Multivariate Covariance Functions for Test Day Production in Danish Dairy Breeds

Jette H. Jakobsen[1], Per Madsen[2], and Jørn Pedersen[1]

[1] *Danish Agricultural Advisory Centre, Udkaersvej 15, Skejby, DK-8200 Aarhus N, Denmark,*
[2] *Danish Institute of Agricultural Sciences, Department of Animal Breeding and Genetics,*
PO Box 50, DK-8830 Tjele, Denmark

Introduction

Several countries are now using test day models for genetic evaluation of milk production traits. More countries are planning to change from a lactations model to a test day model. It has been decided also to introduce a test day model for breeding evaluation of milk production traits in Denmark.

Four breeds are represented within the Danish dairy cattle population. These are Red Danes, Danish Holsteins, Danish Jerseys, and Danish Red Holsteins. The Danish Red Holsteins account for one percent of the purebred Danish dairy cattle whereas Red Danes, Danish Holsteins and Danish Jerseys account for approximately 10, 76, and 13 percent, respectively. Canadian studies have shown that covariance parameters are different across breeds (Schaeffer *et al.,* 2000).

Disadvantages of test day models account size of mixed model equation system due to the huge number of parameters to be estimated.

The purpose of this study was to estimate (co)variance components to be used in the Danish test day model for yield for the Danish dairy breeds and to compare fit of covariance functions within and across lactations within traits and for all traits jointly. Also, heritabilities and genetic correlations within and across traits and across breeds were compared.

Material and Methods

Data

To represent the three major Danish dairy breeds four data set were formed – one data set for the Red Danes (RD), two separate data sets representing two different regions in Denmark for Danish Holsteins (DH), and one data set for the Danish Jerseys (DJ). Each of the four data sets comprised 40,000 – 50,000 cows with data. Test day records were from the time period 1990 to 2000, and records from completed as well as uncompleted lactations. Pedigrees were traced as far back as possible and organized such that both sire models and animal models could be run. The relationship in the sire model took both male and female relationship into account. All records were obtained from the national cattle database at the Danish Agricultural Advisory Centre.

Models

The indirect covariance function approach (Van der Werf *et al.,* 1998) was used to estimate parameters. Five lactation stages (DIM = 5-20, 31-60, 121-150, 211-240, 301-330) were formed within each lactation. For observations obtained in one of the first three lactation stages in first lactation the following effects were included in the model: Herd*Test Day (*r*), Herd*Time Period (*fc*), Calving Year*Calving Month (*fc*), Calving Age*Time period (*fc*), breed proportions (*fr*), heterozygosity (*fr*), animal (*r*), and residual (*r*) where *r* is random effect, *fc* is fixed class effect, and *fr* is fixed regression. For observations in one of the last two lactation stages in first lactation an effect of days carried calf (*fr*) was included. This effect was formed as a deviation from the average pregnancy length at this time of lactation. The same effects were included in models for later lactations, but in addition to this, an effect of previous calving interval (*fc*) was included. Time period was defined in five-year periods. In the last 25 years, the Danish cattle populations have been strongly influenced by foreign breeds. According to Van der Werf and De Boer (1989) heritability estimates for milk production traits were biased upwards when breed proportions and heterosis were not accounted for in a cross bred population. The RD population included four breeds [original RD, American Brown Swiss (ABS), Red Canadian Cattle (RCC), and Nordic Breeds (NB)]. The DH population included two breeds [original

DH and Holstein]. And the DJ population included three breeds [original DJ, New Zealand Jerseys and American Jerseys]. Proportions of original Danish genes were removed from the model as a constraint. For DJ and DH an effect of general heterozygosity was included in the model. General heterozygosity was defined as a summation of all heterosis contributions within each of these breeds. For RD, five heterosis combinations were included in the model. These were RDxABS, RDxRCC, RDxNB, ABSxRHF, and sum of other heterosis combinations in the RD breed.

Estimation of Variance Components

Five lactation stages, three lactations, and three production traits formed 45 different "traits". The diagonals of the 45 x 45 matrices were filled with variances obtained from bivariate animal model analyses, where the second lactation stage of first lactation was kept as reference trait for all analyses to account for selection. Correlations between the 45 "traits" were obtained in sire model analyses. All variance components were estimated using the DMU-package (Madsen & Jensen, 2000). To sum part-matrices and to obtain positive definite genetic and residual (co)variance matrices the estimated submatrices were weighed together using iterative summing of expanded part matrices (Mäntysaari, 1999). Due to numerical problems the procedure was modified to sum on correlation matrices instead of covariance matrices. Additive genetic and residual covariances were obtained from variances from animal models and correlations from sire models.

Fit of Covariance Functions

Different orders of normalized Legendre polynomials (Abramowitz & Stegum, 1965) were fitted within and across lactations for milk, fat, and protein, respectively. And in addition to this, covariance functions were fitted within and across lactations for milk+fat+protein. With other words, covariance functions were fitted in two dimensions. Multivariate fit of covariance functions have earlier been published by Veerkamp and Goddard (1998), but with a slightly different approach. Order of fit within lactation was limited to fourth order polynomials (five lactation stages), and order of fit across lactations was limited to second order polynomials (three lactations).

Table 1. Scenarios and fit of covariance functions.

Scenario	Fit of covariance functions
1	Full fit
2	Full fit –1 within lactation
3	Full fit within lactation and full fit across lactations
4	Full fit –1 within lactation and full fit across lactation
5	Full fit –2 within lactation and full fit across lactation
6	Full fit –3 within lactation and full fit across lactation
7	Full fit within lactation and full fit across lactations + interaction of linear terms
8	Full fit within lactation and full fit across lactations + interaction of quadratic terms
9	Full fit within lactation and full fit across lactations + interaction of linear terms + interaction of quadratic terms + interaction of linear and quadratic terms[1]
10	Full fit within lactation and full fit across lactations + interaction of linear terms + interaction of quadratic terms + interaction of quadratic and linear terms[2]
11	Full fit within lactation and full fit across lactations + interaction of linear terms + interaction of quadratic terms + interaction of linear and quadratic terms[1] + interaction of quadratic and linear terms[2]

[1] Linear term from covariance function across lactations, quadratic term from covariance function within lactation
[2] Linear term from covariance function within lactation, quadratic term from covariance function across lactations

The time dependent variable within lactation (days in milk) was defined in the interval [1, ..., 365], and the time dependent variable across lactation (lactation number) was defined in the interval [1,..,3]. Eleven different fits of covariance functions were compared. In the following they are indicated as scenario 1 to 11. The scenarios are listed in Table 1. Scenario one is a full multi trait fit where a fourth order Legendre polynomial is fitted for all traits and for all lactations. Scenario 2 is also a multitrait fit, but only a third order Legendre polynomial is fitted within traits and lactation. In scenario three a fourth order Legendre polynomial is fitted within lactation and a second order Legendre polynomial is fitted across lactations. This will reduce number of parameters to be estimated from 45 (scenario 1) to 21. In the scenarios 4, 5, and 6 the five- parameter Legendre polynomial within lactation is reduced to four-, three-, and two-parameter functions, respectively. For the rest of the scenarios, a fourth order Legendre polynomial is fitted within lactation, and a second order Legendre polynomial across lactation - and in addition, interactions of linear terms within and across lactations (scenario 7), interaction of quadratic terms within and across lactations (scenario 8), interaction of linear terms + interaction of quadratic terms + interaction of linear term across lactations and quadratic term within lactations (scenario 9), interaction of linear terms + interaction of quadratic terms + interaction of linear term within lactations and quadratic term across lactations (scenario 10), and interaction of linear terms + interaction of quadratic terms + interaction of linear term within lactations and quadratic term across lactations + interaction of linear term within lactations and quadratic term across lactations (scenario 11).

Residual Matrix

The residual matrix (**R**) was separated in a permanent environmental matrix and a temporary environmental matrix using an EM-algorithm (Mäntysaari, 1999). Normalized Legendre polynomials were fitted within lactation. Temporary environment was assumed to be homogeneous within trait and lactation.

Results and Discussion

Covariance functions

Goodness of fit (R-square values) for different fit of covariance functions within and across lactations for milk, fat, and protein separately, and for milk+fat+protein are listed in Table 2 for Danish Holsteins. Results clearly show a limit of 10 for reduction of number of parameters within traits. Results demonstrate a limited advantage of fitting a covariance function across lactations. When fitting covariance functions for milk, fat, and protein simultaneously a tremendous rank reduction can be obtained maintaining a R-square value above 0.99. This result indicates that a larger reduction in number of parameters can be obtained when covariance functions are fitted simultaneously to milk, fat, and protein in favour of fitting covariance functions to each trait separately. Results for reduction of number of parameters are only shown for DH. Goodness of fits is very similar for RD and DJ indicating, that the rank can be reduced with the same number of parameters for all breeds.

Table 2. R^2 for fit of different covariance functions to genetic covariance matrices for milk, fat, protein and milk+fat+protein. Danish Holsteins.

Scenario	# of parms.	Milk	Fat	Protein	# of parms.	Milk+Fat+Protein
			R^2			R^2
1	15	1	1	1	45	1
2	12	0.9758	0.9298	0.9954	36	0.9976
3	7	0.9260	0.8226	0.8995	21	0.9927
4	6	0.9042	0.7919	0.8978	18	0.9906
5	5	0.8872	0.7813	0.8876	15	0.9889
6	4	0.4610	0.7387	0.7682	12	0.9473
7	8	0.9441	0.8869	0.9447	22	0.9945
8	9	0.9506	0.8981	0.9557	23	0.9951
9	10	0.9644	0.9141	0.9571	24	0.9965
10	10	0.9732	0.9187	0.9763	24	0.9973
11	11	0.9868	0.9350	0.9780	25	0.9987

Genetic Parameters

Daily heritabilities for milk, fat, and protein for lactation one, two, and three for RD, DH, and DJ are shown in Figure 1. Third-order Legendre polynomials were fitted for all traits and all lactations for additive genetic effects and permanent environmental effects. Daily heritabilities were lowest in beginning and end of lactation, and higher in first lactation than in later lactations. Heritability curves were similar across breeds with fat in first and second lactation, and milk in third lactation as an exception. The heritability was higher for fat yield in first lactation for RD than for DH and DJ. In second lactation, heritabilities were higher at the end of lactation for RD and DH. For third lactation milk yield, heritabilities were highest for DJ, lower for DH and lowest for RD, but heritabilities followed the same trend across breeds only the level was displaced. The levels of heritabilities as a function of time for DH are higher than estimates obtained by Strabel and Misztal (1999), but very similar to estimates obtained by Liu *et al.* (2001).

Genetic correlations between protein on day 45 and all days of fat yield in first, second, and third lactation are shown in Figure 2 for RD, DJ, and DH. In general, genetic correlations between protein and fat were higher for RD and DJ than for DH.

305-d heritabilities for milk, fat, and protein for RD, DH and DJ, respectively, are shown in Table 3. Heritabilities are similar across breeds. Heritabilities in first lactation were higher than heritabilities in later lactations.

Table 3. 305-d heritabilities for Red Danes (RD), Danish Holsteins (DH), and Danish Jerseys (DJ).

	M1	M2	M3	F1	F2	F3	P1	P2	P3
RD	0.43	0.26	0.22	0.45	0.31	0.28	0.37	0.24	0.22
DH	0.45	0.32	0.29	0.39	0.32	0.31	0.36	0.27	0.27
DJ	0.46	0.29	0.29	0.37	0.24	0.25	0.39	0.24	0.25

Genetic correlations between 305-d yields of milk and fat, milk and protein, and fat and protein, respectively, are shown in Table 4, 5, and 6. In general, genetic correlations between traits were higher for RD and DJ than for DH, and genetic correlations were higher for milk and protein and fat and protein, than for milk and fat. Lower correlations between milk and fat than between the other traits are also seen in e.g. Veerkamp and Goddard (1998), De Roos *et al.* (2001), Jakobsen *et al.* (2002). Genetic correlations between second and third lactation 305-d yield within traits were close to unity indicating possibilities for treating later lactations as repeated observations.

However, genetic correlations across breeds in the range (0.61-0.89) between persistency measures in second and third lactation indicate, that persistency in second and third lactation are different traits (Jakobsen *et al.,* unpublished data). Therefore, if persistency is going to be included in the breeding goal it must be considered to treat later lactations as different traits.

Table 4. 305-d genetic correlations between milk (M) and fat (F) for Red Danes (RD), Danish Holsteins (DH), and Danish Jerseys (DJ).

	M1-F1	M1-F2	M1-F3	M2-F1	M2-F2	M2-F3	M3-F1	M3-F2	M3-F3
RD	0.68	0.57	0.49	0.57	0.53	0.45	0.55	0.51	0.46
DH	0.49	0.23	0.09	0.44	0.41	0.32	0.44	0.45	0.37
DJ	0.76	0.47	0.50	0.66	0.55	0.54	0.69	0.56	0.56

Table 5. 305-d genetic correlations between milk (M) and protein (P) for Red Danes (RD), Danish Holsteins (DH), and Danish Jerseys (DJ).

	M1-P1	M1-P2	M1-P3	M2-P1	M2-P2	M2-P3	M3-P1	M3-P2	M3-P3
RD	0.91	0.84	0.75	0.81	0.82	0.76	0.77	0.80	0.80
DH	0.86	0.70	0.62	0.75	0.82	0.77	0.71	0.82	0.78
DJ	0.91	0.82	0.84	0.79	0.85	0.85	0.82	0.86	0.87

Table 6. 305-d genetic correlations between fat (F) and protein (P) for Red Danes (RD), Danish Holsteins (DH), and Danish Jerseys (DJ).

	F1-P1	F1-P2	F1-P3	F2-P1	F2-P2	F2-P3	F3-P1	F3-P2	F3-P3
RD	0.82	0.80	0.74	0.71	0.75	0.70	0.65	0.69	0.67
DH	0.70	0.70	0.69	0.46	0.70	0.74	0.32	0.61	0.67
DJ	0.88	0.84	0.84	0.64	0.78	0.76	0.67	0.78	0.76

Table 7. 305-d genetic correlations between milk (M), fat (F) and protein (P) for Red Danes (RD), Danish Holsteins (DH), and Danish Jerseys (DJ).

	M1-M2	M1-M3	M2-M3	F1-F2	F1-F3	F2-F3	P1-P2	P1-P3	P2-P3
RD	0.90	0.87	0.91	0.94	0.92	0.94	0.92	0.85	0.93
DH	0.85	0.80	0.97	0.81	0.75	0.94	0.85	0.80	0.96
DJ	0.88	0.89	0.98	0.86	0.87	0.95	0.90	0.92	0.97

Conclusion

These results clearly show the advantage of fitting a covariance functions multivariately. And also an advantage of fitting covariance functions for milk, fat, and protein simultaneously, in favour of fitting covariance function to the three traits, separately. In general, there was no difference across breeds in 305-d heritabilities and only slightly differences for single day heritabilities across breeds. However, genetic correlations between 305-d milk, fat, and protein yield were slightly lower for DH than for RD and DJ. Also genetic correlations between different points in time of protein and fat yield were lower for DH than for the other breeds.

References

Abramowitz, D. & Stegum, I.A. 1965. *Handbook of mathematical functions.* Dover publications, Inc., New York.

De Roos, A.P.W., Harbers, A.G.F. & De Jong, G. 2001. Random regression test-day model in the Netherlands. *Interbull Bulletin 27,* 155-159.

Jakobsen, J.H., Madsen, P., Jensen, J., Pedersen, J., Christensen, L.G. & Sorensen, D.A. 2002. Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. *J. Dairy Sci. 85,* accepted.

Liu, Z., Reinhardt, F. & Reents, R. 2001. Estimating parameters of a random regression test day model for first three lactation milk production traits using the covariance function approach. *Interbull Bulletin 27,* 159-167.

Madsen, P. & Jensen, J. 2000. *A User's Guide to DMU (6, release 4).* Danish Institute of Agricultural Sciences, Department of Animal Breeding and Genetics, Research Centre Foulum, Denmark.

Mäntysaari, E.A. 1999. Derivation of multiple trait reduced rank random regression (RR) model for the first lactation test day records of milk, protein and fat. *Proc. 50th Ann. Mtg. EAAP,* Zurich, Switzerland, Aug. 22-26, 1999.

Schaeffer, L.R., Jamrozik, J., Kistemaker, G.J. & Van Doormaal, B.J. 2000. Experience with a test-day model. *J. Dairy Sci. 83,* 1135-1144.

Strabel, T. & Misztal, I. 1999. Genetic parameters for first and second lactation milk yields of Polish black and white cattle with random regression test-day models. *J. Dairy Sci. 82,* 2805-2810.

Van der Werf, J.H.J. & De Boer, W. 1989. Estimation of genetic parameters in a cross bred population of black and white dairy cattle. *J. Dairy Sci. 72,* 2615-2623.

Van der Werf, J.H.J., Goddard, M.E. & Meyer, K. 1998. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. *J. Dairy Sci. 82,* 2231-2237.

Veerkamp, R.F. & Goddard, M.E. 1998. Covariance functions across herd production levels for test day records on milk, fat, and protein yields. *J. Dairy Sci. 81,* 1690-1701.
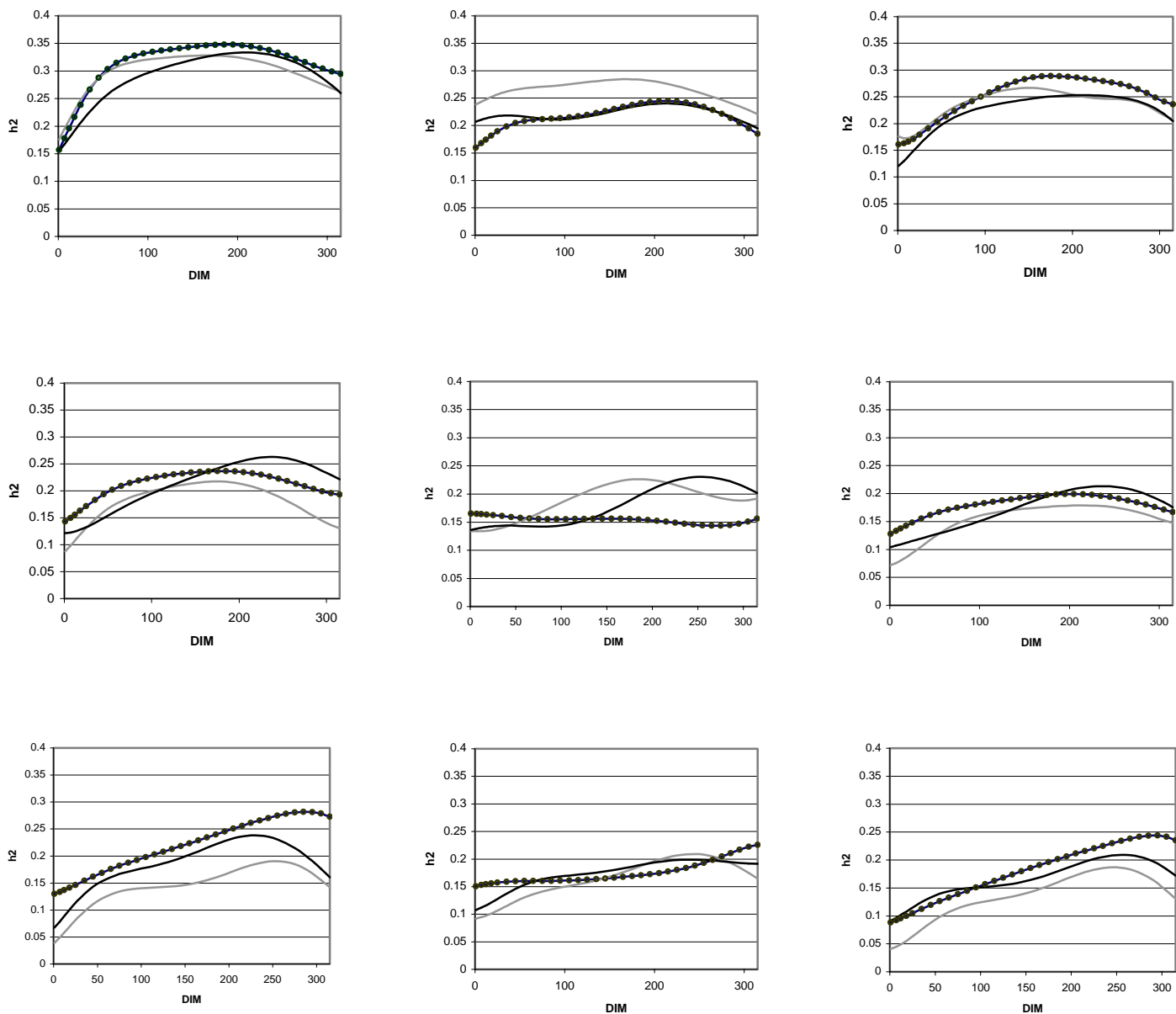
Figure 1. Heritabilities (h2) for milk, fat, and protein (horizontal) as a function of days in milk (DIM) for lactation one, two, and three (vertical) for Red Danes (grey), Danish Holsteins (black), and Danish Jerseys (circles).
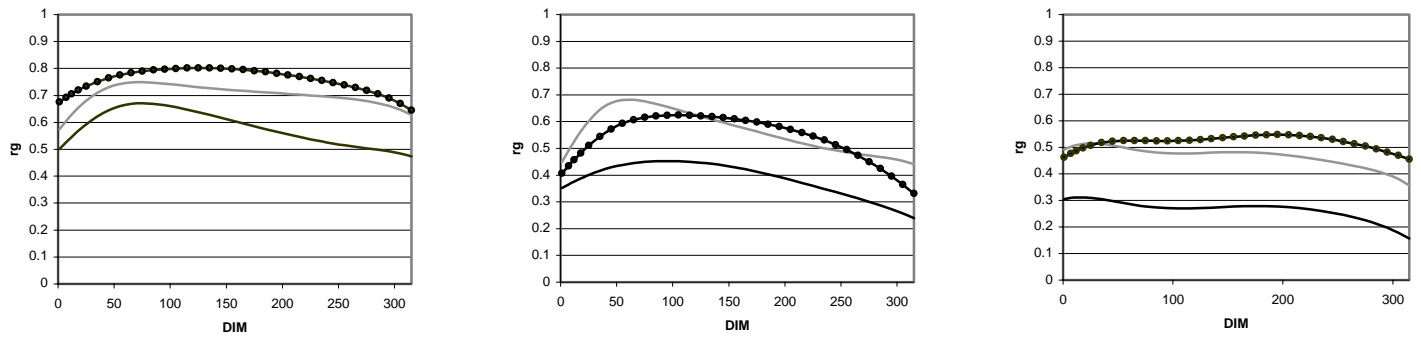
Figure 2. Genetic correlations between protein yield on d 45 in first lactation and fat yield on all days in first (left figure), second (middle figure), and third (right figure) lactation for Red Danes (grey), Danish Holsteins (black), and Danish Jerseys (circles).