\mathbf{Z}_{nq}$ = a matrix with standardized observations and dimension n*q

$\mathbf{B}_{pq}$ = matrix with b_{ij} (i = 1,p;j = p+1,15)

$\mathbf{E}_{nq}$ = matrix with residual elements

n = number of observations

$\mathbf{B}_{pq}$ is estimated according to

$$\mathbf{B}_{pq} = (\mathbf{Z}'_{pn}\, \mathbf{Z}_{np})^{-1}\, \mathbf{Z}'_{pn}\, \mathbf{Z}_{nq} = \mathbf{S}_{pp}^{-1}\, \mathbf{S}_{pq} \qquad (4)$$

where $$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{pp} & \mathbf{S}_{pq} \\ \mathbf{S}_{qp} & \mathbf{S}_{qq} \end{pmatrix}$$

Here **S** is the correlation matrix between the test-day yields x_i.

For a single regression of x_j on the last known test-day yield, x_p, equation (2) changes to

$$\hat{x}_j = \bar{x}_j + \sigma_j * b_{pj} \frac{(x_p - \bar{x}_p)}{\sigma_p} \qquad (5)$$

Estimation of b_{pj} is similar to (4).

From (1) and (5) it follows that:

$$\hat{Y}_R = \sum_{j=p+1}^{14} \bar{x}_j*20 + 25*\bar{x}_{15} + \left(20* \sum_{j=p+1}^{14} \frac{b_{pj}\sigma_j}{\sigma_p} + 25*\frac{b_{p15}\sigma_{15}}{\sigma_p}\right)(x_p-\bar{x}_p)$$

$$= \bar{Y}_R + b^*(x_p-\bar{x}_p) \quad (6)$$

where b^* = regression coefficient of Y_R on x_p.

In a factor analysis model, each standardized test-day yield can be represented by a sum of several common factors, which are mutually independent, and a unique factor. The common factors account for the correlations among the standardized test-day yields, while each unique factor accounts for the remaining variance of that standardized test-day yield. The weights of the common factors are called factor loadings.

When using a factor analysis model, equation (3) changes to

$$\mathbf{Z}_{nm} = \mathbf{A}_{nk}\, \mathbf{B}^*_{km} + \mathbf{E}_{nm} \quad (7)$$

where $\mathbf{Z}_{nm}$ = matrix with standardized observations and dimension n*m
$\mathbf{A}_{nk}$ = matrix with factor scores and dimension n*k
$\mathbf{B}^*_{km}$ = matrix with factor loadings and dimension k*m
k = number of factors
m = total number of tests, which is 15.

$\mathbf{B}^*_{km}$ is estimated in $\mathbf{S} = \mathbf{B}^{*'}_{mk}\, \mathbf{B}^*_{km} + \mathbf{K}^2$ according to the principal factor method (Harman, 1970). For the i-th standardized test-day yield, the diagonal of matrix $\mathbf{K}^2$ represent the unique variances e_i and the diagonal of $\mathbf{B}^{*'}_{mk}\, \mathbf{B}^*_{km}$ represent the communalities (λ_i) i.e. the variance which is shared in common with other standardized test-day yields. Total variance (= 1) equals $\lambda_i + e_i$. The number of factors ,k, was determined by the number of eigenvalues of $\mathbf{S}$ greater than 1.

A can be calculated if p test-day yields are known. In that case (Harman, 1970):

$$\mathbf{A}_{nk} = \mathbf{Z}_{np}\ \mathbf{S}^{-1}_{pp}\ \mathbf{B}^{*'}_{pk} \qquad (8)$$

The remaining q test-day yields can be predicted, using (7):

$$\mathbf{Z}_{nq} = \mathbf{A}_{nk}\ \mathbf{B}^{*}_{kq}.$$

Combination with (8) gives:

$$\mathbf{Z}_{nq} = \mathbf{Z}_{np}\ \mathbf{S}^{-1}_{pp}\ \mathbf{B}^{*'}_{pk}\ \mathbf{B}^{*}_{kq} \qquad (9)$$

where $\mathbf{B}^{*'}_{pk}\mathbf{B}^{*}_{kq}$ is an approximation of $\mathbf{S}_{pq}$. In (9) a test-day yield $\hat{x}_j$ is predicted according to

$$\hat{x}_j = \bar{x}_j + \sum_{i=1}^{p} b'_{ij}\ \frac{(x_i-\bar{x}_i)}{\sigma_i} * \sigma_j \qquad (10)$$

where b'_{ij} is an element in $\mathbf{S}^{-1}_{pp}\ \mathbf{B}^{*'}_{pk}\ \mathbf{B}^{*}_{kq}$. Equation (10) is of similar form to equation (2)

Criteria for comparison.

The correlation matrix **S** was calculated in data set I within classes of age * season of calving. Correlations were then pooled within lactation number 1, 2 and greater than 2. Data set II was used for comparison of MR, SR and FA.

The criteria for comparison were the mean difference between Y_{305} and $\hat{Y}_{305}$, the standard deviation of ($Y_{305}-\hat{Y}_{305}$) and the correlation between Y_{305} and $\hat{Y}_{305}$ and between Y_R and $\hat{Y}_R$. Furthermore, the methods of prediction were compared with theoretical values for the criteria, in which means do not play a role.

In order to calculate the theoretical values of the criteria, let

$$Y_{305} = \mathbf{a'x} \; ; \; Y_D = \mathbf{p'x} \; ; \; Y_R = \mathbf{q'x}; \; \hat{Y}_R = \mathbf{q'B'x};$$

where $\mathbf{x}$ = a vector with test-day yields

$\mathbf{a}$ = a vector with ones and 1.25 in element 15

$\mathbf{p}$ = a vector with ones and zeros

$\mathbf{q}$ = a vector with ones and zeros, but 1.25 in element 15.

$$\mathbf{B'} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{B'}_{qp} & \mathbf{0} \end{pmatrix}$$

Furthermore, let $\mathbf{C}$ be the covariance matrix of $\mathbf{x}$, then $var(Y_{305}) = \mathbf{a'Ca}$, $var(Y_D) = \mathbf{p'Cp}$, $var(Y_R) = \mathbf{q'Cq}$, $var(\hat{Y}_R) = \mathbf{q'B'CBq}$, $cov(\hat{Y}_R{}'Y_D) = \mathbf{q'B'Cp}$ and $cov(Y_{305}, \hat{Y}_{305}) = \mathbf{pCa} + \mathbf{q'B'Ca}$.

Consequently $r(Y_{305}, \hat{Y}_{305}) = \dfrac{\mathbf{p'Ca} + \mathbf{q'B'Ca}}{(\mathbf{a'Ca})^{\frac{1}{2}} \; (\mathbf{p'Cp} + \mathbf{q'B'CBq} + 2\mathbf{q'B'Cp})^{\frac{1}{2}}}$,

$r(Y_R, \hat{Y}_R) = \dfrac{\mathbf{q'B'Cq}}{(\mathbf{qCq})^{\frac{1}{2}} \; (\mathbf{q'B'CBq})^{\frac{1}{2}}}$ and

$$var\,(Y_{305} - \hat{Y}_{305}) = \mathbf{a'Ca} + \mathbf{p'Cp} + \mathbf{qB'CBq} + 2\mathbf{q'B'Cp} - 2\mathbf{p'Ca} - 2\mathbf{q'B'Ca}.$$

Estimation of means

The means in equations (2), (6) and (10) are influenced by environmental factors such as herd, year, age and season of calving. Furthermore, means for future test-day yields or remaining lactation yields are not available in practice. For the estimation of means, test-day records in data set I were adjusted for age and month of calving as described in a previous study (Wilmink, 1986). The base was 24 months for age and Februari for month of calving. Lactation curves were fit per herd by the function (Wilmink, 1986):

$$y_t = a + bt + ce^{-0.05t} \qquad (11)$$

where y_t represents an average test-day production per herd, t the number of days, e is the base of the natural log and a,b,c are parameters, describing the lactation curve. The parameters a and c represents level of production (at t = o), c the rate of increase of production to the peak and b the rate of decline thereafter. The factor -0.05 determines the time of peak production, i.e. about 50 days p.p. in this case.

Herd lactation curves in data set II were computed by the regression of parameters in (11) on the mean test-day yield per herd at 50 days p.p. ($\bar{x}_{50}$) and mean 305-day yield per herd ($\bar{Y}_{305}$) in data set I, both adjusted for age to 24 months and for season at calving. In (11), the decrease in daily production after peak yield is linear. Therefore mean x_{50} and Y_{305} per herd were the major factors in determining a lactation curve. For example, in data set II the parameter a in the herd lactation curve was calculated as $u + v\bar{x}_{50} + w\bar{Y}_{305}$, where u, v and w are regression coefficients. Likewise parameters b and c were calculated. Notice that data set I and II comprise records of cows calved in successive years. From $\bar{x}_{50}$ and $\bar{Y}_{305}$, obtained in the previous year, herd lactation curves for heifers can be calculated in the current year. Lactation curves for an age by month of calving subclass other than the base within the herd were obtained by a back adjustment first for age (multiplicatively) and then for month of calving (additively).
From these lactation curves the desired means $\bar{x}'_t$ and, after summation, $\bar{Y}'_R$ per herd * age in months * month of calving subclass (Type II means) were derived. These means were compared with the means $\bar{x}_t$ and $\bar{Y}_R$, directly calculated in data set II (Type I means).

RESULTS

Estimation of means

Lactation curves were fit per herd on age and season corrected test-day records by (11). Mean values of the parameters are in table 2 per herd level for data set I. Higher herd productions were associated with a higher level of the lactation curve and a slightly lower persistency. The coefficients of determination (R^2) were

Table 2. Mean values of parameters in herd lactation curves ($y = a + t + ce^{-.05}$) per herd level for milk yield (kg) for data set I

	level		
	low	medium	high
Y_{305}	4090	4622	5233
$\hat{a}$	17.0	19.9	22.1
$\hat{b}$	-0.028	-0.029	-0.030
$\hat{c}$	-2.043	-2.043	-3.819
SE ($\hat{a}$)	0.268	0.264	0.273
SE ($\hat{b}$)	0.001	0.001	0.001
SE ($\hat{c}$)	0.794	0.783	0.811
R^{2a}	0.97	0.97	0.97
$\sqrt{MSE}^b$	0.380	0.374	0.388
n^c	141	294	151

a) R^2 = coefficient of determination
b) MSE = mean square error
c) n = number of herds

all 0.97 while the standard deviation for the residuals was less than 0.388 kg milk. Coefficients for the regression of the parameters a, b, c in (11), on the mean x_{50} and Y_{305} per herd in data set I are presented in table 3. The R^2 values were 0.91 0.72 and 0.19 for a, b and c respectively. In fact, factor c was not important in determining the mean x_t and Y_R for prediction if a last test-day yield beyond 50 days p.p. was used for prediction, because c was associated with the lactation curve in the first two months p.p. Factor 'a' was mainly determined by mean x_{50}. The correlation between 'a' and mean x_{50} was 0.95.

With the coefficients obtained in data set I and presented in table 3, lactation curves per herd for data set II were computed. After back adjustment for age and month of calving, type II means were obtained for each herd by age in months by month of calving subclass. In table 4 these type II means are compared with type I means, directly calculated from data set II. As t increased from 50

Table 3. Coefficients for the regression[a] of the parameters a, b and c of the herd lactation curves on mean test-day milk yield at 50 days post partum ($\bar{x}_{50}$) and mean 305-day milk yield ($\bar{Y}_{305}$) per herd

	a	b	c
u	0.05090	-0.003020	6.29810
v	0.12706	-0.000818	-0.06774
w	-0.00079	0.000027	0.00074
R^2 [b]	0.91	0.72	0.19
$\sqrt{MSE}$ [c]	0.598	0.003	1.829

a) a,b, c = u + v $\bar{x}_{50}$ + w $\bar{Y}_{305}$

b) R^2 = coefficient of determination

c) MSE = mean square error

to 210 days p.p. the correlation between type I and type II means decreased from 0.78 to 0.61 and from 0.69 to 0.51 for mean test-day and mean remaining lactation milk yield respectively. The bias was negligible but the standard deviations of type I minus type II were high. This might be caused by the fact that type I means were often based on 1 observation. In total there were 8743 herd * age in months * month of calving subclasses based on 12578 observations.

Table 4. Comparison of type I and type II mean test-day milk yield at t days post partum ($\bar{x}_t$) and mean remaining lactation milk yield ($\bar{Y}_{Rt}$)[a]

	type I		type II			type II minus type I	
	mean[b]	SD	mean	SD	correlation	mean	SD
$\bar{x}_{50}$	25.6	5.3	25.0	4.2	0.78	-0.6	3.3
$\bar{x}_{210}$	16.4	3.5	16.4	2.4	0.61	0.0	2.8
$\bar{Y}_{R50}$	4366	809	4369	630	0.69	3.0	591
$\bar{Y}_{R210}$	1106	264	1092	187	0.51	-14.0	234

a) Type I: calculated in data set II per herd * age (months) * month of calving.

Type II: calculated from the estimated mean herd lactation curve after back adjustment.

b) kg

Table 5. The eigenvalues of the pooled correlation matrix between test-day milk yields

lactation number	Eigenvalues														
1	10.83	1.65	0.80	0.52	0.32	0.23	0.18	0.14	0.10	0.07	0.05	0.04	0.03	0.03	0.01
2	10.54	2.09	0.80	0.48	0.30	0.22	0.16	0.12	0.09	0.07	0.04	0.04	0.03	0.02	0.01
3	10.12	2.13	0.95	0.56	0.33	0.24	0.17	0.13	0.09	0.07	0.05	0.04	0.03	0.03	0.03

Therefore the standard deviation of differences might contain considerable random error variation.

Comparison of prediction methods

The eigenvalues of pooled correlations between test-day milk yield are presented in table 5 in descending order. These eigenvalues correspond to 15 new independent variables, explaining the entire

Table 6. The factor loadings for the test-day milk yields (x_t) in lactation 1

days post partum	factor 1	factor 2
10	-0.59	0.34
30	-0.80	0.44
50	-0.84	0.40
70	-0.86	0.36
90	-0.88	0.27
110	-0.89	0.19
130	-0.90	0.10
150	-0.90	0.00
170	-0.90	-0.07
190	-0.90	-0.15
210	-0.88	-0.26
230	-0.87	-0.34
250	-0.85	-0.40
270	-0.80	-0.44
290	-0.70	-0.38

correlation matrix. The variation explained by such a variable is represented by the eigenvalue. The variation explained by the new variables 1 and 2 was 83 %, 84 % and 82 % for lactation number 1, 2 and greater than 2 respectively. The third eigenvalue was less than 1, and therefore the factor loading matrix consisted of 2 factors. The values of the unrotated factor loadings were close to -1 for factor 1 and ranged from 0.44 to -0.44 for factor 2 (table 6). In the middle of the lactation factor 1 was -0.90 and factor 2 was zero, indicating that factor 1 was associated with level of production and factor 2 with persistency. Similar results were found by Van Arendonk and Fimland (1983).

Theoretical values for the standard deviations of Y_{305}, $\hat{Y}_{305}$ and $Y_{305}-\hat{Y}_{305}$ and the correlation between Y_{305} and $\hat{Y}_{305}$ and between Y_R and $\hat{Y}_R$ for MR,SR and FA are shown in table 7. With regard to these statistics the three prediction methods were not different. The correlation between Y_{305} and $\hat{Y}_{305}$ increased from 0.83-0.89 to 0.99 and between Y_R and $\hat{Y}_R$ from about 0.73-0.78 to 0.83-0.85 as t, the day for the last test p.p., progressed from 50 to 210. Standard deviations for $Y_{305}-\hat{Y}_{305}$ were then decreasing. There were, however,

Table 7. Theoretical comparison of multiple regression (MR), factor analysis (FA) and single regression (SR) for the prediction of a 305-day milk yield (Y_{305})[a)]

			Last test at 50 day post partum			Last test at 210 day post partum		
	Lactation	SD	SD	correlation		SD	correlation	
Method	number	(Y_{305})	($Y_{305}-\hat{Y}_{305}$)	($Y_{305},\hat{Y}_{305}$)	($Y_R,\hat{Y}_{305}$)	($Y_{305}-\hat{Y}_{305}$)	($Y_{305},\hat{Y}_{305}$)	($Y_R,\hat{Y}_R$)
MR	1	750	379	0.86	0.78	118	0.99	0.85
	2	913	481	0.85	0.75	139	0.99	0.85
	≥ 3	955	528	0.83	0.73	153	0.99	0.83
FA	1	750	380	0.85	0.78	121	0.99	0.84
	2	913	482	0.89	0.75	143	0.99	0.84
	≥ 3	955	529	0.83	0.73	157	0.99	0.82
SR	1	750	380	0.86	0.78	121	0.99	0.84
	2	913	482	0.85	0.75	140	0.99	0.85
	≥ 3	955	529	0.83	0.73	153	0.99	0.83

a) kg

small differences between lactation numbers. The 305-day milk yield for heifers was predicted more accurately than for older cows with all methods.

For comparison of MR, FA and SR in data set II, the criteria are presented in table 8. The means, x_i and Y_R, necessary for prediction were of type II. The correlation between Y_{305} and $\hat{Y}_{305}$ increased from about 0.86 to 0.99 and between Y_R and $\hat{Y}_R$ from about 0.78 to 0.86 as t, the day for the last test, progressed from 50 to 210. There were only slight differences between lactation numbers. The 305-day milk yields for heifers were predicted more accurately than for older cows.
The standard deviations of $Y_{305}-\hat{Y}_{305}$ decreased from about 390 kg to 120 kg, from about 460 kg to 137 kg and from about 501 kg to 148 kg for lactation number 1, 2 and greater than 2 respectively for both MR and SR.

Table 8. Comparison of multiple regression (MR), factor analysis (FA) and single regression (SR) for the prediction of a 305-day milk yield[a] (Y_{305}) in data set II) with use of type II means

				Last test at 50 days post partum				Last test at 210 days post partum			
	Lactation	Y_{305}		$Y_{305}-\hat{Y}_{305}$				$Y_{305}-\hat{Y}_{305}$			
method	number	mean	SD	mean	SD	$r(Y_{305},\hat{Y}_{305})$	$r(Y_R,\hat{Y}_R)$	mean	SD	$r(Y_{305},\hat{Y}_{305})$	$r(Y_R,\hat{Y}_R)$
MR[b]	1	4944	777	- 75	390	0.87	0.80	-20	120	0.99	0.88
	2	5836	940	- 96	460	0.88	0.80	-21	137	0.99	0.87
	≥ 3	6374	982	-148	501	0.86	0.78	-29	148	0.99	0.86
FA[b]	1	4944	777	- 75	392	0.87	0.80	-14	125	0.99	0.87
	2	5836	940	- 92	460	0.88	0.80	-12	141	0.99	0.87
	≥ 3	6374	982	-147	502	0.86	0.78	-19	153	0.99	0.85
SR[b]	1	4944	777	- 76	390	0.87	0.80	-21	120	0.99	0.87
	2	5836	940	-100	459	0.88	0.80	-23	137	0.99	0.87
	≥ 3	6374	982	-147	502	0.86	0.78	-30	148	0.99	0.86
SR[c]	1	4944	777	- 79	384	0.87	0.81	-22	120	0.99	0.87
	2	5836	940	- 98	459	0.88	0.80	-24	137	0.99	0.87
	≥ 3	6374	982	-147	501	0.86	0.78	-30	148	0.99	0.86

a) kg

b) prediction factors were pooled within parity 1, 2 and higher

c) prediction factors were per age * season subclass

For FA, the standard deviation of $Y_{305}-\hat{Y}_{305}$ was slightly higher if the last test was known at 210 days p.p. The bias was negative in all situations, indicating an overprediction. If the last test was at 50 days p.p., this overprediction was more than 100 kg for lactation numbers greater than 1. There were no differences between the methods MR, FA and SR. This indicates that the use of the last test only (SR) was as good as the use of all tests (FA, MR), which is in agreement with Dommerholt (1975) and Van Arendonk and Fimland (1983).

Considering the sensivity of prediction factors, correlations were pooled within parities. Using SR, mean prediction error and its standard deviation and correlations between Y_{305} and $\hat{Y}_{305}$ did not change when different sets of prediction factors were used per age * season class (table 8)

When looking at the validity of type II means, the standard deviation of $Y_{305}-\hat{Y}_{305}$ for SR and second or higher parity cows was found to be about 20 kg lower than the theoretical values if the last test was known at 50 days p.p. For heifers, the theoretical value was 10 kg lower. When the last tests were known at 210 days post partum almost no differences were found. Theoretical correlations and those from data set II were nearly equal. In table

Table 9. The prediction of 305-day milk yield[a] (Y_{305}) at high and low level herds, using the last known test-day yield with single regression (SR) and type II means

herd level	lactation number	Y_{305}		Last test at 50 days post partum				Last test at 210 days post partum			
				$Y_{305}-\hat{Y}_{305}$				$Y_{305}-\hat{Y}_{305}$			
		mean	SD	mean	SD	$r(Y_{305},Y_{305})$	$r(Y_R,\hat{Y}_R)$	mean	SD	$r(Y_{305},\hat{Y}_{305})$	$r(Y_R,\hat{Y}_R)$
low	1	4361	604	-114	351	0.84	0.75	-18	116	0.98	0.83
	2	5062	703	-120	382	0.85	0.75	-16	135	0.98	0.81
	≥ 3	5583	770	-154	450	0.82	0.72	-24	139	0.98	0.82
high	1	5509	718	- 18	419	0.82	0.73	-17	123	0.99	0.84
	2	6583	869	- 52	514	0.81	0.70	-15	146	0.99	0.84
	≥ 3	7119	895	-137	537	0.81	0.69	-30	156	0.98	0.82

a) kg

9 mean prediction errors with their standard deviation and correlations between Y_{305} and $\hat{Y}_{305}$ are presented for SR at low and high level herds. Mean prediction error in first lactation was -114 and -18 kg for low and high production level respectively, when the last test was known at 50 days p.p. Difference in bias was therefore 96 kg between both levels. For second and higher parity cows these differences were 68 kg and 17 kg respectively. When the last test was known at 210 days p.p. almost no differences were found between mean prediction errors for low and high level herds. Correlations between Y_{305} and $\hat{Y}_{305}$ were slightly lower than for overall (table 8).

DISCUSSION AND CONCLUSIONS

The unrotated factor loadings, presented in table 6, showed that factor 1 was associated with level of production and factor 2 with persistency. Test-day yields between 130 and 190 days p.p. were most important in determining factor 1, whereas yields at 30-50 and 250-270 days p.p. were most important for factor 2. The weights in factor 2 for yields between 130 and 170 days p.p. were almost zero. From the eigenvalues, factor 1 explained over 72 % of the total variation of all test-day yields for heifers. Therefore the variation in Y_{305} was mainly determined by the variation of test-day yields between 130 and 190 days p.p. Factor 2 explained another 11 %. Because factor 1 and 2 are uncorrelated, a selection for factor 2 as an alternative parameter for persistency, independent of production level, might be interesting from an economic view point.

This study has demonstrated that SR, MR and FA are identical, both theoretically and in dataset II with type II means, for the prediction of 305-day yields. For practical purposes SR is recommended because of its simplicity. With type II means, a use of prediction factors within parity 1, 2, 3 and greater or within age * season of calving classes did not reveal differences in prediction with regard to bias and accuracy. Prediction factors may vary among herd levels. However, prediction factors were computed as a ratio of variances of two variables multiplied by their correlation. When herd level of production increases, the variance of production increases as well (table 9). Dommerholt (1975) showed that the ratio of total and part lactation milk yield was insensitive to herd level of production. Therefore prediction factors are expected to be insensitive to herd level.

With regard to the estimation of means (of type II) the standard deviations of $Y_{305}-\hat{Y}_{305}$ for SR and second or higher parity cows were in most cases lower than theoretical values. Furthermore, the differences in mean prediction error between low and high level herds varied from 1 to 96 kg. The difference in mean 305-day production between low and high level was 1148 kg for heifers. This indicates that means per herd * age at calving in months * month of calving subclasses were rather well estimated. The fact that mean prediction error for low level herds was higher than for high level herds may be caused by year influences. Adjustment for differences in production level between years for the estimation of mean last known test-day yield might be worthwhile, by using rolling population averages for cows of the same age, freshened in the same season and in the same stage of lactation. Further analysis showed that mean prediction error decreased by only 20 kg (Wilmink, 1985).

Mean x_{50} and Y_{305} per herd were used as the main factors in determining a lactation curve. When the parameters a, b, c in (11) were regressed on mean x_{50} and x_{270} per herd, slightly higher values for R^2 were obtained which were 0.94, 0.93 and 0.23 respectively. However the correlations between type I and type II means had the same values as in table 4.

Disregarding the lactation number in calculating the criteria for comparison in data set II, the correlation between Y_{305} and $\hat{Y}_{305}$ was 0.91 for all methods when the last test is at 50 days p.p., in which case $\sigma(Y_{305}-\hat{Y}_{305})$ was 468 kg and level of production was 5903 kg. Because of the proportionality between mean Y_{305} and the standard deviation of $Y_{305}-\hat{Y}_{305}$, a measure of the goodness of prediction would be $100 * \sigma(Y_{305}-\hat{Y}_{305})/ (\bar{Y}_{305})$ which was 7.9 in this study. In Dommerholt's (1975) study the value was 9.9 and 7.7 if the last test was at 30 and 70 days p.p. respectively. In the study of Wiggans (1980), using test-day production, herd average and level of the partial record as information sources, a value of 9.7 was obtained for lactations of 35-64 days in milk. These results indicate that accuracy of prediction is improved by estimation of type II means.

In conclusion, the estimation of means within herd * age * month of calving subclasses improved the accuracy of prediction. No differences were found between multiple regression, factor analysis and single regression of remaining yield on the last known test-day yield as prediction methods. For practical use the last method is recommended.

REFERENCES

Auran, T. (1976). Studies on monthly and cummulative monthly milk yield records. 4. Estimating total lactation from part lactation. Acta Agric. Scand., 26:10-17.

Congleton, W.R. and R.W. Everett (1980). Application of the incomplete gammafunction to predict cummulative milk production. J. Dairy Sci., 63:109-119.

Danell, B. (1982). Studies on lactation yield and individual test-day yields of Swedisch dairy cows. 4. Extension of part lactation records for use in sire evaluation. Acta Agric. Scand., 32:103-114.

Dommerholt, J. (1975). Correctie van de melkgift van koeien voor verschillen in leeftijd, seizoen en lactatiestadium (with English summery). Agricultural research reports 844, Centre for Agricultural. Publishing and Documentation, Wageningen.

Harman, H.H. (1970). Modern Factor Analysis. The univ. of Chicago Press, Chicago and London.

Schaeffer, L.R.; Minder, C.E.; McMillan, I. and E.B. Burnside (1977). Nonlinear techniques for predicting 305-day lactation production of Holsteins and Jerseys. J. Dairy, Sci. 60:1636-1644.

Van Arendonk, J.A.M. and E. Fimland (1983). Comparison of two methods to extend partial milk records. Z. Tierz. und Züchtungsbiol. 100:33-38.

Wiggans, G.R. (1980). Contribution of days since bred, herd average, and sample-day production to accuracy in projecting lactation records. J. Dairy, Sci. 63:984-988.

Wiggans, G.R. and L.D. Van Vleck (1979). Extending partial lactation milk and fat records with a function of last-sample production. J. Dairy Sci. 62:316-325.

Wilmink, J.B.M. (1985). Het voorspellen van een onbekende proefmelking, deellijst en 305-dagen produktie in de lopende lactatie (with English summary). Report, B276 IVO "Schoonoord", Zeist.
Wilmink, J.B.M. (1986). Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. Submitted to Livestock Prod. Sci.

Chapter 4

EFFICIENCY OF SELECTION FOR DIFFERENT CUMULATIVE MILK, FAT AND PROTEIN YIELDS IN FIRST LACTATION

J.B.M. Wilmink

Royal Dutch Cattle Syndicate
P.O. Box 454, 6800 AL Arnhem

Submitted to Livestock Production Science

ABSTRACT

The objective was to determine the efficiency of selection for different cumulative milk, fat and protein yields in first lactation, which was defined as the ratio of correlated and direct annual response in 305-day yield. Cumulative milk, fat and protein yields for different periods in first lactation were calculated from test-day yields of 20260 Black- and White cows, sired by 204 young and 97 proven bulls. Heritabilities and genetic correlations with 305-day yield were calculated from the between- and within-sire variance/covariance components estimated by restricted maximum likelihood.

Genetic correlations with 305-day yield were close to 1.00 for 180-day milk, fat and protein yield, extended to their 305-day equivalents and for cumulative milk, fat and protein yield in a period between 91 and 180 days post partum. Heritabilities for extended 180-day yield and cumulative milk or protein yield between 91 and 180 days post partum were close to heritablilities for the 305-day yield, which were 0.309, 0.372 and 0.331 for milk, fat and protein respectively. The heritability for fat yield between 91 and 180 days post partum was 0.318.

A selection for 305-day yield could be replaced by a selection for milk and protein yield between 91 and 180 days post partum and 180-day fat yield extended to 305 days, in which case the annual response in 305-day yield would be increased by 5, 6 and 4% respectively.

INTRODUCTION

The advantages of using part lactation records of dairy heifers in predicting sires' breeding values for first lactation yield are 1) shortening of the generation interval, 2) reduction of bias by selection and 3) saving in expences for housing and measurement of other traits, for example milking speed, by an early culling of bulls with low breeding values for milk production (Van Vleck and

Henderson, 1961b; Danell, 1982b). With regard to bias by selection, Fimland (1983) pointed out that the largest potential for non random influence in sire evaluation was related to culling of heifers.

Traditionally part lactation records have been extended to 305-day equivalents, when predicting breeding values for sires (Famula and Van Vleck, 1981; Danell, 1982b; Tandon and Harvey, 1984). Other methods base sire evaluations on part lactations such as yields from day 0 to 100, day 101 to 200 and day 201 to 305 post partum (Lederer and Buthman, 1982) or yields from day 0 to 90, day 91 to 180 and day 181 to 270 (Agyemang et al., 1985b). In the Netherlands, breeding values of sires are predicted for milk production traits in the first 100 days and subsequently in the first 305 days of first lactations (Dommerholt and Wilmink, 1983). A selection of young bulls is then carried out in two stages.

Danell (1982, b) pointed out that a disadvantage of extending part lactations to 305-day equivalents is that level of production may vary in time resulting in biased predictions of 305-day yields. Wilmink (1986) presented a method in which different means, being the greatest potential for biased predictions, were estimated from averages within herd in the previous year. It was shown that accuracy of prediction was increased and that sensitivity for differences in production level was greatly reduced and was only restricted to differences from year to year.
However, extended part lactations of different length are essentially different traits. Genetic correlations are less than one when lactation length is shorter than 200 days (Auran, 1976b) and therefore, problems may arise in comparing extended lactations of different length within a herd-year-season. In addition, little is known about the validity of prediction methods to extend incomplete milk, fat and protein records.

In the method, proposed by Agyemang et al. (1985a), extension is only required over short periods. Factors were presented for the extension of part lactations to 90, 180 or 270 days post partum, giving trimester milk and fat yields. A multitrait best linear unbiased prediction (BLUP) analysis was performed to the trimester milk yields by transformation to canonical scale (Agyemang et al., 1985b). This effectively reduces a multitrait to single trait analysis. However, no adjustment can be made for selection since all trimester yields of a cow should be known. If no linear transformation is performed a multitrait analysis for 3 trimester yields becomes difficult when used for practical purposes because of the big increase of the number of equations (Agyemang et al., 1985b). It might therefore be worthwhile to reduce the number of traits to, for example, one or two cumulative yields over a period in first lactation such that the genetic correlation with 305-day yield is the highest, potential of bias by heifer selection has reduced and the annual response in 305-day yield is increased. Such a trait might also be useful for the selection of cows.

The purpose of this study was to estimate heritabilities for cumulative or extended milk, fat and protein yield in different intervals in first lactation, to estimate genetic correlations with 305-day yield and to calculate ratio's of annual responses in 305-day yield when selection was for cumulative or extended yield.

MATERIAL

Test-day milk, fat and protein yields in first lactation of 24692 Black and White cows, calving from May 1981 until June 1984 in 663 herds, drawn at random, were extracted from the milk recording files. All cows were between 21 and 33 months old at calving, had a minimum lactation length of 270 days and were sired by A.I.
If lactation length was less than 305 days, lactation yield was extended to a 305 basis, using the formula proposed by Poutous and Mocquot (1975).
All test-day records were known at an approximately 3 or 4 week interval with an exception for the holiday period. The first test was carried out when a cow was at least 10 days in milk. In order to calculate cumulative yields in different periods, firstly daily yields at 10,......,310 days post partum (pp) were computed by inter- or extrapolation of the two nearest test-day yields, which

were coded as normal. Abnormal codes were assigned, for example, if a cow was in heat. At the end of lactation, an extrapolation of daily yield was performed if the period was not more than 10 days.

Cumulative yields in different periods are designated as M_{i-j} for milk, F_{i-j} for fat and P_{i-j} for protein, where i and j refer to the number of days pp., reflecting the lower and upper limit of a period. Predicted 305-day yields are designated as PM_{1-j}, PF_{1-j} and PP_{1-j} for milk, fat and protein respectively.

Cumulative yields were calculated by summing the daily yields, multiplied by 5, 10 or 20 days. In this case the first interval was always 20 days. For example, $M_{1-30} = 20 * x_{10} + 10 * x_{30}$ and $M_{1-60} = 20 * x_{10} + 20 * x_{30} + 20 * x_{50}$, where x_{10}, x_{30}, x_{50} is the daily milk yield at 10, 30, 50 days pp. respectively. Predicted 305-day milk, fat and protein yield were calculated according to procedures, recommended by Wilmink (1986). Prediction factors and factors for estimating the mean herd lactation curve for milk, fat and protein were derived from the test-day records of all cows, used in that study.

Table 1. Overall means and standard deviations[a] for monthly and 305-day milk, fat and protein yield

Days in lactation	milk (kg) mean	milk (kg) SD	fat ($kg*10^{-1}$) mean	fat ($kg*10^{-1}$) SD	protein ($kg*10^{-1}$) mean	protein ($kg*10^{-1}$) SD
1- 30	582	103	244	53	190	35
31- 60	599	97	239	43	186	30
61- 90	578	96	229	40	183	31
91-120	553	95	221	39	179	31
121-150	523	94	213	38	174	31
151-180	499	94	208	38	169	31
181-210	467	95	201	38	162	32
211-240	441	96	195	39	155	32
241-270	403	96	187	40	145	33
1-305	5080	831	2142	346	1701	272

a) based on 4554 daughters of young bulls

Genetic parameters were estimated by half-sib analysis of daughters of test bulls, requiring an inversion of a matrix with dimension equal to the number of sires. Records of young-bulls with less than 8 daughters or of proven bulls with less than 25 daughters in the data set were removed, resulting in 4554 records of 204 young and 15706 records of 97 proven sires. Total number of records was therefore 20260. Overall means and standard deviations for monthly yields and 305-day yields, based on 4554 records of young sires, are presented in table 1.

METHODS

Estimation of variance/covariance components

The model, applied to the data, was:

$$y_{ijkl} = hys_i + br_j + b_1 age_l + b_2 age_l^2 + s_k + e_{ijkl} \quad (1)$$

where y_{ijkl} = M_{i-j}, F_{i-j}, P_{i-j}, PM_{i-j}, PF_{i-j} and PP_{i-j} of cow l;
hys_i = fixed effect of herd-year-season i (i = 1,4221);
br_j = fixed effect of breed j of the cow (j = 1,4);
b_1, b_2 = coefficients for the regression of y_{ijkl} on age at calving and age at calving squared (in months);
s_k = random/fixed effect of sire k (k = 1,97 for proven sires, k = 98,301 for young sires);
e_{ijkl} = random residual effect.

Proven sires were considered as fixed because their daughters might originate from selective matings.
For young sires $E(s_k)$ = 0 and $var(s_k)$ was set to σ_s^2. Thus only young sires were considered in variance component estimation. Furthermore $E(e_{ijkl})$ = 0 and $var(e_{ijkl})$ = σ_e^2.
Young sires were assumed to be unrelated and uncorrelated with the residuals.

Seasons were defined from February through August and September through January as is done in the Dutch sire evaluation programs.

This resulted in 7 year-seasons. Breed classes were defined for 1) 100%,2) 87.5, 75.0 and 62.5%,3) 50 and 37.5% and 4) 25, 12.5 and 0% Dutch Friesian. All crosses were Dutch * Holstein Friesian. Lactation length was not considered in the model. A preliminary analysis showed that the maximum least square contrasts with respect to the overall mean was 6,6% for $M_{121-150}$ and 8,8% for M_{1-305}. As $M_{121-150}$ is not influenced by pregnancy effects, differences in yield between different lactation lengths were mainly caused by an earlier culling of low producing heifers and/or allowing more inseminations for high producing heifers (Van den Broek, 1986). An adjustment for lactation length would therefore eliminate some genetic variation in production traits.

As all traits were observed on all heifers and the design matrices were equal, a transformation of the original traits was performed to canonical scale in estimating between- and within-sire variance/covariance components. Restricted maximum likelihood (REML) variance component estimates were obtained for variaties on canonical scale and, after back transformation to original scale, the variance/covariance components for variaties on original scale. Computing algoritm is described in Meyer (1985). Iteration was stopped when the change in estimates between successive rounds was below 1%. Standard errors were obtained from the inverse of the information matrix.

Priors were taken from an univariate analysis for model 1. In that analysis REML variance components were estimated and covariance components were calculated from 1) the correlation between sires' breeding' values or between residuals and 2) the estimated variance components. In the multitrait analysis with canonical transformation, on maximum of 4 traits, which were not too closely related, were considered at one time. Consideration of more traits led to divergence when iterating.

Genetic variance/covariance were estimated as 4 times the estimated between sire variance/covariance. Phenotypic variance/covariance was equal to the sum of estimated between- and within-sire variance/covariance.

Efficiency of selection

The annual genetic gain (G_y) in 305-day yield by selection for y (see model 1) in an A.I.-breeding plan, was calculated by (Skjervold and Langholz, 1964):

$$G_y = \frac{R_{SS} + (1-p)R_{SD} + R_{DS} + R_{DD}}{L_{SS} + (1-p)L_{SD} + pL_{YD} + L_{DS} + L_{DD}}$$

where R = genetic response in 305-day yield after one generation selection for y for the paths sires to breed sons (SS), sires to breed dams (SD), dams to breed sires (DS) and dams to breed dams (DD);

L = generation interval for the paths SS, SD, YD (young bulls to breed dams), DS and DD;

p = proportion of first inseminations with young bulls, which was taken to be 0.2.

The genetic response, R, is equal to $i.r.r_g\,\sigma_H$, where i reflects the selection intensity, which was set to 1.4, 2.06, 2.66 and 0.50 for SS, SD, DS and DD respectively, r the repeatability of the estimated breeding value for y, r_g the genetic correlation between y and 305-day yield and σ_H the additive genetic standard deviation of 305-day yield. The repeatability, r, was calculated as $(n/(n + \frac{4 - h^2}{h^2}))^{\frac{1}{2}}$ for the paths SS and SD, where n is the number of daughters, which was taken to be 50 and 30 respectively, and h^2 the heritability of y. For paths DS and DD, it was assumed that selection was based on a cow index including cow's own performance for y and the breeding values of her parents. The repeatability of this index was calculated as $(b + 0.25\,(r^2_{SD} + r^2_{DD}))^{\frac{1}{2}}$, where b is the weighting factor for cow's record in the index and r^2_{SD}, r^2_{DD} are squared repeatabilities of parent's estimated breeding values (Moen, 1978). The term r^2_{DD} was set on 0.397 for all cases, whereas r^2_{SD} was equal to the repeatabilty squared for path SD. When selection was for 305-day yield the generation interval L was equal to 7, 6, 4.5, 4.5 and 2 years for paths SS, SD, DS, DD and YD respectively. For other traits and paths SS, SD and DS, L was diminished by a factor (305-j)/365, where j is the last day post partum of the partial record. The effect of shorter generation interval for path DD was not considered, because R_{DD} is rather small.

Efficiency of selection (E) was calculated as the ratio in percentages between correlated response/year for 305-day yield when selection was for y (G_y) and direct response/year for 305-day yield (G).

RESULTS

Heritabilities and variances between sires

Heritabilities and estimates for variance components between sires for partial, predicted and complete lactation milk, fat and protein yield are presented in tables 2 - 4. Heritabilities were 0.309, 0.372 and 0.331 for M_{1-305}, F_{1-305} and P_{1-305} respectively. Additive genetic standard deviations were 340, 15.4 and 11.0 kg respectively. Considering monthly milk yields (table 2) the heritabilities increased from 0.159 for M_{1-30} to 0.304 for M_{91-120} followed by a gradual decrease to 0.250 for $M_{211-240}$. Heritabilities were highest in mid lactation. Between sire variances increased from 312 kg^2 for M_{1-30} to 410 kg^2 for M_{61-90} and M_{91-120}, followed by a decrease to 302 kg^2 for $M_{211-240}$. For monthly protein yields, the heritabilities and between sire variance component estimates showed a similar pattern as for monthly milk yield (table 4). Highest values were 0.302 and 0.40 kg^2 respectively for $P_{121-150}$. The fat yield estimates (table 3) showed a somewhat irregular pattern. Heritabilities ranged between 0.131 and 0.325 and were lowest for F_{1-30} and $F_{241-270}$. Between sire variance components ranged from 0.48 to 0.86 kg^2.
Heritabilities increased for cumulative yields over a two months period. Maximum values were 0.314 for M_{91-150}, 0.332 for F_{31-90} and $F_{151-210}$ and 0.345 for P_{91-150}. For both M_{91-150} and P_{91-150} the heritabilities were higher than for the 305-day equivalents. Considering cumulative yields over a 3 months period another increase in heritablities was observed. For partial lactation and predicted

Table 2. Heritabilities, genetic and phenotypic correlations with 305-day milk yield, between-sire variance components[a] and efficiency of selection (E)[b] for cumulative (M_{i-j})[c] and extended part lactation (PM_{i-j}) milk yield

	h^2	se	r_g	se	r_p	se	$\hat{\sigma}^2_s$	se[d]	E
M_{1-30}	0.159	0.044	0.611	0.128	0.661	0.008	312	85	76
M_{31-60}	0.252	0.049	0.882	0.071	0.828	0.006	384	74	99
M_{61-90}	0.288	0.063	0.942	0.046	0.867	0.005	410	89	104
M_{91-120}	0.304	0.062	0.953	0.046	0.879	0.005	409	82	104
$M_{121-150}$	0.295	0.069	0.994	0.041	0.889	0.004	386	90	105
$M_{151-180}$	0.286	0.063	0.993	0.043	0.886	0.004	366	79	103
$M_{181-210}$	0.293	0.056	0.949	0.043	0.875	0.004	368	69	100
$M_{211-240}$	0.250	0.049	0.892	0.057	0.845	0.005	302	59	93
$M_{241-270}$	0.257	0.052	0.800	0.078	0.780	0.007	313	63	87
M_{1-60}	0.221	0.046	0.777	0.086	0.784	0.007	1374	286	90
M_{31-90}	0.292	0.060	0.921	0.054	0.880	0.005	1602	327	103
M_{61-120}	0.310	0.065	0.945	0.045	0.895	0.004	1638	339	104
M_{91-150}	0.314	0.066	0.976	0.036	0.914	0.004	1560	326	105
$M_{121-180}$	0.305	0.063	0.996	0.029	0.911	0.004	1502	307	104
$M_{151-210}$	0.301	0.062	0.980	0.036	0.911	0.004	1423	293	102
$M_{181-240}$	0.276	0.055	0.931	0.051	0.886	0.005	1286	256	96
M_{31-120}	0.313	0.066	0.935	0.046	0.908	0.004	3607	756	104
M_{61-150}	0.326	0.068	0.965	0.036	0.925	0.004	3597	745	105
M_{91-180}	0.320	0.067	0.985	0.030	0.933	0.003	3390	706	105
PM_{1-60}	0.251	0.052	0.853	0.070	0.832	0.005	14089	2910	97
PM_{1-120}	0.282	0.059	0.960	0.039	0.926	0.004	19287	4004	103
PM_{1-180}	0.292	0.058	0.993	0.020	0.968	0.002	23270	4594	103
PM_{1-240}	0.294	0.062	0.999	0.009	0.992	0.001	26097	5476	101
M_{1-60}	0.221	0.046	0.777	0.086	0.784	0.007	1374	286	90
M_{1-120}	0.280	0.057	0.897	0.053	0.899	0.004	5574	1126	99
M_{1-180}	0.304	0.056	0.955	0.034	0.953	0.003	2191	2235	102
M_{1-240}	0.302	0.057	0.989	0.017	0.985	0.003	1963	3666	101
M_{1-305}	0.309	0.060	1.000	--	1.000	--	29104	5631	100

a) kg^2

b) ratio of correlated and direct response/year for 305-day milk yield

c) day i to j post partum

d) lower bound for large sample s.e.

305-day yields the heritabilities and between sire variance components showed an increase to the values for 305-day yields, as the length of the partial record increased. Heritabilities increased when 60-day records were extended to 305-day equivalents.

Table 3. Heritabilities, genetic and phenotypic correlations with 305-day fat yield, between-sire variance components[a] and efficiency of selection (E)[b] for cumulative (F_{i-j})[c] and extended part lactation (PF_{i-j}) fat yield

	h^2	se	r_g	se	r_p	se	$\hat{\sigma}_s^2$	se[d]	E
F_{1-30}	0.131	0.023	0.832	0.064	0.645	0.006	71	12	83
F_{31-60}	0.269	0.055	0.916	0.063	0.780	0.006	86	17	99
F_{61-90}	0.294	0.055	0.950	0.046	0.805	0.006	76	15	101
F_{91-120}	0.274	0.056	0.975	0.044	0.827	0.006	62	12	100
$F_{121-150}$	0.245	0.051	0.999	0.033	0.839	0.005	51	10	98
$F_{151-180}$	0.275	0.056	0.999	0.037	0.838	0.005	57	12	99
$F_{181-210}$	0.325	0.056	0.970	0.040	0.834	0.005	67	11	99
$F_{221-240}$	0.280	0.052	0.916	0.055	0.810	0.006	59	11	93
$F_{241-270}$	0.227	0.026	0.873	0.070	0.768	0.006	48	6	86
F_{1-60}	0.229	0.045	0.883	0.070	0.757	0.007	341	66	94
F_{31-90}	0.332	0.055	0.942	0.055	0.842	0.006	343	60	103
F_{61-120}	0.322	0.063	0.969	0.044	0.851	0.005	287	56	103
F_{91-150}	0.297	0.060	0.992	0.035	0.879	0.005	231	51	101
$F_{121-180}$	0.290	0.058	0.998	0.025	0.871	0.004	223	44	100
$F_{151-210}$	0.331	0.060	0.994	0.040	0.876	0.005	254	48	101
$F_{181-240}$	0.325	0.064	0.955	0.048	0.850	0.005	259	50	98
F_{31-120}	0.349	0.076	0.951	0.044	0.881	0.005	703	134	103
F_{61-150}	0.332	0.065	0.983	0.034	0.897	0.004	581	113	103
F_{91-180}	0.318	0.063	0.998	0.021	0.907	0.003	519	102	102
PF_{1-60}	0.268	0.054	0.930	0.062	0.800	0.006	2130	428	100
PF_{1-120}	0.309	0.060	0.975	0.036	0.907	0.004	3249	632	102
PF_{1-180}	0.365	0.071	0.999	0.022	0.962	0.002	4793	922	104
PF_{1-240}	0.366	0.072	0.999	0.009	0.989	0.001	5529	1072	102
F_{1-60}	0.229	0.045	0.883	0.070	0.757	0.007	341	66	94
F_{1-120}	0.333	0.061	0.919	0.049	0.879	0.005	1290	235	100
F_{1-180}	0.347	0.066	0.972	0.029	0.947	0.003	2407	451	102
F_{1-240}	0.369	0.071	0.998	0.016	0.985	0.002	4026	770	102
F_{1-305}	0.372	0.072	1.000	--	1.000	--	5978	1146	100

a) $kg^2 * 10^{-2}$

b) ratio correlated and direct response/year for 305-day fat yield

c) day i to j post partum

d) lower bound for large sample s.e.

Genetic correlations

The genetic correlations between cumulative partial or predicted 305-day yield and 305-day yield are also given in the tables 2 - 4. Cumulative yield in mid-lactation was highly correlated with 305-day yield. The genetic correlations were equal to 1 for all cumulative monthly and two-monthly yields between 121 and 180 days post partum. Genetic correlation were lowest at the beginning and

Table 4. Heritabilities, genetic and phenotypic correlations with 305-day protien yield, between-sire variance components[a] and efficiency of selection (E)[b] for cumulative (P_{i-j})[c] and extended lactation (PP_{i-j}) protein yield

	h^2	se	r_g	se	r_p	se	$\hat{\sigma}_s^2$	se[d]	E
P_{1-30}	0.104	0.033	0.762	0.137	0.623	0.008	22	7	77
P_{31-60}	0.183	0.042	0.953	0.080	0.774	0.006	24	6	96
P_{61-90}	0.228	0.045	0.980	0.062	0.817	0.005	29	6	100
P_{91-120}	0.294	0.054	0.995	0.051	0.837	0.005	38	7	105
$P_{121-150}$	0.302	0.056	0.998	0.048	0.850	0.005	40	7	104
$P_{151-180}$	0.264	0.051	0.999	0.048	0.851	0.005	36	7	101
$P_{181-210}$	0.251	0.052	0.963	0.048	0.847	0.005	35	7	97
$P_{211-240}$	0.202	0.044	0.943	0.057	0.831	0.005	28	6	91
$P_{241-270}$	0.186	0.043	0.899	0.073	0.790	0.006	25	6	86
P_{1-60}	0.177	0.038	0.868	0.085	0.741	0.007	105	23	91
P_{31-90}	0.245	0.053	0.978	0.049	0.842	0.005	114	25	102
P_{61-120}	0.292	0.059	0.990	0.014	0.860	0.004	138	28	104
P_{91-150}	0.345	0.069	0.992	0.031	0.885	0.004	163	32	106
$P_{121-180}$	0.318	0.064	0.999	0.019	0.881	0.004	159	32	104
$P_{151-210}$	0.290	0.060	0.993	0.034	0.889	0.004	145	30	101
$P_{181-240}$	0.245	0.051	0.965	0.048	0.869	0.005	127	26	95
P_{31-120}	0.295	0.062	0.983	0.039	0.883	0.004	290	60	104
P_{61-150}	0.341	0.068	0.989	0.032	0.902	0.004	340	68	106
P_{91-180}	0.351	0.069	0.996	0.025	0.911	0.004	360	71	106
PP_{1-60}	0.203	0.045	0.894	0.079	0.772	0.006	861	90	95
PP_{1-120}	0.302	0.062	0.983	0.036	0.897	0.004	1745	358	104
PP_{1-180}	0.327	0.066	0.995	0.021	0.955	0.002	2368	473	104
PP_{1-240}	0.317	0.065	0.999	0.011	0.988	0.001	2659	539	102
P_{1-60}	0.177	0.038	0.868	0.085	0.741	0.007	105	23	91
P_{1-120}	0.266	0.055	0.951	0.048	0.877	0.005	468	96	100
P_{1-180}	0.324	0.064	0.980	0.029	0.944	0.003	1158	227	103
P_{1-140}	0.328	0.067	0.995	0.015	0.983	0.002	1970	400	102
P_{1-305}	0.331	0.068	1.000	--	1.000	--	3028	617	100

a) $kg^2 * 10^{-2}$

b) ratio of correlated and direct response/year for 305-day protein yield

c) day i to j post partum

d) lower bound for large sample s.e.

the end of lactation for milk, fat and protien yield. Genetic correlation between M_{1-30} and M_{1-305} was only 0.611, whereas it was higher than 0.762 for all other cumulative monthly records. Genetic correlations between part lactation or predicted 305-day yield and 305-day yield showed an increase to 1 as the partial record was lengthened. Genetic correlations between predicted 305-day and 305-day yields were much higher than between part lactation and 305-day yields, especially when the part lactation was no longer than 120 days. Genetic correlations were almost equal to 1 when 180-day records were extended to 305-day equivalents, whereas they were equal to 0.955, 0.972 and 0.980 for milk, fat and protein respectively, if no extension was carried out.

Efficiency of selection

The efficiency of selection of bulls and cows on part lactation records, expressed as the ratio in percentages between correlated and direct response for 305-day yield per year, is shown in tables 2-4. For milk yield, highest efficiencies were found for $M_{121-150}$, M_{91-150}, M_{61-150} and M_{91-180}. For the fat traits F_{31-90}, F_{61-120}, F_{31-120}, F_{61-150} and PF_{1-180} efficiencies were higher than 102. The maximum value of 104 was for PF_{1-180}. Considering protein yield the efficiency was highest for P_{91-150}, P_{61-150} and P_{91-180}. When looking at part lactation records up to 180 days post partum, in all cases an increase in efficiency was found when these records were extended to 305-days or when the first month(s) were deleted in calculating cumulative yield.

DISCUSSION AND CONCLUSIONS

For cumulative monthly milk yield heritabilities were highest in mid-lactation which is in agreement with results, reported by Van Vleck and Henderson (1961a), Auran (1976a) and Danell (1982a). The heritabilities for monthly protein yield followed a similar pattern. The monthly fat estimates for heritabilities and variances between sires showed a somewhat irregular pattern, possibly due to seasonal effects within herd-year-seasons. When comparing cumulative yields over an one or three month period in mid-lactation respectively, the heritability increased by 0.03 for milk, by 0.04 for fat and by 0.05 for protein yield. For cumulative milk and protein yield over two or three months in mid-lactation, heritabilities were close to their 305-day values. For fat yield they were approximately 0.06 lower.

Heritabilities for M_{1-30}, F_{1-30} and P_{1-30} were lower than 0.159, indicating that genetic differences for cumulative yields between cows at the beginning of lactation were considerably lower than over other intervals. Van Vleck and Henderson (1961a) found an increase in heritability from 0.11 for M_{1-30} to 0.22 for M_{61-90} and Danell (1982a) reported an increase from 0.16 for M_{1-30} to 0.27 for $M_{151-180}$ in 1972 data of purebred Swedish Red and White cows. When 60-day yield was considered, the heritability was higher for cumulative yield from day 31 to 60 post partum than day 1 to 60 despite the fact that number of observations was higher in the latter case. In addition the genetic correlation with 305-day yield was 0.033-0.105 higher. This shows clearly that a selection for cumulative yield from day 31 to 60 post partum is much more efficient than a selecion for 60-day records.

Genetic correlations between cumulative yield in the period between 121 and 180 days pp. and 305-day yield was equal to 1 for milk, fat and protein, indicating that these traits are influenced by the same genes. Similar results were obtained by Van Vleck and Henderson (1961a), Auran (1976a) and Danell (1982a) for milk yield and by Danell (1982a) for fat yield.
Genetic correlations between M_{91-180}, F_{91-180} or P_{91-180} and their 305-day equivalents were higher than 0.985. A 0.99 genetic correlation was reported between M_{91-180} and M_{1-305} by Agyemang et al. (1985b). Selection of bulls or cows based on, for example, $M_{121-150}$ will result in the same ranking as selection based on M_{1-305}, provided that repeatabilities are equal.
When 180-day records were extended to 305-day equivalents, the genetic correlation with 305-day yield was also equal to one. Genetic correlations betweeen 305-day yields extended from 60 or 120 day yields and real 305-day yields were lower than one, indicating that extended yields can be different traits.

Results with regard to the efficiencies of selection indicated that selection based on part lactation records is not worthwhile. An extension to 305-day equivalents resulted in higher efficiencies, especially for 60 and 120-day yields. On the other hand efficiencies of selection for 60 and 120-day yields were increased with

3-9% when first 30 or 60-days were omitted in calculating cumulative yield. The reason for this is that both the genetic correlations between monthly and 305-day yield and the heritabilities of cumulative monthly yield showed a more or less curvilinear pattern. In part lactation yield, however, all monthly yields are equal weighted.

The maximum increase in annual respons of 305-day yields was when bulls and cows were selected for $M_{121-150}$, M_{91-150}, M_{61-150} or M_{91-180} for milk, PF_{1-180} for fat and P_{91-150}, P_{61-150}, P_{91-180} for protein yield. Apart from repeatabilities, a selection based on, for example, M_{61-150} will result in small changes of ranking of cows or bulls, when compared with selection based on M_{1-305}. The genetic correlation was 0.965 and selection based on M_{61-150} may lead to selection for higher peak yield resulting lower persistency. In order to circumvent such undesired increases of correlated responses, the trait with highest correlated response in 305-day yield and a close to one genetic correlation with 305-day yield may be chosen as the trait to select for.
As a result selection based on $M_{121-150}$, M_{91-180}, PF_{1-180} and P_{91-150} or P_{91-180} is one of the best alternatives for selection based on 305-day milk, fat and protein yield. M_{91-180} and P_{91-180} may be chosen in stead of $M_{121-150}$ and P_{91-150} because of slightly higher heritabilities.

Efficiency of selection was defined as a ratio in percentages between correlated and direct response for 305-day yield per year. This ratio is determined by the genetic correlation, the repeatability of predicted breeding value and the generation interval. The repeatability of a sire's breeding value is determined for the most part by the heritability and number of daughters. For cows it is mainly determined by the heritability. The heritabilities of M_{91-180}, PF_{1-180} and P_{91-180} were 0.320, 0.365 and 0.351 respectively. The highest values of all other alternatives for a selection on 305-day yield were 0.326, 0.369 and 0.341 for milk, fat and protein yield respectively. Thus, as far as the repeatability was concerned, M_{91-180}, PF_{1-180} and P_{91-180} were also interesting as alternatives for selection for 305-day yield.

The efficiencies of selection based on PM_{1-180} or PP_{1-180} were 2% lower than M_{91-180} or P_{91-180}. The genetic correlation of these extended yields with 305-day yields were equal to 1 and their heritabilities were close to the values for 305-day yield. On the other hand the efficiency of selection for F_{91-180} was 2% lower than selection for PF_{91-180}. Consequently, choosing for one standard measure for practical simplicity, selection for M_{91-180}, F_{91-180} and P_{91-180} yield slightly better results than selection for PM_{1-180}, PF_{1-180} and PP_{1-180}.

The effect of heifer selection was not considered in estimating genetic parameters because of saving computer expences. It may be worthwhile to re-estimate genetic parameters for the best alternatives in order to look at the effects on genetic parameter estimates when incomplete records are included.

In conclusion selection for 305-day milk, fat and protein yield can be replaced by selection for M_{91-180}, PF_{1-180} and P_{91-180} resulting in 5%, 4% and 6% additional annual response in 305-day yield for milk, fat and protein respectively. In considering these traits in sire evaluation, potential of bias by selection in daughters has decreased and influences of days open are absent. When performing a two-trait BLUP analysis to, for milk yield, M_{31-60} and M_{91-180}, the effect of selection on sire evaluation is minimized if selection is on milk yield. In such a two-trait BLUP, breeding values for sires based on progeny records, will be predicted when some daughters have completed records for M_{31-60}. Such breeding values can serve as a preliminary selection criterium. The final selection can be carried out when enough daughters have completed records for M_{91-180}. The same procedure can be followed for selection of bull-dams. When a cow has completed M_{31-60}, cows with best indices for M_{31-60} may be assigned as potential bull-dams who are inseminated with the best bulls. A final selection can be carried out on M_{91-180}.

Acknowledgement

I am indebted to Dr. Karen Meyer who provided her EQREML programs.

REFERENCES

Agyemang, K.; E.C. Clapp and L.D Van Vleck (1985a). Extension factors for trimester yields of milk and fat. J. Dairy Sci. 68:953-957.

Agyemang, K.; E.C. Clapp and L.D. Van Vleck (1985b). Variance/covariance components associated with trimester yields of milk and fat and multiple sire evaluation for trimester yields. J. Dairy Sci. 68:1233-1240.

Auran, T. (1976a). Studies on monthly and cumulative monthly milk yield records. 3. Estimates of genetic and fenotypic parameters. Acta Agr. Scand. 26:3-9.

Auran, T. (1976b). Studies on monthly and cumulative monthly milk yield records. 4. Estimating total lactation from part lactation. Acta Agr. Scand. 26:10-17.

Broek Van Den, H. (1986). De invloed van het aantal open dagen op de 305-dagen produktie (with English summary). Report, Royal Dutch Cattle Syndicate, Arnhem.

Danell, B. (1982a). Studies on lactation yield and individual test-day-yields of Swedish dairy cows. 2. Estimates of genetic and fenotypic parameters. Acta Agr. Scand. 32:82-92.

Danell, B. (1982b). Studies on lactation yield and individual test-day yields of Swedish dairy cows. 4. Extension of part lactation records for use in sire evaluation. Acta Agr. Scand. 32:103-114.

Dommerholt, J. and J.B.M. Wilmink (1983). Fokwaardeschatting van stieren voor 100-dagen produkties van de dochters (HDN-fokwaarde). Report B-126, I.V.O. "Schoonoord", Zeist

Famula, T.R. and L.D. Van Vleck (1981). Sire evaluation by only extended partial milk and fat records. J. Dairy Sci. 64:484-490.

Fimland, E. (1983). Some properties of using the additive genetic relationship in the mixed model technique. Z. Tierz. und Züchtgsbiol. 100:361-374.

Lederer, J.A. and H. Buthmann (1982). Die Zuchtwertschätzung von Bullen und Kühen mit Hilfe des BLUP-Verfahrens nach einem Mehrmerkmalsmodell. Der Tierzüchter 8:322-325.

Meyer, K. (1985). Maximum likelihood estimation of variance components for a multivariate mixed model with equal design matrices. Biometrics 41:153-165.

Moen, R.A. (1978). A note on a special case of estimating breeding values. Acta Agric. Scand. 28:148-150.

Poutous, M. and J.C. Mocquot (1975). Etudes sur la production latière des bovins. III. Relations entre critères de production duree de lactation et intervalle entre le 1er et le 2ème vêlage. Ann. Genet. Sèl. Anim. 17:181-189.

Skjervold, H. and H.J. Langholz (1964). Factors affecting the optimum structure of A.I.-breeding in dairy cattle. Z. Tierz. und Züchtungsbiol. 80:26-40.

Tandon, P.K. and W.R. Harvey (1984). Best linear unbaised prediction of sire breeding values from part lactations of daughters. J. Dairy Sci. 67:2399-2406.

Van Vleck, L.D. and C.R. Henderson (1961a). Estimates of genetic parameters of some functions of part lactation milk records. J. Dairy Sci. 44:1073-1084.

Van Vleck, L.D. and C.R. Henderson (1961b). Use of part lactation records in sire evaluation. J. Dairy Sci. 44:1511-1518.

Wilmink, J.B.M. (1986). Comparison of different methods of predicting 305-day milk yield using means calculated from within herd lactation curves. Submitted to Livestock Prod. Sci.

Chapter 5

EFFECTS OF INCOMPLETE RECORDS ON RELATIONS AMONG CUMULATIVE YIELDS IN FIRST LACTATION AND ON EXTENSION OF PART LACTATIONS

J.B.M. Wilmink

Royal Dutch Cattle Syndicate
P.O. Box 454, 6800 AL Arnhem

Submitted to Livestock Production Science

ABSTRACT

Genetic and phenotypic correlations among cumulative milk, fat and protein yield in different periods and between extended and actual 305-day yield of 3206 first lactation records were estimated by restricted maximum likelihood, with and without including incomplete lactations i.e. lactations with fewer than 270 days in milk. Including incomplete lactations, all genetic correlations among cumulative yields through the second month, second-trimester and the first 305 days varied from 0.872 to 1.007. Heritabilities ranged from 0.254 to 0.327 for second-month, from 0.288 to 0.338 for second-trimester and from 0.300 to 0.353 for 305-day milk, fat and protein yield. Between-sire variances and heritabilities were increased for second-month yields when incomplete lactations were considered. Results suggested selection for milk yield.

Genetic correlations between extended 60-day and actual 305-day yield were 0.864, 0.911 and 0.869 for milk, fat and protein yield when incomplete lactations were considered. Phenotypic correlations were 0.840, 0.805 and 0.781 which were close to expected. Incomplete lactations revealed differences from completed lactations in level of production and rate of decline after peak yield. Extension of incomplete lactations with means and extension factors for complete lactations is justified. Selection for second-trimester yield is an interesting alternative for selection on 305-day yield.

INTRODUCTION

Predicted breeding values for dairy sires, based on 305-day yields of heifers, may be biased by selection. Fimland (1983) showed that the largest potential of bias in sire evaluation was not related to non-random mating but rather to culling of heifers.

In a previous paper (Wilmink, 1986c), efficiency of selection for cumulative yields calculated over different periods in first lactation or for extended part lactation yields was investigated.

Selection for second-trimester yield, i.e. yield from day 91 to 180 post partum, resulted in highest overall additional gain in 305- day yield, 5% for milk, 2% for fat and 6% for protein. Genetic relations of second-trimester yield with 305-day yield were equal to unity, meaning that the two yields are the same traits genetically. Selection for extended 180-day yields resulted in almost as much additional gain in 305-day yield as selection for second trimester yield. An advantage of selecting for second-trimester or extended 180-day yields is that potential of selection bias in sire evaluation has decreased.

Eriksson (1981) and Pollak et al. (1984) among others, showed that a multi-trait analysis adjusts for selection when selection was for the trait in the model or for a correlated trait and when all records were available. It is therefore expected that a two-trait analysis for second-month and second-trimester yields will minimize the potential of bias. Second-month yield instead of 60-day yield is prefered because of higher heritability and genetic correlation with 305-day yield (Wilmink, 1986c). Applying a two-trait analysis, estimates of genetic parameters for second-month and second-trimester yield, based on both complete and incomplete records should be available. To examine the effect of selection, such estimates might be compared with estimates of genetic parameters, when incomplete lactations are not considered.

On the other hand, in-progress and incomplete lactations may be extended to 305-day equivalents, allowing the inclusion of incomplete lactations in a single trait analysis of 305-day yield. Although the genetic correlations of extended part lactations, less than 180 days in milk, with 305-day yields are less than 1 (Wilmink, 1986c), a single-trait analysis on extended part lactation and 305-day yield is still of interest because of greater simplicity and less computation compared to a multi-trait analysis. However, prediction methods are developed from complete lactations. A use of these methods for extending incomplete lactations has not been validated.

The purpose of this study was 1) to estimate genetic and phenotypic relations among cumulative milk, fat and protein yield from second-month, second-trimester and 305-day lactation, with and without including incomplete lactations, and 2) to validate the prediction method, presented in Wilmink (1986b), for extending incomplete lactations.

MATERIAL

The data comprised individual test-day and cumulative monthly milk, fat and protein yields, of Black and White heifers (data set I; Wilmink, 1986c). Test-day yields were known with an equal interval of 20 days, starting at day 10 post partum. No restric-

Table 1. Means and standard deviations (sd) for the traits analyzed, based on all records (data set I) or records of young bulls (data set II)

	no. of records	% missing[b)]	milk (kg) mean	milk (kg) sd	fat ($kg*10^{-1}$) mean	fat ($kg*10^{-1}$) sd	protein ($kg*10^{-1}$) mean	protein ($kg*10^{-1}$) sd
data set I								
x_{10}[a)]	25244	0	18.8	3.8	8.09	2.07	6.27	1.30
x_{30}	25165	0	19.9	3.4	8.04	1.66	6.16	1.10
x_{50}	24986	1	19.6	3.5	7.88	1.56	6.09	1.09
x_{70}	24811	2	19.1	3.4	7.69	1.50	6.02	1.09
x_{90}	24621	2	18.5	3.4	7.49	1.45	5.94	1.10
x_{110}	24405	3	17.9	3.3	7.30	1.41	5.85	1.10
x_{130}	24166	4	17.3	3.3	7.14	1.39	5.75	1.10
x_{150}	23917	5	16.7	3.3	7.00	1.37	5.63	1.10
x_{170}	23662	6	16.1	3.3	6.87	1.37	5.51	1.11
x_{190}	23364	7	15.5	3.3	6.74	1.38	5.38	1.12
x_{210}	23090	9	15.0	3.3	6.60	1.38	5.24	1.14
x_{230}	22777	10	14.3	3.4	6.47	1.40	5.09	1.15
x_{250}	22083	13	13.6	3.4	6.31	1.43	4.92	1.17
x_{270}	20261	20	12.9	3.4	6.14	1.46	4.73	1.18
x_{290}	16017	36	12.3	3.4	6.00	1.50	4.58	1.20
data set II								
Y_{31-60}[a)]	3923	0	587	100	233	44	182	31
Y_{91-180}	3735	5	1546	279	632	112	511	90
Y_{1-305}	3206	18	5066	844	2133	349	1688	274
Y_{1-60}	3923	0	1158	196	472	93	368	62
Y_{61-305}	3206	18	3889	699	1652	283	1315	229
x_{50}	3923	0	19.6	3.4	7.72	1.49	6.02	1.06
PY_{1-60}	3923	0	5046	751	2125	301	1678	234

a) x_t = test-day yield at day t post partum
Y_{i-j} = cumulative yield from day i to j post partum
PY_{1-60} = 60-day yield extended to 305 day equivalent
b) % missing in relation to x_{10} or Y_{31-60}

tion was made on length of lactation.

Data set I consisted of records of 97 proven and 204 young sires and was used for describing the lactation curves pertaining to the different classes of length of lactation.

Data set II was formed out of data set I with records of 31 proven and 175 young sires to estimate between and within-sire variance and covariance components. Number of sires in data set II was diminished because inversion of a matrix with dimension number of sires by number of traits was required. Further, records in a herd-year-season were removed when these were of proven sires only. In other herd-year-seasons, for each proven sire one record was kept at random to save computer time.

Cumulative yields were designated as Y_{i-j}, where i and j denote days post partum, and test-day yields were designated as x_t, where t refers to number of days post partum. Finally 60-day lactations, extended to 305-day equivalents, were designated as PY_{1-60}.

An outline of the traits analysed with their means and standard deviations is in table 1.

METHODS

Estimation of variance and covariance components

Between and within sire variance and covariance components in data set II were estimated by a multivariate mixed model (1) for two traits:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_{11} & 0 \\ 0 & X_{21} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} + \begin{pmatrix} X_{12} & 0 \\ 0 & X_{22} \end{pmatrix} \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} + \begin{pmatrix} Z_1 & 0 \\ 0 & Z_2 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad (1)$$

Where $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ = a vector of observations for traits 1 and 2;

$\begin{pmatrix} \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{pmatrix}$ = a vector of herd-year-seasons and breed effects, for traits 1 and 2;

$\begin{pmatrix} \mathbf{b}_{12} \\ \mathbf{b}_{22} \end{pmatrix}$ = a vector of linear and quadratic regression coefficients for age at calving for traits 1 and 2;

$\begin{pmatrix} \mathbf{s}_{1} \\ \mathbf{s}_{2} \end{pmatrix}$ = a vector of sire contributions for traits 1 and 2;

$\begin{pmatrix} \mathbf{e}_{1} \\ \mathbf{e}_{2} \end{pmatrix}$ = a vector of residuals for traits 1 and 2.

Incidence matrices $\mathbf{X}_{11}$, $\mathbf{X}_{21}$, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ were known and matrices $\mathbf{X}_{12}$ and $\mathbf{X}_{22}$ consisted of covariables age at calving and age at calving squared. In model 1, proven sires were considered fixed because their offspring originate from selective matings. For young sires the following assumptions were made with respect to expectation and variance:

$$E \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{and var} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{I}t_{11} & \mathbf{I}t_{12} \\ \mathbf{I}t_{12} & \mathbf{I}t_{22} \end{pmatrix}$$

where $\mathbf{I}$ is the identity matrix and t_{11}, t_{12} and t_{22} are variance and covariance components between sires. The expectation and variance structure of residuals were assumed to be:

$$E \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{and var} \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{I}r_{11} & \mathbf{I}r_{12} \\ \mathbf{I}r_{12} & \mathbf{I}r_{22} \end{pmatrix}$$

where r_{11}, r_{12} and r_{22} are variance and covariance components within sires. Furthermore, cov $(\mathbf{s}, \mathbf{e}) = \mathbf{0}$.

Seasons were defined from February through August and from September through January. Breed classes were defined for 1) 100%, 2) 87.5, 75.0 and 62.5%, 3) 50 and 37.5% and 4) 25, 12.5 and 0% Dutch Friesian. Crosses were Dutch * Holstein Friesian.

To examine the effect of selection, variance and covariance components were estimated with and without incomplete lactations, that latter defined as lactations terminated before 270 days in milk. When incomplete lactations were omitted, traits 1 and 2 were known for each heifer with equal design matrices, and estimation of variance and covariance components was according to Meyer (1985) using a transformation of original traits to canonical scale. When incomplete lactations were considered, the algoritm described by Meyer (1986) was used. Each algorithm is iterative and computation of variance and covariance components was stopped when differences between two succesive rounds were less than 1%. Priors were derived from an univariate analysis for each trait and from the correlation between the sire breeding values and between residuals.

Genetic variance or covariance were calculated as 4 times the estimated between-sire variance or covariance component. Phenotypic variances or covariances were the sum of estimated between- and within-sire variance or covariance components.

Extending incomplete lactations

The validity of the prediction method (Wilmink, 1986b) for extending incomplete lactations was examined two ways: 1) a comparison of the rate of decline of lactation curves for incomplete and complete lactations, and 2) a comparison of expected and estimated genetic and phenotypic correlations between extended 60-day and actual 305-day yields.

First generalized least square (GLS) means were estimated for all test-day yields of lactations of different lengths in data set I with an univariate mixed model. By fitting functions to these GLS means, it can be judged whether different means and prediction factors should be used to extend incomplete lactations and lactations in progress. At each stage of lactation i.e., at day 10, 30.... post partum, test-day yields were analysed by (2):

$$\mathbf{Y} = \mathbf{X}_1\mathbf{b}_1 + \mathbf{X}_2\mathbf{b}_2 + \mathbf{X}_3\mathbf{b}_3 + \mathbf{Zs} + \mathbf{e} \qquad (2)$$

where $\mathbf{Y}$ = a vector of test-day yields at day t post partum;

$\mathbf{b}_3$ = a vector of fixed effects for classes of lactation length;

$\mathbf{X}_3$ = a known incidence matrix.

Other elements in (2) are as in (1); in model 2, proven sires were fixed. Young sires were considered random with:

$$E\begin{pmatrix}\mathbf{s}\\ \mathbf{e}\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix} \quad \text{and var} \begin{pmatrix}\mathbf{s}\\ \mathbf{e}\end{pmatrix} = \begin{pmatrix}\mathbf{I}t_i & 0\\ 0 & \mathbf{I}r_i\end{pmatrix}$$

Variance components t_i and r_i were estimated by restricted maximum likelihood using Fisher's method of scoring (Searle, 1979). After solutions were converged, the mixed model equations were solved for $\mathbf{b}_3$. GLS means were estimated from b_3 and from the subclass mean for the lactation length class with the highest number of observations.

Nine classes of lactation length were defined: 51-90, 91-130, 131-170, 171-210, 211-250, 251-290, 291-330, 331-370 and more than 370 days in milk.

The following function was fit to the GLS means per lactation lenght class:

$$f(t) = a + bt + ce^{-0.05t} + de^{0.04(t-305)} \qquad (3)$$

where t is the number of days in milk; a, level of production; b, rate of decline after peak yield, i.e. about 50 days post partum; c, rate of incline at the beginning of lactation; and d, increased rate of decline at the end of lactation. Parameter d was zero when lactations were shorter than 250 days. Equation (3) was a modification of the function described in Wilmink (1986a) in order to fit the biological pattern of lactation curves as well as possible. Inclusion of a quadratic term dt^2 (Wilmink, 1986a) in stead of $de^{0.04(t-305)}$ yielded a fit that was as close as (3), but which had the disadvantage that parameters b and d each represented rate of decline after peak yield. The parameters b in (3) for different lactation length classes, were compared.

Second, expected genetic and phenotypic correlations between PY_{1-60} and Y_{1-305}, when incomplete lactations were included, were derived and compared with estimated values. PY_{1-60} was chosen because it was assumed to be highly affected by incomplete lactations. To derive the expected correlation, PY_{1-60} and Y_{1-305} were written as:

$$PY_{1-60} = Y_{1-60} + \overline{PY}_{61-305} + b^* (x_L - \bar{x}_L) \quad (4)$$

and

$$Y_{1-305} = Y_{1-60} + Y_{61-305} \quad (5)$$

where b^*, $\overline{PY}_{61-305}$ and $\bar{x}_L$ are prediction factor, average remaining and last test-day yield. Averages $\overline{PY}_{61-305}$ and $\bar{x}_L$ were derived from within herd lactation curves, which were estimated from the mean x_{50} and mean Y_{1-305} within herd in the previous year (Wilmink, 1986b). The expected genetic and phenotypic correlation between PY_{1-60} and Y_{1-305} was calculated from:

$$var(Y_{1-305}) = var(Y_{1-60}) + var(Y_{61-305}) + 2\, cov(Y_{1-60}, Y_{61-305}) \quad (6)$$

$$var(PY_{1-60}) = var(Y_{1-60}) + b^{*2} var(x_L) + 2b^* cov(Y_{1-60}, x_L) \quad (7)$$

$$cov(PY_{1-60}, Y_{1-305}) = var(Y_{1-60}) + b^* cov(Y_{1-60}, x_L) + b^* cov(Y_{61-305}, x_L) \quad (8)$$

In equations (6), (7) and (8) genetic and phenotypic variances and covariances were estimated by model 1, including incomplete lactations. Covariances with means were assumed to be zero. The prediction factor b^* was calculated as the phenotypic value for $cov(Y_{61-305}, x_{50})/var(x_{50})$, which was the value for extending incomplete lactations and lactations in progress. The standard error of b^* was neglected in deriving $var(PY_{1-60})$ and $cov(PY_{1-60}, Y_{1-305})$, because it should be very small (Wilmink, 1986c). Expected heritability for PY_{1-60} and genetic and phenotypic correlations between PY_{1-60} and Y_{1-305} were derived from (6), (7) and (8).

Correlations, calculated from equations (6) to (8), tell the relation between PY_{1-60} and Y_{1-305}, including incomplete lactations when the means $\overline{PY}_{61-305}$, $\bar{x}_L$ and factor b*, which depend on age and season of calving and lactation length, are correct. These correlations are compared with the correlation between PY_{1-60} and Y_{1-305}, estimated by model 1, when incomplete and lactations in progress were extended with the method in Wilmink (1986b).

RESULTS

Overall means and standard deviations are in table 1 for traits analyzed in data set I. GLS means for length * stage of lactation classes were based on these data. There was a gradual increase of missing records to 10% for test-day yields taken before 240 days post partum. Then there was another increase from 10 to 36% for test-day yields from 240 to 300 days post partum. For data set II means and standard deviations were based on records of young bulls because these determined, for the most part, estimates of variance and covariance components. Relative to the number of records for Y_{31-60} about 5% were missing for Y_{91-180} and about 18% for complete lactations. This 18% agreed well with the average percentage of incomplete records in the Dutch sire evaluation programme.

Relations among cumulative yields

Heritabilities, genetic and phenotypic relations among Y_{31-60}, Y_{91-180} and Y_{1-305} are summarized in table 2. For milk yield, the estimated between-sire variances increased by 33, 26 and 28% for Y_{31-60}, Y_{91-180} and Y_{1-305} respectively, when incomplete lactations were considered; heritabilities increased by 36, 11 and 18% respectively. For fat and protein yield, only estimates for Y_{31-60} and Y_{91-180} were affected by selection. Heritability

Table 2. Estimated between sire variance components and genetic parameters (standard error between brackets) for cumulative yields in first lactation

	σ^2_s	se[c]	genetic parameters[d] Y_{31-60}	Y_{91-180}	Y_{1-305}
milk (kg)					
Y_{31-60}[a]	514[b]		0.306	0.758	0.840
	336	91	0.224 (0.061)	0.746 (0.010)	0.825 (0.008)
Y_{91-180}	3943		0.872	0.332	0.937
	3140	822	0.863 (0.085)	0.299 (0.079)	0.932 (0.005)
Y_{1-305}	30179		0.891	0.987	0.300
	23636	6699	0.846 (0.098)	0.974 (0.047)	0.254 (0.079)
fat ($kg*10^{-1}$)					
Y_{31-60}	115		0.327	0.640	0.785
	75	20	0.232 (0.061)	0.632 (0.013)	0.774 (0.009)
Y_{91-180}	509		0.872	0.288	0.907
	542	130	0.890 (0.093)	0.336 (0.081)	0.901 (0.006)
Y_{1-305}	5689		0.892	0.997	0.353
	5586	1326	0.901 (0.090)	0.995 (0.035)	0.358 (0.086)
Protein ($kg*10^{-1}$)					
Y_{31-60}	37		0.259	0.669	0.790
	25	7	0.186 (0.055)	0.659 (0.012)	0.777 (0.009)
Y_{91-305}	385		0.887	0.338	0.915
	397	90	0.904 (0.094)	0.386 (0.089)	0.911 (0.006)
Y_{1-305}	3126		0.934	1.007	0.327
	2973	760	0.946 (0.087)	0.998 (0.029)	0.326 (0.084)

a) Y_{i-j} = cumulative yield from day i to j post partum

b) first line : incomplete lactations included (s.e. not computed)
second line: " " not included

c) s.e. = lower bound for large sampling standard error

d) on the diagonal: heritabilities
above the diagonal: phenotypic correlations
below the diagonal: genetic correlations

increased by about 40% for Y_{31-60} but decreased by 13% for Y_{91-180}; between-sire variance changed by about +50 and -5%. Genetic and phenotypic correlations changed slightly when incomplete lactations were considered, except for genetic correlation between Y_{31-60} and Y_{1-305}, which increased by 0.045. Genetic correlation between Y_{91-180} and Y_{1-305} was almost equal to 1 in each case.

Extension of incomplete lactations

GLS means for length * stage of lactation classes were fit by functions (equation (3)). Results are in table 3. Values for parameter b, representing rate of decline of yield after peak yield, were generally constant for lactations between 211 and 370 days post partum. For shorter lactations however, b decreased. Plots of functions and GLS means are presented in figure 1. Clearly, curves of lactations shorter than 200 days showed a higher rate of decline after peak yield than lactation longer than 200 days especially for fat and protein yield.

Curves of lactations longer than 250 days showed an increased rate of decline at the end, especially for protein yield (figure 1), possibly due to effects of pregnancy. Of these lactations, less than 14% were coded as culled (table 3). Increased rate of decline

Table 3. The parameters of the function[a)] fitted to lactations of different length and the percentages coded as culled

lactation length			milk (kg)				fat ($kg*10^{-3}$)				protein ($kg*10^{-3}$)			
class	days	%culled[b)]	a	b	c	d	a	b	c	d	a	b	c	d
1	51- 90	98	17.9	-0.045	-2.4		702	-1.51	-21[c)]		542	-1.10	12[c)]	
2	91-130	95	18.9	-0.041	-3.1		737	-1.26	-31[c)]		566	-0.86	-11[c)]	
3	131-170	93	18.8	-0.035	-3.0		723	-0.95	0[c)]		565	-0.67	4[c)]	
4	171-210	84	18.9	-0.034	-2.7		734	-0.92	-23		587	-0.71	-12[c)]	
5	211-250	45	20.0	-0.032	-3.2		772	-0.79	- 3[c)]		614	-0.66	-25	
6	251-290	8	21.4	-0.032	-3.5	-3.0	819	-0.80	6[c)]	-66[c)]	650	-0.63	-25	-169
7	291-330	3	21.9	-0.032	-3.8	-1.4	841	-0.83	- 3[c)]	-20[c)]	659	-0.60	-28	- 89
8	331-370	6	21.9	-0.031	-3.9	-0.1[c)]	848	-0.83	- 5[c)]	28	655	-0.52	-25	- 48
9	371-	14	22.0	-0.029	-3.7	0.6	849	-0.75	3[c)]	22	654	-0.46	-16[c)]	- 24

a) $y = a + bt + ce^{-0.05t} + de^{0.04(t-305)}$
where t = days post partum
d = 0 for classes 1-5

b) % culled from the herd for some reasons, as given by the farmer

c) not significant (p <.05)

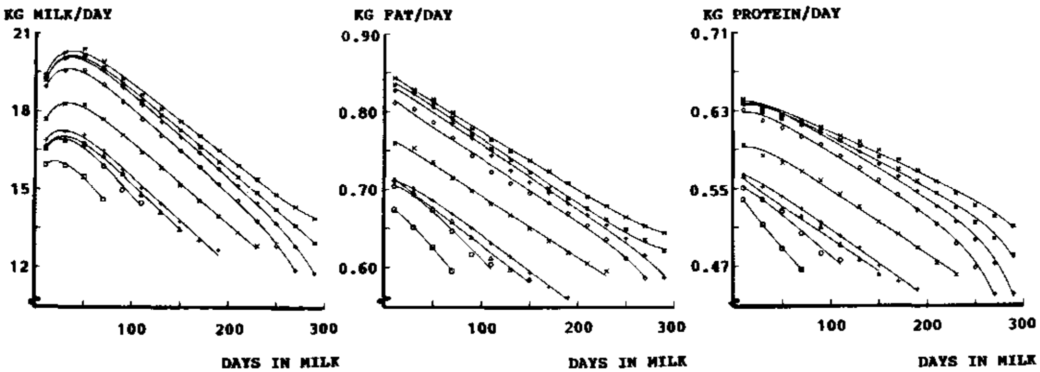

Figure 1. Estimated generalized least square means and lactation curves fit to these means. At 50 days post partum lines from below to upper refer to lactation length class 1 to 9 respectively

at the end (parameter d, table 3), was highest for lactations between 251 and 290 days, and decreased for longer lactations. About 45% of lactations between 211-250 days were coded as culled. GLS means for this class appeared to have a constant rate of decline after peak yield (figure 1).

Estimates for between-sire variance components, heritabilities and correlations between Y_{1-305} and PY_{1-60} are in table 4. When 60-day yield were extended to 305-day yield, for complete and incomplete lactations, estimates for between-sire variances and for heritabilities of PY_{1-60} milk, fat and protein, and Y_{1-305} milk increased substantially. The genetic correlation between PY_{1-60} and Y_{1-305} increased from 0.80 to 0.86 for milk, whereas there were small differences for fat and protein yield.

Expected values for between-sire variance components, heritabilities and genetic and phenotypic correlations between Y_{1-305} and PY_{1-60}, for milk and fat agreed with estimates obtained from analysis of Y_{1-305} and PY_{1-60} when incomplete records were considered.
For protein expected genetic correlation was 0.065 higher than estimated genetic correlation, but expected heritability was 0.042

Table 4. Estimates for between sire variance component, heritability and genetic and phenotypic relations between 305-day yield (Y_{1-305}) and extended 60-day yield (PY_{1-60})

	Between sire component		heritability		$r(Y_{1-305}, PY_{1-60})$	
	Y_{1-305}	PY_{1-60}	Y_{1-305}	PY_{1-60}	genetic	phenotypic
milk (kg)						
I[a]	23664	12115	0.254	0.219	0.804	0.828
II	30662	18103	0.309	0.300	0.864	0.840
III	--	19557	--	0.276	0.879	0.839
fat ($kg*10^{-1}$)						
I	5582	1778	0.358	0.223	0.926	0.794
II	5784	2502	0.356	0.298	0.911	0.805
III	--	2952	--	0.305	0.892	0.788
protein ($kg*10^{-1}$)						
I	2975	766	0.326	0.179	0.882	0.773
II	3110	1242	0.329	0.267	0.864	0.781
III	--	1310	--	0.225	0.934	0.791

a) I = estimates, not including incomplete lactations
II = estimates, including incomplete lactations
III = expected estimates, including incomplete lactations

lower than estimated heritability. The estimates of phenotypic correlations for protein were in agreement with expected phenotypic correlations.

DISCUSSION AND CONCLUSIONS

The purpose of using a two-variate analysis to estimate the variance and covariance components was to account for bias due to selection on dairy performance (Rothschild et al 1979; Meyer and Thompson 1984). Using all records estimates of sire variances increased for milk, fat and protein Y_{31-60}, milk Y_{91-180} and milk Y_{1-305}. Results suggest that decisions about culling were most directly related to milk yield.

Heritabilities for Y_{31-60} showed highest increase followed by Y_{91-180} including incomplete lactations. For Y_{1-305}, only estimates related to milk were changed. Estimates of genetic and phenotypic correlations differed slightly irrespective of whether incomplete lactations were considered, except for genetic correlation between Y_{31-60} and Y_{1-305} for milk, which was 0.891 including incomplete lactations and 0.846 excluding incomplete lactations. This difference, however, might be not significant because of sampling errors. Genetic parameters, considering incomplete lactations agreed with estimates presented in Wilmink (1986c). Only for Y_{31-60}, the heritabilities were higher in this study which means that selection on Y_{91-180} is still one of the best alternatives to selection on Y_{1-305}. An increased heritability for Y_{31-60} means that this cumulative yield has become more interesting as a selection objective then previously (Wilmink, 1986c).

Relative to number of records of young sires for Y_{31-60}, about 5% was missing for Y_{91-180} and 18% for Y_{1-305}. Hence, bias potential in sire evaluation has been reduced significantly, when Y_{91-180} is used instead of Y_{1-305}. The bias in sire evaluation for 305-day yield will be minimized if all records for Y_{31-60} and Y_{91-180} are considered in a multitrait analysis.

The level of production was lower for incomplete than for complete lactations. This agrees with Dommerholt (1975) and Auran (1977) who analyzed incomplete lactations for milk yield. The rate of decline after peak production was equal for lactations terminated between 211 and 370 days. Shorter lactations tended to have higher rates of decline, especially for protein yield. Results for milk yield agree with Dommerholt (1975), who found increased rate of decline for lactations shorter than 200 days, but are different from Auran (1977), who reported that rate of decline after peak yield was lower for incomplete than for complete lactations.

With regard to lactation curves, results suggest that other extension factors are required for lactations shorter than 200 days. Level of production is much more affected by lactation length than rate of decline (figure 1). In prediction equation (4), $\overline{PY}_{i-j}$ and $\bar{x}_L$ are related to level of production, whereas b* is related to differences in rate of decline. In theory, therefore, level of production should be considered first when extending incomplete lactations. It is not known, however, what the reasons are for a faster decline of lactations of less than 200 days in milk. If illness and bad management are the reasons, it is not justified to use other prediction factors for incomplete lactations. Furthermore, differences in level of production for complete lactations within herd by age by season of calving subclasses may be as large as differences in level between complete and incomplete lactations. Adjustment for difference in level from subclass averages is performed by $b^*(x_L - \bar{x}_L)$ in (4).

In addition, using means and prediction factors for incomplete lactations different from those for lactations in progress, one does not know whether a lactation becomes incomplete or will remain in progress at the time of prediction of 305-day yield. For terminated lactations only, distinction can be made. Therefore the question whether it is valid to extend incomplete lactations with procedures used for lactations in progress.
Genetic correlation between Y_{1-305} and PY_{1-60} was increased for milk but remained constant for fat and protein yields. Expected genetic and phenotypic relations between PY_{1-60} and Y_{1-305} agreed with genetic parameters estimated from Y_{1-305} and PY_{1-60}. This supports the concept that it is justified to extend incomplete lactations (Norman et al, 1985) and that the same procedure for extending lactations in progress can be used.

Heritabilities and between sire variances for milk, fat and protein PY_{1-60} were increased when incomplete lactations were extended with a procedure for lactations in progress. This means that differences between predicted breeding values will increase when incomplete lactations are extended.

Extended incomplete lactations can be used in the ordinary sire evaluation scheme for 305-day yield to account for selection bias. This has the advantage that it is simple and requires much less computer time than a multitrait analysis. In this study, all extended part lactations were 60 days and a possible extension bias, which is influenced for the most part by the means in equation (4) was not investigated. Famula and Van Vleck (1981) reported an overestimation of 305-day yield when extending incomplete lactations. Use of correct means becomes even more important when lactations of different lengths are extended. Therefore the merit of using extended incomplete lactations to account for selection bias in sire evaluation has to be compared with a multitrait analysis before it can be applied in practice.

In conclusion, genetic and phenotypic correlations changed slightly when incomplete lactations were considered. Second trimester yield is correlated, genetically, close to 1 with 305-day yield and is an interesting trait for selection. Potential of bias in sire evaluation by selection of heifers then has decreased substantially. Another reduction in selection bias might be expected by using multitrait analysis or by extending incomplete lactations. For the latter, it is justified to use the same procedure as for extending lactations in progress.

Acknowledgement

I am very grateful to Dr. Karen Meyer for supplying her REML programs.

REFERENCES

Auran, T. (1977). Studies on monthly and cumulative mothly milk yield records. V. Estimating total lactation from part lactation records from culled cows and from cows with different production levels. Acta Agric. Scand. 27:190-196.

Dommerholt, J. (1975). Correctie van de melkgift van koeien voor verschillen in leeftijd, seizoen en lactatiestadium (with English summary). Agric. Research reports 844, Centre for Agricultural. Publishing and Documentation, Wageningen.

Eriksson, J. (1982). Estimating sire's genetic value for milk yield in first and second lactation by different mixed model procedures, using selected second lactation records. Acta Agric. Scand. 32:193-206.
Famula, T.R. and L.D. Van Vleck (1981). Sire evaluation by only extended partial milk and fat records. J. Dairy Sci. 64:484-490.
Fimland, E. (1983). Some properties of using the additive genetic relationship in the mixed model technique. Z. Tierz. und Züchtungsbiol. 100:361-374.
Meyer, K. (1985). Maximum likelihood estimation of variance components for a multivariate mixed model with equal design matrices. Biometrics 41:153-165.
Meyer, K. (1986). Between algorithms: a "short cut" restricted maximum likelihood procedure to estimate variance components. J. Dairy Sci. 69:1904-1916.
Meyer, K. and R. Thompson (1984). Bias in variance and covariance component estimators due to selection on a correlated trait. Z. Tierz. und Zuchtüngsbiol. 101:33-50.
Norman, H.D.; F.N. Dickinson and J.R. Wright (1985). Merit of extending completed records of less than 305 days. J. Dairy Sci. 68:2646-2654.
Pollak, E.J.; J. Van Der Werf and R.L. Quaas (1984). Selection bias and multiple trait evaluation. J. Dairy Sci. 67:1590-1595.
Searle, S.R. (1979). Notes on variance component estimation: a detailed account of maximum likelihood and kindred methodology. Paper BU-673-M, Cornell University, Ithaca.
Rothschild, M.F.; C.R. Henderson and R.L. Quaas (1979). Effects of selection on variances and covariances of simulated first and second lactations. J. Dairy Sci. 62:996-1002.
Wilmink, J.B.M. (1986a). Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. Submitted to Livestock Prod. Sci.
Wilmink, J.B.M. (1986b). Comparison of different methods of predicting 305-day milk yield using means calculated from within herd lactation curves. Submitted to Livestock Prod. Sci.
Wilmink, J.B.M. (1986c). Efficiency of selection for different cumulative milk, fat and protein yields in first lactation. Submitted to Livestock Prod. Sci.

Chapter 6

GENERAL DISCUSSION

Introduction

Individual cow milk recording provides the basis to control herd management and genetic improvement (Craven and Warren, 1980). After introduction of the European common market quota system, farmers' income from producing milk can be increased only by decreasing production costs. As a consequence, the need to control herd management and genetic improvement has increased.

Thus far, lactation indices are used in herd management for ranking cows on expected 305-day yield in current lactation relative to herd average (Dommerholt et al., 1977a; Craven and Warren, 1980; Bratt et al., 1986). The lactation index, reflecting cow's relative production capacity within herd, can be used to decide whether to inseminate a cow. In addition, herd indices are calculated by averaging all test-day yields, adjusted for age and season of calving and lactation stage per sampling (Dommerholt, 1977b), or by predicting the herd milk yield for the next four weeks (Craven and Warren, 1980). Herd indices can be used to control feeding and health of the dairy herd. For genetic improvement, selection of bulls and cows is based on predicted breeding values for 305-day yields (Philipsson et al., 1978; Philipsson and Danell, 1985). Cows and bulls are ranked relative to a common base, which is usually defined by population.

Recently, Van Arendonk (1985) developed new management guides, representing future profitability of a cow, for supporting decisions on replacement of cows and on whether to inseminate a cow. The base in determining these guides is estimation of cow's relative production capacity within herd. Records from milk recording programs are not comparable within herd because of differences in age and season of calving, and lactation stage.

A base requirement in calculating guides for management and in

predicting breeding values is correct adjustment for age and season of calving and lactation stage.
In chapters 1 and 2 it was concluded that adjustment of milk, fat and protein yield for age at calving and lactation stage was improved by using mathematical functions. Advantages of using mathematical functions are that correction factors can be derived for every smallest subclass, i.e. month for age at calving, and day for lactation length, and that level of production can be updated if it has changed in time. Adjustment factors for test-day fat and protein yield are presented in chapter 2. This allows a calculation of a herd index for fat plus protein yields. Such an index reflects the amount of energy produced per sampling better than a herd index for milk yield.

In chapter 3, prediction of 305-day milk yield is investigated, which is useful for herd management and early selection of bulls and cows. Early selection may increase genetic progress in 305-day yield and reduces potential of bias by selection. These aspects are studied in chapters 4 and 5.

In this chapter, the general discussion of this thesis will be confined to 1) prediction of future production, 2) early selection of bulls and cows, 3) relation of second trimester yield with persistency and 4) definition of the breeding goal.

Prediction of future production

In chapter 3, prediction of future production was in two parts: estimation of averages within class of systematic environmental effects and use of prediction factors, derived from regression of remaining lactation yield on known test-day yields. It was shown that use of last known test-day only was as accurate as use of all known test-day yields in current lactation. Prediction factors were found to be insensitive to differences in herd by age by season of calving subclasses whereas within subclass averages were sensitive to these differences. Within subclass averages were obtained by back-adjusting within herd lactation curves for age and season of calving. Within herd lactation curves were estimated from simple herd averages in the previous year for test-day yield at day 50 post partum and 305-day yield.

The advantage of this procedure is, when used with correct adjustment factors for test-day yields, that bias in prediction is restricted to differences in herd level of production from year to year. In chapter 3, prediction bias was 76 kg when extending 60-day milk yield of heifers. Using these extended 60-day yields in sire evaluation, bias in predicting breeding values might be 152 kg. Further research should be directed to development of a procedure for better estimation of herd averages of test-day yield at day 50 post partum and 305-day yield and for accounting for herd-year-season interactions.

In studying different methods to predict 305-day yields, milk yield was the only trait considered. The method, recommended in chapter 3, was applied to fat and protein yield in chapter 4 and 5. Phenotypic correlations between extended part lactation fat or protein yields and their 305-day equivalents were only slightly lower than for milk yield (chapter 4, tables 2-4). Consequently, fat and protein yield is predicted as accurate as milk yield when using the method presented in chapter 3.

Use of regression methods in predicting 305-day milk yield, results in generally smaller prediction error than use of ratio methods or exponential functions (Danell, 1982c). The dependency of the regression method on correct within subclass averages for last known test-day and for remaining lactation yield was seen as a serious drawback. However, in the ratio method, in which remaining part lactation yield is estimated by multiplying last known test-day yield with a ratio factor averages are also involved in estimating ratio factors (Van Arendonk, 1981), and such ratio factors are dependent on effects such as level of production, and age and season of calving (Wiggans and Van Vleck, 1979; Danell, 1982c).

By estimating the average lactation curve, pertaining to a group of cows future test-day yields or part lactation yield in an arbitrary period can be predicted unbiasedly (Wilmink, 1985). Prediction of future test-day yields is valuable in feeding advisory programs (Herland, 1980). Furthermore, part lactation yields in next lactation can be predicted when age and season of calving for next parturition are known. Averages are taken from the average lactation curve for that herd by age by season of calving subclass. Prediction factors may be derived from phenotypic correlations between test-day yields in successive lactations and phenotypic standard deviations of test-day yields. Predicted yields in next lactations may be used to predict the amount of milk within herds over an arbitrary period, which is useful to the farmer in the present European common market milk quota system.

Early selection of bulls and cows

In chapter 4, it was concluded that the highest increase in annual genetic gain for 305-day yield was obtained when selection was for second trimester milk and protein yield and extended 180-day fat yield. Choosing one measure for milk, fat and protein, selection based on extended 180-day yields resulted in slightly lower expected average correlated response in 305-day yield than selection based on second trimester yields. Considering single trait selection and using genetic parameters and selection intensities in chapter 4, annual genetic gain amounted to 77.6 kg milk, 3.64 kg fat and 2.54 kg protein cow^{-1} year^{-1} if selection was based on 305-day yield. Single trait selection based on second trimester yield resulted in an additional 3.9 kg milk, 0.07 kg fat and 0.15 kg protein cow^{-1} year^{-1}. Corresponding figures were 2.3 kg milk, 0.15 kg fat and 0.10 kg protein cow^{-1} year^{-1} when selection was based on extended 180-day yields. In The Netherlands selection on net-profit index is presently used. Net-profit index is a combination of predicted breeding values for milk, fat and protein, weighed by -0.125, 4.75 and 7.5 respectively (Wilmink and De Graaf, 1986). Selection based on net-profit index for second trimester yields resulted in an additional 2.9 kg milk, 0.05 kg fat and 0.13 kg protein cow^{-1} year^{-1}.

Klir et al. (1983) reported expected responses per generation

(selection intensity = 1) to sire selection of 738 pounds milk and 30.1 pounds fat, when selection was based on 270-day yield.
Using the sum of generation intervals in chapter 4, genetic progress was 34.8 pounds milk and 1.4 pounds fat cow^{-1} $year^{-1}$. When selection was for second trimester yield and account was made for shortening generation interval, expected correlated response was 36.0 pounds milk and 1.4 pounds fat cow^{-1} $year^{-1}$ in the first 270 days of lactation. These results indicate that selection for second trimester or extended 180-day yield is attractive.

Presently, sires' breeding values for 305-day yields in first lactation are predicted by single trait Best Linear Unbiased Prediction (BLUP) methods (Henderson, 1973). In chapter 5, a two-trait BLUP analysis for second month and second trimester yield (MT_{180}) or a single-trait BLUP for extended and complete lactations (ST_{EXT}) was recommended to select early and minimize bias by selection of heifers.

To judge whether to use MT_{180} or ST_{EXT} in practice, breeding values for second trimester yields and extended 180-day yields, including incomplete lactations of less than 180 days in milk, were estimated by MT_{180} and ST_{EXT}, respectively, from data used in chapter 4. No restriction was made on lactation length and sires were required to have at least 2 daughters; this resulted in 29796 cows sired by 1094 bulls. Incomplete and 180-day lactations were extended using the prediction procedure in chapter 3. The model used was described in chapter 4 (model 1) for single-trait and in chapter 5 (model 1) for two-trait analysis. MT_{180} and ST_{EXT} were compared with MT_{305}, a two-trait analysis for second month and 305-day yield. 305-day breeding values from MT_{305} were chosen as reference because in MT_{305} sires were evaluated for 305-day yield, which is the present situation, and adjustment was made for selection of heifers. In addition, second trimester (ST_{180}) and 305-day (ST_{305}) yields were analyzed by single-trait BLUP methods.

Rank correlations between sires evaluated by the different alternatives and having more than 10 effective daughters (= number of daughters after absorption of fixed effects) in ST_{305} were at

Table 1. Rank correlations between sires, evaluated for milk, fat and protein yield in first lactation

	MT_{180}	MT_{305}	ST_{180}	ST_{EXT}	ST_{305}
milk					
MT_{180}[a]	1.000				
MT_{305}	0.976	1.000			
ST_{180}	0.989	0.960	1.000		
ST_{EXT}	0.981	0.982	0.968	1.000	
ST_{305}	0.943	0.965	0.953	0.948	1.000
fat					
MT_{180}	1.000				
MT_{305}	0.985	1.000			
ST_{180}	0.988	0.970	1.000		
ST_{EXT}	0.985	0.984	0.972	1.000	
ST_{305}	0.961	0.979	0.965	0.960	1.000
protein					
MT_{180}	1.000				
MT_{305}	0.985	1.000			
ST_{180}	0.991	0.972	1.000		
ST_{EXT}	0.988	0.984	0.976	1.000	
ST_{305}	0.963	0.975	0.967	0.957	1.000

[a] MT_{180} = multi trait analysis for second trimester yield
MT_{305} = " " " " 305-day yield
ST_{180} = single trait " " second trimester yield
ST_{EXT} = " " " " extended 180-day lactations and incomplete lactations
ST_{305} = " " " " 305-day yield
Number of sires = 241

least 0.943 for milk, 0.960 for fat and 0.957 for protein yield (table 1). Rank correlations with MT_{305} were at least 0.960 for milk, 0.970 for fat and 0.972 for protein yield.

In table 2 the distribution of sires over deciles is given, when selecting the top 30% on MT_{305}. If the best 10% bulls (n = 24) are chosen on MT_{305}, 1 to 5 bulls would be in the second 10% class if selection is on MT_{180}, ST_{180}, ST_{EXT} or ST_{305}.

Table 2. Distribution of estimated breeding values for milk, fat and protein yield: deciles of MT_{180}, ST_{180}, ST_{EXT} or ST_{305} (vertical) by deciles for MT_{305} (horizontal), selecting the top 30% for MT_{305}[a)]

	MT_{180} by MT_{305}			ST_{180} by MT_{305}			ST_{EXT} by MT_{305}			ST_{305} by MT_{305}		
	1	2	3	1	2	3	1	2	3	1	2	3
milk												
1	21[b)]	3		20	4		21	3		20	4	
2	3	17	4	4	14	6	3	16	5	3	16	7
3		4	16		6	15		5	16	1	6	15
4			4			3			3			2
fat												
1	21	3		21	3		23	1		22	2	
2	3	18	3	3	17	4	1	20	3	2	19	3
3		3	17		3	16		3	16		3	17
4			4		1	4			5			4
protein												
1	20	4		19	5		20	4		21	3	
2	4	18	2	5	16	3	4	17	3	3	17	4
3		2	19		3	15		3	19		4	16
4			3			6			2			4

a) see table 1

b) all entries are number of sires (total number = 241)

Rankings of sires on MT_{180} and ST_{EXT} were more highly correlated with rankings on MT_{305}, probably due to better adjustment for heifer selection, than with rankings on ST_{180} and ST_{305}. Therefore, it is better to select on MT_{180} or ST_{EXT} than on ST_{305}. In general, sire rankings on second trimester, extended 180-day and 305-day yields were close but, due to a shorter generation interval, selection on MT_{180}, ST_{EXT} or ST_{180} has to be preferred.

Rank correlations between MT_{180} and ST_{180} were at least 0.988. Both MT_{180} and ST_{180} can be used in evaluating sires for second trimester yield. The advantage of using MT_{180} instead of ST_{180} is that 1) a priliminary selection can be carried out when daughters of young bulls have only second month yields and 2) potential of bias, due to selection of heifers, has been minimized. A disadvantage, however, is that much more computer power is required. In ST_{EXT} young bulls can be selected preliminary when daughters have extended records (less than 180 days in milk). A final selection can be carried out when enough daughters have extended 180-day yields. Problems with ST_{EXT}, however, may arise. For example, there is a problem when young sires having only extended 60-day records are compared with sires having mostly extended 180-day records because of prediction bias (chapter 3) and because extended 60-day and extended 180-day records are different traits (chapter 4).

Effects of pregnancy on milk production starts 4-5 months after conception (Danell, 1982a). Therefore second trimester or extended 180-day yields are not influenced by number of days between parturition and conception, which is a valuable phenomenon in cow evaluation. In cow evaluation, second trimester yields seem more attactive than extended 180-day yields because of a possible higher prediction bias for excellent cows.

Relations of second trimester yields with persistency

It has been argued that persistency of production will decrease when selecting for cumulative yields early in lactation (Van Vleck, 1975, cit. by Danell, 1982c). The benefits of persistent milk production are reported by Danell (1982b). Using materials and methods in chapter 4, genetic and phenotypic correlations were estimated between two measures of persistency and second trimester or 305-day yield. Persistency was defined as the regression coefficient (b), fit to test-day yields after 50 days post partum (Wilmink, 1980), usually negative in value, and as the ratio (P) of cumulative yields from 180 to 270 days and from 0 to 90 days post partum in percentages (Danell, 1982b). Genetic correlations between b and second trimester yield were -0.237, -0.266 and 0.003 for milk, fat and protein (table 3). Corresponding correlations between b and 305-day yield were -0.169, -0.223 and -0.004. P was correlated positively with second trimester and with 305-day yield. Genetic correlations were about 0.20-0.25 for milk, 0.08

Table 3. Genetic parameters for second trimester yield (Y_{91-180}), 305-day yield (Y_{1-305}), rate of decline of the lactation curve after 50 days post partum (b) and ratio of third trimester and first trimester yield (P)

	Estimated variances for sires		genetic parameters[a)]			
	between	within	Y_{91-180}	Y_{1-305}	b	P
milk (kg)						
Y_{91-180}	3451	39042	0.325	*	-0.123	0.146
Y_{1-305}	29456	348222	*[b)]	0.312	-0.088	0.180
b	$10.52*10^{-6}$	$177.9*10^{-6}$	-0.237	-0.169	0.223	0.838
P	7.18	87.0	0.209	0.250	0.886	0.305
fat ($kg*10^{-1}$)						
Y_{91-180}	503	6000	0.309	*	-0.086	0.052
Y_{1-305}	5938	58321	*	0.370	-0.097	0.006
b	$17.86*10^{-6}$	$414.4*10^{-6}$	-0.266	-0.223	0.165	0.824
P	6.72	131.1	0.074	0.084	0.939	0.195
protein ($kg*10^{-1}$)						
Y_{91-180}	363	3742	0.354	*	0.042	0.172
Y_{1-305}	3066	33551	*	0.335	0.075	0.192
b	$3.06*10^{-6}$	$207.1*10^{-6}$	0.003	-0.004	0.058	0.839
P	3.76	111.5	0.325	0.320	0.942	0.130

a) on diagonal: heritabilities, range SE: 0.021-0.072
above diagonal: phenotypic correlations, range SE: 0.003-0.015
below diagonal: genetic correlations, range SE: 0.028-0.206
b) not estimated

for fat and 0.32 for protein yield. Heritablities for P were 0.305, 0.195 and 0.130 for milk, fat and protein, whereas corresponding heritabilities for b were 0.03-0.08 lower in value.

An important advantage of persistent production is a better distribution of feed requirements over the entire lactation. Because feed requirements are determined by the amount of fat, protein and lactose in milk (Van der Honing, 1984) it is more interesting to look at the genetic correlations of b or P with second trimester and 305-day yield for fat and protein. In that case, there was almost no difference in genetic correlations between b and second trimester or 305-day yield and no difference in genetic correlations between P and second trimester or 305-day yield (table 3). Correlated responses for b will be negative for fat and zero for protein if selection is for second trimester or 305-day yield. As a result, rate of decline after peak yield will increase. Defining persistency as P, correlated responses are expected to be almost zero for fat and positive for protein. Positive relations of P with yield can be explained because P will increase in value if level of lactation curve has increased but b remains constant (Danell, 1982b).

In conclusion, there is no difference in direction of expected correlated responses in persistency when selection is on second trimester or 305-day yield.

Definition of breeding goal

Breeding goal was defined as 305-day yield in first lactation (g_1). Because second trimester, extended 180-day and 305-day yield can be considered as genetically identical traits, it is not expected that genetic relations of second trimester and extended 180-day yields with later lactation yields will deviate from genetic relations between 305-day and later lactation yields. Therefore it seems reasonable that use of life-time production as a breeding goal and use of first lactation records as observations has no impact on choice of alternatives for 305-day yield during first lactation as reported in chapter 4.

When studying use of later lactation records in sire evaluation, breeding goal has to be clearly expressed. In a review, Strandberg (1985) concluded that life-time profit seemed to be the best breeding goal. When genetic correlation between first and second lactation was below 0.90 first and second lactation records could be considered in sire evaluation only for selection of sires to breed sons. Consideration of lactations later than first and second seemed not to be worthwhile.

An evaluation of first and second lactation records, which has to be done simultaneously because of high culling rate after completing first lactation, will lead to increased generation interval anyway. To minimize this increase, cumulative yields over a certain period in second lactation, for example from day 60 to 150 post partum, which are highly correlated with second 305-day yield, might be interesting to use as alternatives for 305-day yields in second lactation.

Genetic correlations between first and second lactation milk yield were found to be higher than 0.90 in literature, when account was made for selection on first lactation yield in genetic parameter estimation (Strandberg, 1985).
Using a genetic correlation between first and second lactation records of 0.91, Meyer (1983) found that accuracy of estimated breeding values was increased by 5-6% when first and second lactation records were analyzed simultaneously rather than only first lactation records. Meyer (1983) concluded that gain in accuracy was too small to warrant delay of selection. An advantage of using later lactations was improvement of the data structure due to more links among sires; this caused about half of the increase in accuracy. Young sires were evaluated more accurately, even if they had no second lactation records. In the Netherlands, the number of links between sires was increased by more than 100% (D.S. Koorn, pers. communication) after implementation of the maternal grandsire model (Quaas et al., 1979). Consequently, gain in accuracy caused by improvement of data structure when considering second lactations, would be less in the Dutch than in the British situation as described by Meyer (1983). When parts of second lactations are considered, the increase in generation interval will be still about 1 year, if selection for first lactation yield is based on second trimester yield. As a result, the need for considering a part of second lactation in sire evaluation is limited, when selection for first lactation yield is on second trimester yield.

REFERENCES

Bratt, G; K. Pettersen and H. Funke (1986). Integration of information from A.I.-service, milk recording and veterinary diagnoses for management and health purposes. Contribution to the 37th annual EAAP meeting, Budapest.

Craven, J.A. and M.E. Warren (1980). Economic interpretation of management information derived from milk records. Contribution to the 31th annual EAAP meeting, München.

Danell, B. (1982a). Studies on lactation yield and individual test-day yields of Swedish dairy cows. 1. Environmental influence and development factors. Acta Agr. Scand. 32:65-81.

Danell, B. (1982b). Studies on lactation yield and individual test-day yields of Swedish dairy cows. 3. Persistency of milk yield and its correlation with lactation yield. Acta Agr. Scand. 32:93-101.

Danell, B. (1982c). Studies on lactation yield and individual test-day yields of Swedish dairy cows. 4. Extension of part-lactation records for use in sire evaluation. Acta Agric. Scand. 32:103-114.

Dommerholt, J.; S.R. Sijbrandij and W.M.G. Wismans (1977a). De individuele standaardkoeproduktie en lactatiewaarde. Bedrijfsontwikkeling 8:319-321.

Dommerholt, J.; S.R. Sijbrandij and W.M.G. Wismans (1977b). De bedrijfstandaardkoeproduktie. Bedrijfsontwikkeling 8:323-326.

Henderson, C.R. (1973). Sire evaluation and genetic trands. Proc. of Animal Breeding and Genetics Symp. in honor of Dr. J.L. Lush, ASAS and ADSA, Champaign.

Herland, P.J. (1980). Milk records in feeding management. Contribution to the 31th EAAP meeting, München.

Klir, J.J.; K. Agyemang and L.D. van Vleck (1983). Expected responses to selection on different parts of the lactation. Gen. research 1983-1984. Report to Eastern A.I. Cornell Univ.

Meyer, K. (1983). Scope for evaluating dairy sires using first and second lactation records. Livestock Prod. Sci. 10:531-553.

Philipsson, J. and B. Danell (1985). Survey of methods for sire evaluation of production traits practised in different countries. Contribution to the 36th annual EAAP meeting, Kallithea, Greece.

Philipsson, J.; J. Dommerholt, E. Fimland, C. Gaillard, L. Gjøl-Christensen, J. Lederer, A.E. McElintock and J.C. Moquot (1978). Problems in cow evaluation and current use of cow index. Report of a working group on cow evaluation. Livestock Prod. Sci. 5:3-18.

Quaas, R.L.; R.W. Everett and A.C. McElintock (1979). Maternal grandsire model for dairy evaluation. J. Dairy Sci. 62:1648-1654.

Strandberg, E. (1985). Estimation procedures and parameters for various traits affecting life time milk production: a review. Report 67, Swedish Univ. of Agric. Sci., Uppsala.

Van Arendonk, J. (1981). Comparison of methods to extend part lactation records. Contribution to the 32th EAAP meeting, Zagreb.

Van Arendonk, J. (1985). Studies on replacement policies in dairy cattle. Thesis, Agric. Univ. Wageningen.

Van der Honing, Y. (1984). Factors affecting energy utilisation of dairy cows. Contribution to the 35th EAAP meeting, The Hague.
Wiggans, G.R. and L.D. van Vleck (1979). Extending partial lactation milk and fat records with a function of last-sample production. J. Dairy Sci. 62:316-325.
Wilmink, J.B.M. (1980). Breeding for persistency in milk production. 2. A comparison of some persistency measures and the relationship with live weight change during lactation (in Dutch, with English summary). Student Thesis, Agric. Univ. Wageningen.
Wilmink, J.B.M. (1985). Het voorspellen van een onbekende proefmelking, deellijst en 305-dagen produktie in de lopende lactatie. Report B-276, I.V.O. "Schoonoord", Zeist.
Wilmink, J.B.M. and F.M. de Graaf (1986). Genetische parameters voor 305-dagen produkties in de eerste lactatie en de bijstelling van de netto melkgeld index (with English summary). Report Royal Dutch Cattle Syndicate, Arnhem.

SUMMARY

Data of milk recording provides the basis to control herd management and genetic improvement of cows. Different management guides can be presented to dairy farmers. Breeding values are predicted for 305-day yields in order to select bulls and cows. However, breeding values should be predicted as early as practicable, so as to increase genetic progress and minimize prediction bias by selection.

To calculate mangement guides and predict breeding values, adjustment of individual milk, fat and protein records for age and season of calving and prediction of 305-day yields should be made. For early selection, a number of information sources other than 305-day yield in first lactation, may be interesting. The research in this thesis was directed towards 1) improvement of procedures to correct milk, fat and protein records 2) prediction of 305-day yield and 3) development of a procedure to predict breeding values for 305-day milk, fat and protein as early as possible.

In the first chapter, influence of age at calving on 305-day milk, fat and protein yield was studied in relation to herd and population level of production. Age differences per herd level of production were estimated by a repeatability model using 49669 305-day lactation records of pure-bred Dutch Friesians, calving from June 1979 until June 1982. Herds were grouped in 3 levels according to average 305-day milk yield. Second and third degree polynomial functions and a linear function including log of age at calving were fit to estimated age differences within and over parities for medium herd level. Best fit was a second degree polynomial for parity 1 (R^2=0.996) and for parity 2 (R^2=0.995), seperately. Age differences in third and higher parities were fit best by the linear function including log of age at calving (R^2=0.967). Using these functions, age adjustment factors were calculated. At medium herd level, age factors increased from 0.717 for heifers, aged 22-23 months at calving to 1.000 for cows of about 7.5 years old. Age factors between herd levels differed by only 0.017. It was concluded that one set of age factors could be

used for all herd levels. However, age factors differed by 0.021-0.068 from factors estimated on records of cows calving in 1970-1971 and producing about 1,000 kg less milk on heifer basis. These differences were less if level of production was updated. It was concluded that a more accurate age adjustment was obtained by using mathematical functions and updating mean production of heifers.

Age differences at various stages of lactation for test-day milk, fat and protein yield were estimated by generalized least square methods on records of 14275 Dutch Friesians (chapter 2). Mathematical functions of days in milk and age at calving (in months) were fit to age differences. Adjustment for age at calving and stage in lactation by factors calculated from these functions was compared with adjustment by factors derived from constant estimates. F-values for age at calving decreased by 0.5, 0.8 and 0.4 for milk, fat and protein, when factors derived from functions were used. Adjustment factors for month of calving were derived from constant estimates. It was concluded that use of mathematical functions in adjustment of test-day records for age at calving and lactation stage resulted in more precise correction.

In chapter 3, regression and factor analysis models for predicting 305-day milk yield, using means calculated from within herd lactation curves, were compared. In multiple regression and factor analysis all known test-day yields in current lactation were used as information sources, whereas last known test-day yield was used in single regression. Prediction of 305-day milk yield was split in 1) calculation of averages within herd-age- season subclasses and 2) estimation of predicting factors. Averages were calculated from lactation curves pertaining to a group of cows of the same age, freshening in the same month and performing the lactation in the same herd. These lactation curves were obtained by regression of the parameters in the lactation curve on mean test-day yield at 50 days post partum and mean 305-day yield per herd and by back-adjustment for age and season of calving.

Using these averages the correlation between predicted and realized 305-day milk yield for parity 1, 2 and 3 and later increased from 0.86 to 0.99 when day post partum for the last known test progressed from 50 to 210. 305-day milk yield was predicted more accurately for heifers than for older cows. In all cases average predicted minus realized 305-day yield was positive which was influenced by year influences. Correlations between predicted and realized 305-day milk yield and standard deviations of prediction errors agreed with theoretical values. No differences were found between methods. Prediction factors were insensitive to differences in age and season of calving, whereas average prediction error within low level herds differed by only 96 kg from average prediction error in high level herds when last known test was at 50 days post partum. Using the estimation procedure of means, accuracy of prediction was improved. The single regression model was recommended for prediction purposes in practice.

In chapter 6 it was concluded that 305-day fat and protein yield was predicted as accurate as 305-day milk yield, when utilizing the prediction method as presented in chapter 3.

Efficiency of selection defined as the ratio of correlated and direct response in 305-day yield on different cumulative or extended milk, fat and protein yield in first lactation, was studied (chapter 4). Genetic parameters were estimated from records of 20260 Black and White heifers, sired by 204 young and 97 proven bulls. Genetic correlations between cumulative milk, fat and protein yield in mid-lactation (120-180 days post partum) or extended 180-day yields and the 305-day equivalents were unity. Genetic correlations between second trimester and 305-day yield were 0.985 for milk, 0.998 for fat and 0.996 for protein. Heritabilities for second trimester milk and protein yield and for extended 180-day fat yield were close to values for 305-day equivalents. Heritabilities for 305-day yield were 0.309 for milk, 0.372 for fat and 0.331 for protein.
Heritability for second trimester fat yield was 0.318. It was concluded that selection on second trimester milk and protein yield and extended 180-day fat yield were the best alternatives to

select for 305-day yield, in which case annual genetic response in 305-day yield would be increased by 5, 6 and 4% respectively. Selection on second trimester fat yield resulted in 2% additional annual response in 305-day yield. Considering efficiency of selection on part lactation yields, an increase in efficiency was obtained when cumulative yields in the first months of lactation were eliminated in calculating part lactation yields.

Inclusion of incomplete first lactations to estimate genetic parameters for second month, second trimester and 305-day yield did increased heritabilities for second month and second trimester yield (chapter 5). Genetic variances were increased for second month milk, fat and protein yield, and for second trimester and 305-day milk yield, indicating that milk yield was most directly related to culling decisions. Genetic correlations between second month or second trimester yield and 305-day yields were almost not affected when incomplete first lactations were included. As a result, selection on second trimester yields would still be the best alternative to selection for 305-day yield.

When incomplete lactations were extended genetic correlation between 60-day milk yield, extended to 305 days, and 305-day milk yield was increased, but remained constant for fat and protein. Heritabilities for extended 60-day yields increased when incomplete lactations were included. It was concluded that it was justified to extend incomplete lactations with the procedures presented in chapter 3.

Alternatives for predicting breeding values for 305-day milk, fat and protein yield, together with minimization of prediction bias by selection of heifers, are discussed in the last chapter. Selection on breeding values for second trimester or extended yields would result in higher annual gain for 305-day yield than selection based on 305-day yields, whereas potential bias by selection of heifers was minimized. When predicting breeding values for extended yields, some bias by extending part lactations might be encountered.

Genetic correlations between persistency and second trimester yield were not different from genetic correlations between persistency and 305-day yield for fat and protein. It was concluded that selection for second trimester yield is an attractive alternative for selection for 305-day yield.

SAMENVATTING

In de bedrijfsvoering en veeverbetering spelen melkcontrolegegevens een centrale rol. Ter ondersteuning van de bedrijfsvoering worden na elke melkproduktiecontrole de bedrijfsstandaardkoe-produktie en de lactatiewaarde berekend. De lactatiewaarde beoogt het rangschikken van koeien op basis van de fenotypische produktiecapaciteit binnen het bedrijf, terwijl de bedrijfsstandaardkoeproduktie het gemiddeld produktieniveau van de veestapel op een zeker moment aangeeft. Daarnaast kunnen melkcontrolegegevens gebruikt worden in het berekenen van kengetallen ter ondersteuning van beslissingen over het al dan niet insemineren en opruimen van koeien. Ter ondersteuning van de veeverbetering worden, voor zowel koeien als stieren, fokwaarden geschat voor 305-dagenprodukties in de eerste lactatie.

De basis in het berekenen van kengetallen ter ondersteuning van de bedrijfsvoering en het schatten van fokwaarden voor 305-dagenprodukties is a) het corrigeren van de routinematig verzamelde melkproduktiegegevens voor verschillen in leeftijd en seizoen van afkalven en lactatiestadium en b) het voorspellen van de 305-dagenprodukties. Voorts kunnen cumulatieve produkties over willekeurige perioden in de eerste lactatie, die eerder bekend zijn dan de 305-dagenproduktie in de eerste lactatie, interessant zijn voor de vroegtijdige selectie van koeien en stieren. Het doel van dit onderzoek was 1) het verbeteren van methoden om melk-, vet- en eiwithoeveelheid te corrigeren voor leeftijd, seizoen van afkalven en lactatiestadium, 2) het verbeteren van de methode om 305-dagenprodukties te voorspellen en 3) het ontwikkelen van een procedure om de fokwaarden van stieren en koeien voor 305-dagenprodukties zo snel en nauwkeurig mogelijk te voorspellen.

Correctie voor leeftijd, seizoen en lactatiestadium

In hoofdstuk 1 is de invloed van leeftijd bij afkalven op de hoeveelheid melk, vet en eiwit gedurende 305 dagen in de lactatie onderzocht in relatie met het bedrijfs- en populatieniveau. Leeftijdseffecten werden geschat per bedrijfsniveau op basis van

49.669 lactaties van FH-koeien die afkalfden tussen juni 1979 en juni 1982. De bedrijven waren verdeeld in 3 groepen op grond van de gemiddelde 305-dagenproduktie. Het verloop van de leeftijdscontrasten in produktie werd beschreven met verschillende functies van leeftijd bij afkalven. Het bleek dat het verloop van de leeftijdscontrasten in produktie in de 1e en 2e lactatie afzonderlijk het best werd beschreven door een kwadratische functie (de procent verklaarde variantie bedroeg resp. 99,6 en 99,5). Het verloop van de leeftijdsverschillen in produktie in 3e en hogere pariteiten werd het best beschreven door een lineaire functie met leeftijd en de natuurlijke logaritme van leeftijd bij afkalven als onafhankelijke variabelen (96,7% variatie verklaard). Met behulp van deze functies en het populatieniveau zijn leeftijdscorrectiefactoren berekend per bedrijfsniveau-klasse. De leeftijdscorrectiefactoren namen toe van 0,717 voor jonge vaarzen tot 1,000 voor ongeveer 7,5 jaar oude koeien. Bij gelijke leeftijd verschilden de leeftijdscorrectiefactoren voor de hoog- of laag-niveau bedrijven maximaal 0,017 van die voor de midden-niveau bedrijven. Er kon derhalve volstaan worden met 1 set correctiefactoren voor alle bedrijven.

Leeftijdscorrectiefactoren verschilden echter 0,021-0,068 van die welke in de midden jaren '70 geschat zijn op basis van lactatiegegevens van koeien die gemiddeld 1000 kg minder melk produceerden bij een leeftijd van ca. 24 maanden. Deze verschillen werden kleiner wanneer aan de hand van de ontwikkelde methode de correctiefactoren werden aangepast voor het hogere populatieniveau. Er is geconcludeerd dat het gebruik van functies leidt tot een nauwkeuriger correctie voor leeftijd bij afkalven.

De effecten van verschillen in leeftijd op melk-, vet- en eiwitprodukties per dag op verschillende momenten in de lactatie werden geschat met behulp van gegevens van 14275 FH-koeien (hoofdstuk 2). De leeftijdsverschillen in produktie zijn beschreven met functies van leeftijd bij afkalven (in maanden) en het aantal dagen in lactatie. Hierdoor is het mogelijk voor elke combinatie van leeftijd (in maanden) en aantal dagen in lactatie correctiefactoren af te leiden voor leeftijd bij afkalven en lactatiestadium, hetgeen resulteert in een nauwkeuriger correctie. Voorts zijn correctiefactoren bepaald voor de invloed van maand van afkalven. Deze zijn berekend uit de schattingen voor het effect seizoen van afkalven.

Voorspellen van 305-dagenproduktie

In hoofdstuk 3 zijn regressie- en factoranalyse modellen voor het voorspellen van 305-dagenmelkprodukties met elkaar vergeleken. In het multipele regressie- en factoranalyse model zijn alle bekende proefmelkingen in de lopende lactatie gebruikt als informatiebronnen, terwijl in het enkelvoudige regressiemodel de laatst bekende proefmelking aangewend is. Bij het voorspellen van de 305-dagenprodukties is onderscheid gemaakt tussen het berekenen van gemiddelden binnen bedrijf-leeftijd-seizoen klassen en het schatten van de voorspellingsfactoren. Gemiddelden waren berekend uit lactatiecurven voor een groep koeien van gelijke leeftijd, die gekalfd had in hetzelfde seizoen en hun lactatie voltooide op hetzelfde bedrijf. De parameters in deze lactatiecurven waren geschat uit bedrijfsgemiddelden voor de produktie op dag 50 na afkalven en voor de cumulatieve 305-dagenprodukties in het afgelopen jaar. Dit resulteerde in lactatiecurven voor een groep vaarzen (24 maanden bij afkalven), die afkalfde in februari. Lactatiecurven voor andere groepen koeien werden verkregen door terugcorrectie voor leeftijd en seizoen van afkalven.

De correlatie tussen de voorspelde en gerealiseerde 305-dagenmelkhoeveelheid nam toe van 0,86 tot 0,99 wanneer het aantal dagen in lactatie toenam van 50 tot 210. De correlaties tussen voorspelde en gerealiseerde 305-dagenmelkgift en de standaardafwijkingen van de voorspellingsfouten kwamen goed overeen met de theoretische waarden. Er werden geen verschillen gevonden tussen de onderzochte methoden. De 305-dagenprodukties van vaarzen werden nauwkeuriger voorspeld dan die van oudere koeien. Het bleek dat de voorspellingsfactoren ongevoelig waren voor de effecten leeftijd en seizoen van afkalven. De voorspelde lijsten kwamen gemiddeld hoger uit dan de gerealiseerde lijsten, hetgeen verklaard kon worden door jaarinvloeden. Wanneer de laatste proefmelking bekend was bij 50 dagen na afkalven verschilde de gemiddelde voorspellingsfout bij laag-niveau bedrijven slechts 96 kg van de gemiddelde voorspellingsfout bij hoog-niveau bedrijven. Er is geconcludeerd dat de nauwkeurigheid van voorspelling is toegenomen wanneer de noodzakelijke gemiddelden zijn afgeleid van lactatiecurven voor een groep koeien. De enkelvoudige regressie op de laatst bekende proefmelking is aanbevolen voor routinematig gebruik.

In hoofdstuk 4 bleek dat de voorspellingsmethode in hoofdstuk 3 tevens goed voldoet bij de voorspelling van vet- en eiwithoeveelheid. De genetische correlaties tussen voorspelde en gerealiseerde 305-dagenproduktie bedroegen 0.853 voor melk-, 0,930 voor vet- en 0,894 voor eiwithoeveelheid, wanneer de lactatielengte gelijk was aan 60 dagen. De fenotypische correlaties waren achtereenvolgens 0,832, 0,800 en 0,772.

Vroegtijdige selectie van koeien en stieren

In hoofdstuk 4 is nagegaan in hoeverre een selectie op melk-, vet- en eiwithoeveelheden, berekend over verschillende intervallen in de 1e lactatie, leidt tot een hogere genetische respons per jaar in 305-dagenproduktie. De benodigde genetische parameters werden geschat op basis van gegevens van 20260 vaarzen, die afstamden van 204 proef- en 97 fokstieren. De genetische correlaties tussen de cumulatieve produkties in het midden van de lactatie (120-180 dagen na afkalven) en 305-dagenprodukties waren gelijk aan 1. Hetzelfde gold voor de genetische relatie tussen 180-dagenprodukties, voorspeld tot 305 dagen, en 305-dagenprodukties. De genetische correlaties tussen de produktie van 90 tot 180 dagen na afkalven (het tweede trimester) en de 305-dagenprodukties bedroegen 0,985 voor melk-, 0,998 voor vet- en 0,996 voor eiwithoeveelheid. De erfelijkheidsgraden voor de melk- en eiwithoeveelheid in het tweede trimester en voor de geëxtrapoleerde 180-dagenvethoeveelheid kwamen overeen met de erfelijkheidsgraden voor 305-dagenprodukties. De erfelijkheidsgraden voor 305-dagenprodukties waren gelijk aan 0,309 voor melk-, 0,372 voor vet- en 0,331 voor eiwithoeveelheid. De erfelijkheidsgraad voor de tweede-trimestervethoeveelheid bedroeg 0,318. Er is geconcludeerd, dat de erfelijke verbetering in 305-dagenproduktie het best gerealiseerd kan worden door middel van een selectie op melk- en eiwithoeveelheid in het tweede trimester en geëxtrapoleerde 180-dagenvethoeveelheid. Vergeleken met een directe selectie op 305-dagenprodukties leidt dit tot een extra genetische vooruitgang per jaar van 5% voor melk-, 4% voor vet- en 6% voor eiwithoeveelheid. Een selectie op vethoeveelheid in het tweede trimester resulteert in 2% extra erfelijke respons/jaar in 305-dagenvethoeveelheid. Voorts is geconcludeerd dat het beter is om bij een selectie op een voortschrijdend totaal de eerste proefmelkingen niet mee te nemen in de berekening. Het aantal proefmelkingen in het begin van de lactatie dat weggelaten kan worden hangt af van de lengte van de deellijst.

Het meenemen van afgebroken lijsten bij het berekenen van genetische parameters leidde tot een verhoging van de erfelijkheidsgraden voor de cumulatieve hoeveelheid in de tweede maand en tweede trimester van de lactatie (hoofdstuk 5). De additief genetische varianties namen dan toe voor de melk-, vet- en eiwithoeveelheid in de tweede maand en de melkhoeveelheid in het tweede trimester en 305 dagen. Melkhoeveelheid speelde een sterke rol bij het nemen van beslissingen over opruimen van vaarzen. De genetische correlaties tussen produkties in de tweede maand of trimester na afkalven en 305-dagenprodukties werden nauwelijks beïnvloed door het meenemen van afgebroken lijsten. Derhalve is selectie op tweede-trimesterprodukties nog steeds het beste alternatief voor een selectie op 305-dagenprodukties.

De genetische correlatie tussen de geëxtrapoleerde 60-dagen- en 305-dagenmelkhoeveelheid nam toe wanneer ook afgebroken lijsten werden geëxtrapoleerd (hoofdstuk 5). Voor vet- en eiwithoeveelheid werd geen verandering waargenomen. Voorts namen de erfelijkheidsgraden toe van geëxtrapoleerde 60-dagenlijsten wanneer ook onvolledige lijsten werden meegenomen. Geconcludeerd is, dat het gerechtvaardigd is om afgebroken lijsten te extrapoleren met de methode in hoofdstuk 3.

Tenslotte zijn in het laatste hoofdstuk alternatieven voor het schatten van fokwaarden voor 305-dagenprodukties bediscussieerd. Een selectie op fokwaarden voor tweede-trimesterprodukties of geëxtrapoleerde 180-dagenlijsten resulteert in een hogere erfelijke vooruitgang per jaar voor 305-dagenprodukties. Het effect van selectie van vaarzen op de geschatte fokwaarden is dan aanzienlijk gereduceerd. Echter, fokwaarden voor geëxtrapoleerde 180-dagenlijsten kunnen onzuiver geschat zijn omdat de gemiddelde fout van het extrapoleren van lijsten kan verschillen tussen stieren. Voorts bleek dat de genetische relatie tussen persistentie en vet- of eiwithoeveelheid in het tweede trimester niet verschilde van de genetische relatie tussen persistentie en vet- of eiwithoeveelheid in 305 dagen. Derhalve is het schatten van fokwaarden voor produkties in het tweede trimester van de lactatie een zeer aantrekkelijk alternatief voor het schatten van fokwaarden voor 305-dagenprodukties.

CURRICULUM VITAE

Hans Wilmink werd geboren op 13 augustus 1952 te Enter, waar hij op een veehouderijbedrijf opgroeide. In 1970 behaalde hij het diploma Gymnasium B aan het toenmalige Gymnasium "St. Alberti" te Zenderen. Na twee jaar studeren aan de Technische Hogeschool Twente en enige tijd werken op het ouderlijk bedrijf, begon hij in 1974 met de studie aan de Landbouwhogeschool. In 1980 slaagde hij voor het doctoraal-examen met veefokkerij als verzwaard hoofdvak en veevoeding en de wiskundige statistiek als keuzevakken. Na ruim een jaar lesgeven aan de Rijks Middelbare Landbouwschool in Tiel werd hij begin 1982 door de toenmalige landelijke rundveeverbeteringsorganisaties als wetenschappelijk medewerker gedetacheerd bij het Instituut voor Veeteeltkundig Onderzoek in Zeist, speciaal belast met foktechnisch onderzoek. Sinds 1985 is de auteur teamleider van de afdeling onderzoek van het Nederlands Rundvee Syndicaat te Arnhem.