GENETIC STUDIES ON TEST DAY YIELDS IN DAIRY CATTLE

Banwari Lal Pander

Ph. D.

University of Edinburgh

1992

CONTENTS

DECLARATION

I declare that this thesis has been composed by me.

Specific contributions of others are acknowledged.

B. L. Pander

August 1992

ACKNOWLEDGEMENTS

I thank Prof. W. G. Hill and Dr. R. Thompson for supervising this study. I am very grateful to them for their able guidance, invaluable advice and constant encouragement during the course of this study.

I thank the Commonwealth Commission in the U. K. for providing Commonwealth Scholarship, the Milk Marketing Board of England and Wales for providing data, Mr. G. Swanson for arranging transfer of data and helpful comments on several manuscripts.

I also thank to Dr. Geoff Simm for providing data on daily milk yield records.

Thanks to Mrs. S. Brotherstone for handling of data and computational help. I also thank to all the post-graduate and post-doctoral fellows in the group, especially Mr. B. Grundy and Dr. P. Visscher, for their help in programming.

I wish to thank:

- my parents, especially my grandmother who kindly agreed to allow me to take up this study at Edinburgh.
- my wife, Kamlesh and sons, Gaurav and Kaushal.
- friends and colleagues at home.
- Haryana Agricultural University for granting study leave.

PUBLICATIONS

Paper in press from this thesis is:

Pander, B. L., Hill, W. G. and Thompson, R. 1992. Genetic parameters of test day records of British Holstein-Friesian heifers. *Anim. Prod. (in press).*

ABSTRACT

The main aim of this study was to estimate the genetic parameters of test day records of British Holstein-Friesian heifers and to determine how best to use these parameters for genetic prediction of lactation performance from test day records. The possibility of reducing the frequency of recording to less often than monthly and procedures for the inclusion of part records in genetic evaluations were also investigated.

Estimates of genetic parameters of test day and lactation records were obtained from data on 47736 heifers in 7973 herds, progeny of 40 proven and 707 young sires, using multivariate restricted maximum likelihood methods with a sire model. Average values of heritability estimates for test day records of milk, fat and protein yields and fat and protein contents were 0.36, 0.23, 0.29, 0.36 and 0.36, respectively. Generally, heritability estimates for test day records were lowest at the start and highest in mid and late lactation. Heritability estimates for lactation records of these traits were 0.49, 0.39, 0.43, 0.63 and 0.47, respectively.

Average values of genetic correlations between adjacent TD records of these traits were high (0.92 to 0.97), and the correlations decreased as the interval between tests increased. Genetic correlations of lactation milk yield with fat and protein yields and contents were 0.72, 0.94, -0.56 and -0.53, respectively. Estimates of genetic correlation of test day records with corresponding lactation traits were also high (0.76 to 0.99), being highest in mid lactation.

The effect of increasing the interval between recordings from one month to two months was studied by analysis of alternate monthly test day records. Heritability estimates of bi-monthly test day records were similar to the average of the estimates for the two corresponding monthly test day records, as were genetic and phenotypic correlations between bi-monthly test day records and lactation records. The cost of recording could be reduced by having 6-7 equally spaced tests over the entire lactation with very little loss in accuracy.

The accuracy of prediction of breeding value for lactation performance from test day records was studied using a genetic selection index including successive test day records. It was found that the accuracy of a repeatability model including successive test day records was not far below that of an optimal index. For inclusion of part TD records in current genetic evaluation in the U. K., a method based on phenotypic indices of successive TD records would result in increased accuracy and reduced bias by inclusion of records of less than 200 days.

A repeatability model including successive TD records was recommended for the prediction of breeding values for heifer lactation records as an alternative to multivariate BLUP under the present computational facilities. Test day records in progress could easily be handled under the repeatability model.

CHAPTER 1

INTRODUCTION

The aim of dairy cattle breeding is to improve animals for production traits. Genetically superior animals are identified on the basis of their predicted breeding values from the phenotypic values for 305-day lactation yields of milk, fat and protein. The phenotypic values used for genetic evaluation are predicted (currently in the United Kingdom (U. K.) by linear interpolation) from test day (TD) records usually taken at monthly intervals over the lactation period of 305 days. Thus prediction of breeding values for 305-day lactation is a two step procedure, i.e prediction of breeding values from predicted phenotypes. Further the predicted phenotypes for 305-day yields of milk, fat and protein may be slightly biased and inaccurate. Therefore there is a need to find alternative selection criteria based on actual TD records such as an index or linear function of TD records.

Construction of a selection index would need genetic and phenotypic parameters of TD records and their association with lactation records.

The aim of the present study is to estimate the genetic parameters of TD records of milk, fat and protein yields and to apply these parameters to optimise prediction of breeding values for 305-day lactation from TD records.

Published studies on genetic and phenotypic parameters of TD records are reviewed in chapter 2. These parameters of TD records are not known, however, for British population. Estimation of genetic and

phenotypic parameters of TD records and their association with lactation records in British Holstein-Friesian heifers are discussed in chapter 3.

Presently in the U. K., only a fraction of cows (about 50%) (MMB, 1989) are included in the national genetic evaluation programme. This reduces the effectiveness of current genetic evaluation programme. One of the main limiting factors in broadening the selection base is the high cost of recording. This can be reduced by less frequent than the current monthly recording. The consequences of increasing the interval between recordings on the accuracy of genetic prediction of lactation performance are discussed in chapter 4.

The records for 305-day lactation used in this study have been predicted from TD records, therefore the phenotypic correlations between TD records and predicted lactation records may be different from those between TD records and actual lactation records (sum of daily yields). Estimation of phenotypic correlations between TD records and actual lactation records and between daily milk yield records throughout the lactation are discussed in chapter 5.

Since the current evaluation programme is a two step procedure as discussed above, a selection index or multivariate BLUP combining all the TD records would be the most accurate method for the prediction of breeding values for total lactation. Completed records of less than 200 days do not qualify for inclusion in the current genetic evaluation programme in the U. K. Neglecting records of less than 200 days and including the records of varying length (200 to 305 days) may result in biased sire proofs. In the U. S. A. all the records of less than 305 days are projected to 305 day lactation for genetic evaluation. The projection methods may be slightly inaccurate and the

genetic correlation between projected and completed records may not always be unity. A selection index or multivariate BLUP approach could handle the records of varying length simultaneously without the need for projecting the records in progress. Prediction of breeding values for total lactation from TD records, and inclusion of records in progress in the form of TD records without projecting, are discussed in chapter 6. Finally, discussions and conclusions from this study are dealt with in chapter 7.

CHAPTER 2

LITERATURE REVIEW

This review is divided into six part: (1) methods of prediction of phenotype for 305-day lactation yields of milk, fat and protein, (2) environmental factors causing variation in TD yields, (3) genetic parameters of test day yields, (4) extension of records in progress (RIP), (5) problems of inclusion of RIP in genetic evaluations and their solutions, and (6) genetic evaluations on part and complete lactation / test day records.

2.1 Methods for Phenotypic Prediction of Lactation Yield

There have been several attempts to predict 305-day yield from test day (TD) yields. McDaniel (1969) reviewed the accuracy of various sampling procedures for phenotypic prediction of lactation yield. The prediction error (SD of difference between predicted and actual yield) in predicted 305-day yield (predicted by linear interpolation) increased with the increase in interval between tests. Weekly testing gave extremely reliable results with less than 2% prediction error. Monthly testing proved to be the next best with an error of 3%. The error of prediction was found to be more for fat yield than for milk yield. Bi-monthly sampling resulted in 30% more error (i. e. 4%) than monthly sampling. Recently, Anderson *et al.* (1989) compared the effect of frequency (monthly, bi-weekly and weekly) and spacing of sampling (equal and unequal intervals) on accuracy and precision of prediction of total lactation yield. They predicted lactation yield by two methods: (1) linear interpolation between sample points with

adjustment of TD yields in the early and the late part of the lactation by "Shook Factors" (Shook et al., 1980) (cited by Anderson *et al.*, 1989), and (2) fitting a non linear function to the lactation curve as proposed by Wood (1967). They concluded that all sampling methods tended to overestimate the actual yield, probably due to overestimation of the ascending phase of the lactation curve. Biases (predicted-actual yield) in prediction of lactation yield were similar for the methods of prediction. The largest bias occurred when the post peak period was less frequently sampled. The prediction errors were slightly lower for the non linear method than for linear interpolation.

From these studies it is clear that errors in estimation of 305-day yield from monthly TD yields are of the order of 3%. Most countries are now using monthly sampling for the prediction of the lactation yield by linear interpolation. In the U.S.A., some adjustment is made for monthly TD yields in the first and the last part of the lactation (Shook, 1975) (cited by Anderson *et al.*, 1989). The Milk Marketing Board (MMB) of England and Wales are also using linear interpolation without any adjustment (British Standards 4866 (1972) method 3). The MMB have suggested a new method of predicting 305-day yield which is based on the average parameters of lactation curves within environment groups (i. e. herds or herd-year-season) (Gnanasakthy, 1989). From these parameters, expected daily yields of the average cow are estimated and expressed as a proportion of the total expected yield (sum of expected daily yields) for the lactation.

All the above mentioned results are based on the data from the experimental herds where the actual lactation yield is known. In all the progeny testing programmes milk yield is recorded once in a month

and actual lactation yield is not known. Thus continued prediction of 305-day yield by a method developed long back and which is based on the phenotypic regressions can not be justified.

2.2 Environmental Factors causing Variation in TD Yields

Knowledge of variation in TD yields due to environmental factors is essential for correct estimation of genetic parameters and breeding values. Numerous environmental factors influencing TD yields have been reported in the literature. The most common are herd, age and month of calving, length of first test period (interval between calving and first test), days open, calving interval and lactation length.

2.2.1 Herd Effect

The herd effect comprises herd level of prodution (herd average), management and feeding practices. The usual way of removing the herd effect is to include the herd as herd-year-season (HYS) in the model. The herd effect may also be studied by fitting the herd average as a covariable if computation facilities are limited. Auran (1973) studied herd effect as the regression of TD milk yields on herd average and found that the herd effect varied for various TD milk yields. The reduction in sums of squares (SS) due to herd average was comparatively less for the first and the last TD yields than for the TD yields in mid lactation (Table 2.1). This may be due to poor agreement (low correlation) between herd average and TD yields in the early and the late part of the lactation. He observed that herd level explained 72% of the herd influence on the lactation yield. The corresponding figures for first TD yield, TD yields in mid

Table 2.1: Percent reduction in sums of squares of TD milk yields due to various environmental factors.

	TD									
REFERENCES	1	2	3	4	5	6	7	8	9	10
				AGE AT CALVING						
3	41.5	40.6	35.9	30.6	23.8	16.6	7.9	1.9	1.0	2.0
4	2.6	2.5	2.4	2.1	1.8	1.7	1.5	1.4	0.9	0.5
				MONTH OF CALVING						
3	1.8	2.9	3.8	4.0	4.3	4.8	7.9	7.8	5.4	2.8
4	1.0	2.5	4.1	3.3	2.7	3.0	3.9	5.3	6.2	5.4
				LENGTH OF FIRST TEST PERIOD						
3	2.3	0.4	1.4	1.2	1.2	1.5	2.5	3.8	4.7	3.8
4	4.0	0.6	1.0	0.6	0.5	0.5	0.6	1.1	2.2	3.4
				DAYS OPEN						
4	0.2	0.4	0.7	1.0	1.3	1.3	1.8	3.0	7.6	18.8
				REGRESSION ON HERD AVERAGE						
3	12.7	17.6	19.9	21.7	22.8	22.6	21.4	15.9	10.1	5.1
4	9.8	15.5	16.6	18.0	18.5	18.0	17.4	15.4	12.1	6.5

3: Auran (1973);
4: average estimates from Danell (1982a).

lactation and the last 2-3 TD yields were 60, 66 to 68 and 40 to 60 percent, respectively. Thus herd average should not be used to remove the herd effect particularly for TD yields in early and the last part of the lactation. More recently, Meyer *et al.* (1989) reported that herd-year-month of test (HYMT) removed more variation from TD yields of milk, fat and protein compared to HYS, indicating the importance of time of test.

2.2.2 Age Effect

Age at first calving affects first lactation yield significantly (Lee and Hickman, 1972; Amir *et al.*, 1978 and Danell, 1982a). Heifer TD yields of milk, fat and protein are also affected by age (Auran, 1973; Danell, 1982a; and Wilmink, 1987a). Age at calving is the second major factor, contributing 10-20% of the variation in 305-day yield and TD yields (Syrstad, 1965; RØnningen, 1967; and Auran, 1973). The age effect was largest (in terms of reduction in SS) on the first TD milk yield and then reduced gradually as the lactation advanced (Table 2.1) indicating that the heifers are gradually maturing towards the end of the lactation. Auran (1973) reported much higher figures for reduction in SS due to age than Danell (1982a). The higher estimates reported by Auran (1973) are probably due to the different model and age groupings used and correction of data for calving interval before analysis, because calving interval is negatively correlated with age at calving.

Table 2.2: Least squares constants of TD milk yields (kg) for month of calving.

MONTH OF CALVING	REFERENCE	TD 1	3	5	7	9
JAN.	3	0.82	1.09	1.48	0.73	-0.82
	4	0.39	0.74	1.12	0.44	-1.13
FEB.	3	0.63	1.36	0.71	0.22	-1.29
	4	0.58	1.07	0.65	0.08	-1.89
MAR.	3	0.69	1.25	0.72	-0.85	-1.56
	4	0.34	0.87	0.00	-1.39	-1.76
APR.	3	0.92	0.69	0.32	-1.68	-1.29
	4	0.40	0.31	-0.21	-1.81	-1.15
MAY	3	0.79	0.20	-1.02	-1.72	-1.15
	4	0.33	-0.06	-1.39	-1.30	-0.31
JUN.	3	-0.60	-0.49	-1.92	-1.04	-0.26
	4	0.27	-0.87	-1.40	-0.46	0.19
JUL.	3	-0.86	-1.15	-1.34	-0.31	0.72
	4	-0.76	-1.51	-0.82	-0.05	0.85
AUG.	3	-1.22	-2.01	-1.12	0.22	1.40
	4	-0.55	-1.09	-0.09	0.49	1.35
SEP.	3	-0.92	-1.02	-0.07	1.17	2.39
	4	-0.57	-0.37	-0.31	0.84	1.91
OCT.	3	-0.68	-0.42	0.24	1.51	1.32
	4	-0.46	0.12	0.47	1.11	1.06
NOV.	3	-0.08	0.29	0.73	1.37	0.26
	4	-0.20	0.29	0.51	1.25	0.57
DEC.	3	0.51	0.60	1.28	0.81	-0.03
	4	0.24	0.51	0.86	0.81	0.31

3: Auran (1973);
4: Danell (1982a).

2.2.3 Month / Season of Calving

Seasonal effects reflect variation in climatic factors and availability of feed and fodder. The effect of month of calving, measured as the magnitude of least squares constants, was found to be much greater on the last 2 or 3 TD yields than the first (Table 2.2) (Syrstad, 1965; Spike and Freeman, 1967; Auran, 1973; Danell, 1982a; and Wilmink, 1987a). It may be due to differences in management practices, availability of feed and fodder and climatic factors. Body reserves can supply part of the energy needs during early lactation, hence TD yields in the early part of lactation are likely to be less influenced by month of calving. Miller *et al.* (1967) and Danell (1982a) observed an interaction between month of calving and stage of lactation. This interaction may be viewed as differences in climatic factors and availability of feed and fodder during different stages of lactation. A careful examination of least squares constants for monthly TD yields of milk revealed that calving months which are favourable (i.e. with positive least squares constants) at the beginning of the lactation have adverse effects at the end (least squares constants become negative), and vice versa (Table 2.2).

Length of first test period, days open, calving interval, lactation length and stage of lactation also influence the TD yields (Table 2.1). Length of first test period was studied by Auran (1973) and Danell (1982a); days open and calving interval by Danell (1982a); lactation length by Lindgren *et al.* (1980) and stage of lactation by Lindgren *et al.* (1980) and Wilmink (1987a).

Table 2.3: Heritability estimates (X100) of TD milk yields and predicted 305-day lactation milk yield (LMY).

REFERENCES	TD 1	2	3	4	5	6	7	8	9	10	LMY
1	11	17	22	19	19	15	14	14	12	08	-
2	14	19	20	20	22	24	23	23	22	20	25
3	20	18	20	18	22	25	22	20	23	16	-
4a	26	24	22	27	23	22	22	27	23	24	31
4b	16	15	18	22	24	27	27	27	23	20	30
4c	21	20	27	27	31	30	26	19	20	12	30
5	15	19	20	21	21	25	23	14	22	08	-

1: Van Vleck and Henderson (1961a);
2: Keon and Van Vleck (1971);
3: Auran (1976), range of S. E. = 0.05-0.08;
4a, 4b and 4c: Danell (1982b), range of S. E. = 0.02-0.07;
5: Meyer *et al.* (1989), range of S. E. = 0.03-0.06.

Table 2.4: Estimates of genetic (below diagonal)and phenotypic (above diagonal) correlation (X100) among TD milk yields and LMY.

		TD										
REFERENCES		1	2	3	4	5	6	7	8	9	10	LMY
TD1	1	-	60	53	50	46	41	33	25	14	06	58
	2	-	76	69	64	59	55	50	44	36	25	70
	3	-	75	67	61	57	55	48	42	30	18	68
	4	-	70	63	57	53	48	42	38	29	20	68
	5	-	59	51	46	43	40	34	25	25	20	-
TD2	1	100	-	80	72	61	55	48	36	25	17	76
	2	96	-	84	79	73	68	63	55	46	34	82
	3	93	-	83	76	66	68	63	54	43	27	81
	4	95	-	78	71	66	59	53	46	35	23	77
	5	90	-	63	57	51	47	42	33	28	26	-
TD3	1	94	92	-	82	69	62	56	45	30	22	81
	2	90	98	-	86	80	75	70	62	51	37	86
	3	75	93	-	85	79	75	69	60	48	31	86
	4	86	95	-	81	75	69	61	53	42	29	83
	5	88	93	-	63	56	56	47	38	33	28	-
TD4	1	96	86	96	-	79	71	64	52	60	24	85
	2	82	84	98	-	86	81	76	68	57	42	89
	3	77	99	94	-	85	79	67	64	50	33	87
	4	79	92	96	-	82	74	67	58	46	33	85
	5	83	91	94	-	63	56	51	42	37	29	-
TD5	1	91	80	100	101	-	82	73	60	42	27	85
	2	78	91	96	100	-	86	81	73	63	47	90
	3	80	98	95	100	-	84	77	68	54	36	88
	4	75	88	94	97	-	81	72	63	50	36	87
	5	80	89	90	93	-	63	56	45	40	34	-
TD6	1	74	58	83	90	98	-	83	67	48	32	85
	2	73	84	93	97	99	-	86	79	69	53	90
	3	68	92	92	98	100	-	84	73	59	39	89
	4	69	80	90	92	97	-	80	68	55	40	85
	5	72	79	87	89	92	-	62	50	43	36	-
TD7	1	82	55	91	98	102	103	-	78	57	39	83
	2	65	79	88	94	97	99	-	85	74	58	89
	3	70	84	92	98	96	102	-	82	66	45	87
	4	65	79	87	92	94	98	-	77	61	44	84
	5	70	79	87	87	91	95	-	57	49	38	-
TD8	1	69	43	70	70	82	97	103	-	77	55	78
	2	54	68	79	86	91	95	98	-	82	66	85
	3	66	76	84	96	86	97	95	-	79	60	84
	4	63	73	81	83	87	95	97	-	74	55	80
	5	51	65	74	75	80	87	92	-	55	48	-

CONTD.....

TD9	1	65	50	74	71	94	115	100	100	-	75	66
	2	45	58	70	77	84	88	93	98	-	79	79
	3	61	69	70	79	82	96	95	99	-	82	76
	4	48	53	64	68	75	87	88	95	-	70	72
	5	51	53	60	61	65	72	80	88	-	61	-
TD10	1	16	24	42	45	67	97	78	98	108	-	53
	2	36	47	56	62	69	74	79	84	95	-	65
	3	32	35	33	53	60	75	73	84	91	-	59
	4	25	34	45	49	58	73	78	79	90	-	57
	5	39	44	41	52	54	52	47	61	76	-	-
LMY	1	89	79	94	95	101	98	99	88	94	71	-
	2	78	88	94	97	99	98	97	87	89	77	-
	3	78	90	89	95	97	103	99	96	94	73	-
	4	78	87	94	94	96	99	95	95	85	71	-
	5	-	-	-	-	-	-	-	-	-	-	-

1: Van Vleck and Henderson (1961a);
2: Keon and Van Vleck (1971);
3: Auran (1976), maximum S. E. = 0.21;
4: Danell (1982b), range of S. E. = 0.02-0.14;
5: Meyer *et al.* (1989).

Table 2.5: Pooled estimates (average) (X100) of heritability (diagonal), genetic (below diagonal) and phenotypic (above diagonal) correlation among TD milk yields and LMY.

		TD										
		1	2	3	4	5	6	7	8	9	10	LMY
1	TD1	*18*	68	61	56	52	48	41	35	27	18	66
2	TD2	95	*19*	78	71	63	59	54	45	36	25	79
3	TD3	87	94	*21*	79	72	67	61	52	41	32	84
4	TD4	83	92	96	*22*	79	72	66	57	45	36	87
5	TD5	81	89	95	98	*23*	79	72	62	50	40	88
6	TD6	71	79	89	93	97	*24*	79	68	55	45	87
7	TD7	70	75	89	94	96	98	*22*	76	61	57	86
8	TD8	61	65	78	82	85	94	96	*21*	73	57	82
9	TD9	54	57	68	71	80	89	91	96	*21*	73	73
10	TD10	30	37	43	52	62	74	71	81	90	*15*	58
LMY		83	86	93	95	98	99	97	94	90	73	*29*

From these studies it is clear that the major environmental sources of variation in TD yields of milk, fat and protein are HYMT / HYS and age at calving, together accounting for 60-70% of the total variation identified. Although the interval between calving and first test does not explain much variation in TD records, it is important for making correction due to differences in day of lactation. One of the similarity among all these studies was that the effect of some of these factors was different for TD yields in different stages of lactation. HYMT explained more variability in TD yields than HYS indicating the importance of month of test.

2.3 Estimates of Genetic Parameters

2.3.1 Heritability Estimates

The estimates of heritability of TD milk yields were lower than of 305-day milk yield (Searle, 1961; Van Vleck and Henderson, 1961a; Keon and Van Vleck, 1971; Auran, 1976a; and Danell, 1982b). The same pattern was observed for fat and protein yields. In these studies, heritability estimates for TD milk yields were higher in mid lactation (Table 2.3 and 2.5). Estimates for TD fat yield (Table 2.8 and 2.10) and protein yield (Table 2.11) were also higher in mid and late lactation. Recently Wilmink (1987c) and Meyer *et al.* (1989) obtained similar results using restricted maximum likelihood (REML) procedures. The heritability estimates of cumulative TD yields were equal to or greater (0.22-0.31) than of predicted 305-day yield (Wilmink, 1987c).

These studies indicate that TD yields in mid lactation have consistently higher heritability estimates than TD yields at the

Table 2.6: Standard deviations of TD yields of milk, fat and protein.

REFERENCES	TD									
	1	2	3	4	5	6	7	8	9	10
					MILK YIELD (kg)					
3	4.9	4.8	4.0	4.4	3.7	3.5	3.6	4.1	4.8	5.0
4a	3.2	3.2	3.2	3.1	3.1	3.1	3.2	3.4	3.9	5.0
4b	3.8	3.8	3.8	3.6	3.6	3.6	3.6	3.8	4.2	5.3
5	3.0	3.0	2.7	2.5	2.3	2.2	2.1	2.1	2.0	1.9
POOLED	3.7	3.7	3.4	3.4	3.3	3.2	3.3	3.5	4.0	4.5
					FAT YIELD (kg)					
4a	0.17	0.15	0.14	0.13	0.13	0.13	0.13	0.14	0.16	0.22
4b	0.19	0.16	0.15	0.14	0.14	0.14	0.14	0.15	0.17	0.22
5	0.14	0.13	0.12	0.11	0.10	0.09	0.09	0.10	0.09	0.09
POOLED	0.17	0.15	0.14	0.13	0.13	0.13	0.13	0.14	0.15	0.20
					PROTEIN YIELD (kg)					
5	.103	.092	.085	.076	.070	.068	.067	.068	.066	.068

3: Auran (1973);
4a and 4b: Danell (1982a);
5: Meyer *et al.* (1989).
Pooled: weighted average (weights are the number of records).

Table 2.7: Percent missing records with advancement in lactation compared to day 10 of lactation (Wilmink, 1988).

DAY OF LACTATION	PERCENT MISSING RECORDS
10	0
30	0
50	1
70	2
90	2
110	3
130	4
150	5
170	6
190	7
210	9
230	10
250	13
270	20
290	36

start and the end of lactation. Using pooled estimates (average) of heritability (Table 2.5) and phenotypic standard deviation (Table 2.6), it is observed that the genetic variance of TD milk yields is similar except for the first TD, for which it is lower. The phenotypic variance is higher at the beginning and at the end than the mid part of the lactation. Thus the lower estimate of heritability for first TD milk yield reflects a high phenotypic variance and low genetic variance and for the last TD yield a higher phenotypic variance.

As the lactation advances the number of cows in milk reduces due to culling or completion of lactation (Table 2.7). Wilmink (1988) examined the effect of culling on the estimates of genetic parameters of cumulative yields. The sire component of variance and heritability of cumulative yield to the second month increased when extended incomplete records were included in the analysis but genetic correlations among cumulative yields and between cumulative and predicted 305-day yield did not change much. However, Meyer *et al.* (1989) did not observe any change in heritability estimates from successive multivariate anlysis of TD records.

2.3.2 Genetic and Phenotypic Correlations

Estimates of genetic correlations between predicted 305-day yields and TD yields of milk (Table 2.4 and 2.5) and fat (Table 2.9 and 2.10) were higher (0.70 to 0.99) for TD yields in mid lactation than for TD yields at the begining and the end of the lactation. Genetic correlations among TD yields of milk, fat and protein were also higher during mid lactation. All studies (Auran, 1976a; Wiggans and Van Vleck, 1978; and Danell, 1982b) revealed the same pattern.

Table 2.8: Heritability estimates (X100) of TD fat yields and predicted lactation fat yield (LFY).

REFERENCES	TD 1	2	3	4	5	6	7	8	9	10	LFY
1	19	13	15	07	16	10	08	08	10	07	-
4a	24	19	19	20	17	17	18	21	21	21	28
4b	16	11	12	15	16	20	20	19	17	16	22
5	12	11	12	11	11	13	12	11	15	09	27

1: Van Vleck and Henderson (1961a);
4a and 4b: Danell (1982b), range of S. E. = 0.02-0.07;
5: Meyer *et al.* (1989), range S. E. = 0.02-0.04.

Table 2.9: Estimates of genetic (below diagonal)and phenotypic (above diagonal) correlation (X100) among TD fat yields and LFY.

			TD										
REFERENCES			1	2	3	4	5	6	7	8	9	10	LFY
1	TD1	4	-	61	53	46	42	39	34	31	25	18	64
		5	-	46	38	34	33	32	28	22	22	10	-
2	TD2	4	85	-	67	58	53	47	42	38	30	19	73
		5	82	-	48	42	39	36	33	26	23	21	-
3	TD3	4	82	98	-	68	60	54	49	43	34	23	77
		5	74	82	-	49	43	39	37	31	28	24	-
4	TD4	4	70	91	91	-	69	62	55	48	39	28	78
		5	68	82	84	-	50	44	41	35	31	24	-
5	TD5	4	68	84	92	96	-	69	59	52	42	30	78
		5	64	77	84	84	-	51	46	39	34	29	-
6	TD6	4	55	76	84	88	96	-	69	58	48	33	78
		5	60	75	73	84	80	-	52	42	37	31	-
7	TD7	4	49	73	78	86	89	102	-	69	53	37	77
		5	56	70	76	79	83	85	-	48	42	31	-
8	TD8	4	44	64	71	80	84	96	96	-	66	46	74
		5	59	67	69	81	84	89	92	-	47	40	-
9	TD9	4	32	41	37	61	72	84	91	95	-	65	68
		5	59	68	71	74	78	86	86	89	-	49	-
10	TD10	4	14	18	33	40	54	71	79	81	81	-	53
		5	46	30	57	52	49	54	56	69	70	-	-
LFY		4	70	84	89	92	97	98	96	92	81	66	-
		5	-	-	-	-	-	-	-	-	-	-	-

Table 2.10: Pooled estimates (average) (X100) of heritability (diagonal), genetic (below diagonal) and phenotypic (above diagonal) correlation among TD fat yields and LFY.

		TD										
		1	2	3	4	5	6	7	8	9	10	LFY
1	TD1	*19*	54	46	40	38	36	31	27	24	14	66
2	TD2	84	*13*	58	50	46	42	38	32	27	20	73
3	TD3	78	90	*15*	59	52	47	43	37	31	24	77
4	TD4	69	86	78	*14*	60	53	48	42	35	26	78
5	TD5	66	75	88	90	*17*	60	53	46	38	30	78
6	TD6	57	71	78	86	88	*16*	61	50	43	32	78
7	TD7	52	69	77	82	86	92	*16*	59	48	34	77
8	TD8	51	65	70	80	84	92	94	*15*	57	43	74
9	TD9	45	54	59	67	75	85	88	92	*16*	57	68
10	TD10	30	24	45	46	51	62	67	75	75	*13*	53
LFY		70	84	89	92	97	98	96	92	81	66	*26*

Table 2.11: Estimates (X100) of heritability (diagonal), genetic (below diagonal) and phenotypic (above diagonal) correlation among TD protein yields (Meyer *et al.*, 1989)

		TD									
		1	2	3	4	5	6	7	8	9	10
1	TD1	*10*	44	38	33	32	32	28	21	21	21
2	TD2	74	*10*	49	43	38	37	35	28	26	26
3	TD3	70	83	*17*	50	43	40	39	33	30	26
4	TD4	69	77	83	*15*	51	45	43	39	31	26
5	TD5	62	73	82	86	*14*	51	46	41	34	31
6	TD6	60	72	85	86	79	*19*	53	45	36	33
7	TD7	54	71	69	80	83	82	*20*	53	43	37
8	TD8	63	63	44	66	76	77	86	*16*	52	45
9	TD9	63	73	66	69	72	77	81	83	*11*	54
10	TD10	72	71	53	73	72	79	80	83	71	*07*

However, Searle (1961) reported negative genetic correlations between TD yields of fat at the begining and the end of lactation. Sampling errors of these estimates were high (0.20) indicating low precision of these estimates. Phenotypic correlations among TD records and between TD and lactation records were lower than the coresponding genetic correlations but followed the same trend. Recent studies by Wilmink (1987c), Meyer *et al.* (1989) and Meinert *et al.* (1989) are in line with the earlier findings.

Wilmink (1987c) reported genetic correlations close to unity between predicted 305-day yield and extended 180-day milk, fat and protein yields or cumulative milk, fat and protein yield between 91 and 180 days post-partum, which is obvious because cumulative yield and 305-day yield are functions of TD yields.

Studies on genetic parameters indicate that TD yields in the mid part of the lactation have the highest heritabilities and genetic correlations among themselves and with predicted 305-day yield, indicating the scope for improvement in 305-day yield by index selection incorporating TD records. Danell (1990) reviewed the studies on inheritance of different parts of lactation and arrived at similar conclusions.

Phenotypic correlations between TD records and corresponding lactation records may be biased because lactation records are predicted from the TD records.

2.4 Projection of Records in Progress

Several methods have been used for the projection / extension of RIP / part records but the ratio method (Lamb and McGilliard, 1960a; 1960b; Van Vleck and Henderson, 1961c; and Syrstad, 1964), regression

techniques (Madden *et al.*, 1959; Van Vleck and Henderson, 1961b; 1961d; 1961f) and non linear techniques (Wood, 1969; Nelder,1966; and Schaeffer *et al.*, 1977) are most commonly employed. A brief description of these methods follows:

2.4.1 Ratio Method

$$\hat{TL}_{305} = C_i PL_i;\ C_i = TL_{305}/PL_i$$

2.4.2 Regression Method

$$\hat{TL}_{305} = \bar{TL}_{305} + \Sigma\, b_i(TD_i - \bar{TD}_i)$$

Where,

$\hat{TL}_{305}$ = projected 305-day yield;

TL_{305} = predicted 305-day yield;

C_i = extension factor;

PL_i = cumulative yield up to i^{th} day of lactation;

TD_i = yield for i^{th} TD;

b_i = are regression coefficients of TL_{305} on TD_i.

2.4.3 Gamma Function

$$Y_t = at^b e^{-ct}$$

2.4.4 Inverse Polynomial

$Y_t = t/(b_0 + b_1 t + b_2 t^2)$

Where,

Y_t = yield on the middle day of t^{th} week;

a, b, c, b_0, b_1 and b_2 = constants;

e = base of natural logrithm.

2.4.5 Exponential Model

$Y_{ij} = A \exp(-\beta(i-t_0)[(1-\exp(-B(i-t_0))] / B \exp(\epsilon_{ij})$

Where,

Y_{ij} = milk yield on i^{th} day of lactation of j^{th} cow;

t_0 = lag time parameter and may indicate when a cow's udder begins to lactate prior to calving;

B = slope of lactation curve during the increasing phase;

A = parameter associated with peak production;

β = slope during the decline phase;

ϵ_{ij} = residual effect.

Auran (1976b) compared ratio and regression methods and concluded that the ratio method, which estimates the remaining part of the lactation from the last TD yield, is better because of its simplicity and high precision. Danell (1982c) used these methods and showed that the extension factors vary with stage of lactation and level of

production in the herd, suggesting that these methods still need refinement. The regression coefficients used in these equations also depend on the day of lactation.

Schaeffer *et al.* (1977) compared the exponential model with the ratio method and concluded on the basis of bias (predicted-actual yield) and error (SD of difference between predicted and actual yields) that these methods have the same accuracy. However, Rao and Sundaresan (1980) found that ratio and multiple regression methods are better than non-linear methods but Congleton and Everett (1980) stated that non-linear equations have smaller prediction errors. Agyemang *et al.* (1985a) and Bar Anan *et al.* (1986) studied the effect of environmental factors on extension factors. The model which included all possible environmental factors had slightly higher accuracy of prediction.

Recently Wilmink (1987b) compared single and multiple regression of unknown TD yields on known TD yields and factor analysis (Harman, 1970) models and did not find any difference between these three methods, but estimation of mean within herd-age-month of calving subclasses improved the accuracy of prediction. In the factor analysis approach, each standardised TD yield was represented by the sum of several common factors, which are mutually independent, and a unique factor. The common factors account for covariances among the standardised TD yields and the unique factors account for the remaining variance of the standardised TD yields. Goddard (1987) suggested a method using genetic regresssion coefficients of yield of complete lactation on part lactation for extending part lactations, but it is based upon the assumption of a genetic correlation of unity between complete and part lactation.

Results from these studies show that all the above described methods of projecting RIP are similar in terms of phenotypic accuracy and bias. In the U. S. A., RIP are projected by a method based on regression of yield in the remaining part of lactation on last TD yield (Wiggans and Dickinson, 1985). For records of less than 155 days herd average is also included in the projection. The MMB are not including projected records in their sire evaluation, only records with lactation length of 200 days or more are included.

2.5 Problems Associated with Inclusion of RIP in Genetic Evaluation and their Solutions

Although the projected 305-day records have high genetic correlations with complete records (predicted 305-day records), many problems have been reported when projected records are included in sire evaluation. First, inclusion of incomplete projected records may lead to more change in sire proofs than expected when more complete records are added later (Agyemang, 1985a). Second, Famula and Van Vleck (1981) reported that sire proofs calculated entirely from projected incomplete records were overestimated as compared to proofs from complete records. Third, extension factors were found to be sensitive to level of production and yearly variation in the shape of the lactation curve (Danell, 1982c). All these problems are due to differences in variances of projected and complete records and a genetic correlation of less than unity between projected and complete records.

Various alternatives have been suggested to overcome these problems. Agyemang (1985b) suggested dividing the complete lactation into parts (trimesters of lactation), analysing these parts as

separate traits by multi-trait methodology, and combining these part lactation evaluations into an overall evaluation. The proposal of Agyemang (1985b) requires extension of incomplete records within a trimester, which is less noisy, but multivariate anlysis is computationally demanding. Weller (1988) suggested differential weighting of partial records, where the elements of the incidence matrix for sires are set equal to the regression of the sire effect of projected records on the sire effect of the complete records. For complete records the elements of the incidence matrix are set equal to 1 in usual way. Diagonal elements of the residual matrix were actual error variances of projected and completed records. VanRaden *et al.* (1991) pointed out that this method could not be directly implemented and requires more computation. They suggested expansion of projected records by the ratio of sire SD for completed and projected records in order to match the genetic variance of projected records to that of completed records. Since projected records have greater error variance, less weight is given to projected records in animal model evaluation.

All the methods presented above except the multivariate approach assume a genetic correlation of unity between the projected and completed records, which is not the case, particularly for projected records from the first one or two TD records. Variances of projected records could be scaled but the genetic correlations of less than unity could only be handled properly by a multivariate analysis.

2.6 Genetic Evaluations on Partial and Complete Lactation / TD Records

Hickman (1971) and Van Vleck (1975) reported that if selection is based upon the first half of the lactation, it will change the shape of lactation curve because the genetic correlation between different parts of the lactation is less than unity. Van Vleck and Henderson (1961e; 1961g) observed that preliminary evaluation of sires may be based upon the first five TD milk yields and meaningful decisions could be made on the basis of 7-8 months TD yields. To get an accuracy nearly equal to that based upon 10 monthly records(not actual lactation yield), 7-8 months TD records are required. In these studies approximately equal heritability of 8 and 10 months cumulative yield (0.22 and 0.23, respectively) and genetic correlation of 0.99 between 8 and 10 months yield were assumed, implying that both traits are the same, and sires may be selected on their daughter's 8 months cumulative yield. The accuracy of the progeny test can be increased from 0.61 to 0.78 by including 20 five-month part records in addition to 10 complete records which is obviously due to increased number of daughters and a genetic correlation of 0.94 between 5 month and 10 month cumulative yield.

Famula and Van Vleck (1981) reported that correlations between sire proofs with complete lactation and with only extended part lactation records from 130 to 160 days were 0.93 or 0.94 regardless of the method of extension. However these correlations reduced drastically when extended 60-80 days records were used. In this study the same daughters were used for prediction of sire proofs from predicted complete lactation and extended part records from day 130-160 of lactation and the heritability was also assumed the same. A high

correlation between sire proofs from complete and extended records is due to similar heritability estimate and the high genetic correlation between complete and extended records.

Tandon and Harvey (1984) compared the number of daughters required for sire evaluation on part lactation and complete lactation records for the same accuracy and concluded that an evaluation with 18 daughters per sire from 150-day milk yield gives an accuracy equivalent to evaluation from predicted 305-day milk yield for 14 daughters. They assumed a lower heritability estimate of 150-days yield than of 305-days yield (0.32 and 0.38, respectively) and a genetic correlation of less than unity (rg = 0.93) between 150-days and 305-days yield. When the additive genetic relationships among sires were included then the increase in accuracy of evaluation by part lactation was more than by complete lactation.

Wilmink (1987c) predicted that efficiency of selection (ratio of correlated and direct annual response in predicted 305-day yield) on cumulative milk and protein yield between 91-180 days post partum, and extended 180 day fat yield would be 1.05, 1.06 and 1.04, respectively. Such predictions depend on the heritability of part / extended lactation records and genetic correlation between part / extended and complete records, and generation interval used in this study. The major part of the increased correlated response in predicted 305-day yield can be explained by reduced generation interval due to selection based on the early part of the lactation.

Danell (1979) constructed a selection index combining 10 TD milk records and concluded that small changes in heritabilities of TD records resulted in drastic changes in index weights, but the accuracy of the index, aggregate response and response in single TD

were not affected. The changes in index weights influenced the index values for each individual with differential ranking of bulls with different heritabilities.

Results indicate that selection on records from only the first half of the lactation is not so accurate for making final selection decisions because the heritability for cumulative yield over the first half of the lactation is less than for complete lactation and the genetic correlation between first half and complete lactation is also less than unity. But the TD records in the early part of lactation could serve as guidelines for culling of sires / cows of very low breeding value in a sequential selection programme.

2.7 Conclusions

Estimates of heritability of TD records and their genetic correlations with predicted 305-day records indicate scope for improvement via direct selection on an index of TD records. There is no information on genetic parameters of TD records of British Holstein-Friesian heifers and as herd-year-month (of test) explained more variation in TD records compared to herd-year-season, it is worth calculating genetic parameters of TD records using herd-year-month (of test) instead of herd-year-season in the statistical model.

Inclusion of records in progress in genetic evaluations along with completed records have been addressed in the literature by extending RIP to 305-day records. The projection methods used may not be accurate. Thus, there is a need to address this problem on TD basis without projecting them to 305-day records.

Phenotypic correlations between TD records and predicted 305-day records may be biased as the latter is predicted from TD records. No attempts have been made to estimate this bias in studies in the past. Phenotypic correlations among records less than one month apart are also not known.

Studies on genetic evaluations on the basis of part records indicate that the final selection decision should not be based only on the basis of first five TD yields. It will be worthwhile, however, to explore the possibility of selection on five TD records taken at two months interval over the entire lactation, which would reduce the cost of recording.

Presently genetic evaluations are based on predicted phenotype for 305-day records of milk, fat and protein. This predicted phenotype may be slightly biased and inaccurate. Genetic evaluations on an index of TD records may provide an alternative.

The present study will address some of the above mentioned needs and problems.

CHAPTER 3

GENETIC PARAMETERS OF TEST DAY RECORDS AND THEIR ASSOCIATION WITH 305-DAY LACTATION RECORDS

3.1 Introduction

In cow and sire evaluation programmes in the United Kingdom (U.K.) and elsewhere, selection is based on 305-day lactation yields of milk, fat and protein which are predicted from individual test day (TD) records usually taken at monthly intervals. Selection on TD yields of milk, fat and protein in the early part of lactation would result in a reduction in generation interval, cost of recording and cost of maintaining cows and bulls with low breeding values. Alternatively cost of recording may also be reduced by less frequent than monthly recording. In addition to increasing selection intensity, the use of test day records as such or as extended 305-day records can reduce the bias due to culling of heifers before the completion of 200 days of lactation, the minimum length of a lactation to qualify for inclusion in sire and cow evaluation in the U.K. TD records may also be used to increase the accuracy of sire selection by including part records in addition to complete 305-day records, or sires may be selected earlier with the same accuracy. Even for completed lactation, selection on a properly weighted index of TD records could be more accurate than selection on predicted phenotypic records for 305-day production.

All the previous studies reviewed in Chapter 2 indicated that TD records in mid lactation have a high heritability and a high genetic correlation with total lactation. There have been no such analyses on

Holstein-Friesian heifers in Britain. This study reports the effect of environmental factors on TD records and estimates of genetic parameters of TD records and their relationship with 305-day lactation records.

3.2 Material and Methods

3.2.1 Data

Data consisted of test day records obtained by National Milk Records (NMR) of the Milk Marketing Board (MMB), of milk yield and of fat and protein contents of British Holstein-Friesian heifers in 7973 herds, having their first test between November 1988 and October 1989. TD milk yield is the addition of all the individual weighings taken during a 24 hour period from noon to noon. Milk samples are also taken for milk composition analysis at the same time. TD yields of fat and protein were calculated by multiplying TD milk yield by fat and protein contents. The 305-day lactation yields of milk, fat and protein were predicted by linear interpolation using the MMB's method (British Standards 4866 (1972) Method 3). Fat and protein contents for 305-day lactation were calculated by dividing the lactation fat and protein yields by lactation milk yield. Lactations shorter than 200 days were not used for the prediction of 305-day lactation yields. The following conditions (both lower and upper limits inclusive) were set for a record to be included in the analysis: (i) age at calving was restricted between 20 and 40 months, (ii) first test had to be between day 4 and 45 of lactation, and (iii) the interval between consecutive tests was kept between 20 and 50 days. A few other records (4) were deleted as suspected outliers by subjective judgement based on histograms.

After these edits, two data sets were constructed. Data were sorted according to number of daughters per sire, and progeny of the 40 most widely used sires and 707 ''young'' sires (random sires) born between 1980 and 1985 (both inclusive) were extracted for the analysis. Although young sires born in 1980/81 would have had their progeny test results in 1986/87 and could have been used widely after 1986/87, in this data set the highest number of daughters per young sire was 56, suggesting that none were in widespread use in these data.

Data set 1 comprised records on heifers having at least 8 tests (TD1 to TD8) and data set 2 was a subset of these having at least 10 tests (TD1 to TD10). The structure of these data and others described subsequently is shown in Table 3.1. The number of daughters per old sire varied between 378 and 2432 while the corresponding figures for young sires were between 8 and 56 for data set 1. About one-fifth of the data were on offspring of the young sires. On average, 56% of herd-year-month of first test (HYMT) subclasses had only one record, 38% had 2-5 and 6% had 6-12 records. Seventy-two percent of HYMT subclasses had only one sire, 28% had 2-5 and 0.15% had 6-9 sires. HYMT classes with only one sire would contribute information only on within sire variance, those with one record only on other fixed effects. For data set 2 the proportion of the records belonging to young sires was similar to data set 1 and, since it is a subset of data set 1, the proportion of HYMT subclasses having only one record increased by 2% and of those having 6-12 records decreased by 1.2%. Similar changes were observed for the distribution of sires over the HYMT subclasses.

Table 3.1: Structure of different data sets.

DATA SET	NO.OF RECORDS	OLD SIRES	YOUNG SIRES	BASE+ ANIMALS	HYMT CLASSES	DF RESIDUAL
1	47736	40	707	232	21698	25052
2	34029	40	705	231	17014	16032
3	43482	40	624	200	20402	22207
4	40376	40	117	31	19264	20917
5	45254	40	582	220	20912	23493
6	14998	56	141	-	6549	8245
7	68275	40	708	-	26619	40903
8	76878	40	708	-	29603	46522
9	133263	40	708	-	53036	79474
10	105845	40	146	-	42967	62687

+ BASE ANIMALS = Number of paternal grandsires;
HYMT = Herd-year-month of first test;
DF = Degrees of freedom after taking account of paternal grandsires.

Table 3.2: Percent missing records with stage of lactation (test number) as compared with first test (total observations = 50014).

TEST NO.	MISSING RECORDS	
	NO.	PERCENTAGE
1	0	0.00
2	427	0.80
3	624	1.20
4	833	1.66
5	1046	2.09
6	1320	2.64
7	1731	3.46
8	2278	4.55
9	4583	9.16
10	15985	31.96

Initially it was decided to include the first 8 tests (data set 1) in the analysis in order to avoid the effect of culling bias on genetic parameter estimates because only 4.5% of the heifers do not have 8 tests while about 32% do not have 10 tests (Table 3.2). Data set 2 was analysed to estimate the genetic parameters of tests 9 and 10.

3.2.2 Statistical Methods

In current sire evaluation models for lactation production traits used by the MMB, environmental effects are fitted as herd-year-season of calving along with month of calving as a cross-classified effect (Quaas *et al.*, 1979; MMB, 1979). Since herd-year x month interaction is important (Chauhan and Hill, 1986), herd-year-month of calving is probably more appropriate effect for lactation traits. But for TD records, month of test is likely to be more important than month of calving due to environmental effects specific to the time of test (see section 2.2.3 of Chapter 2). In a preliminary analysis to determine proper partitioning of the environmental variation, the following models were tested: (i) herd and month of first test as cross-classified effects, (ii) herd-year-season (HYS) and month of first test as cross-classified effects (3 seasons were defined as groups of 4 consecutive months of test, i.e. November to February, March to June and July to October), (iii) herd-year-month of first test (subsequently adopted for the genetic analysis), and (iv) herd-year-month of calving (HYMC).

For the estimation of genetic parameters, multivariate restricted maximum likelihood (REML) (Patterson and Thompson, 1971) analyses were conducted on both the data sets using REML.PK programs (Meyer, 1986). A sire model was used and relationships among sires through

paternal grandsires were also included. The mixed linear model used was:

$$\mathbf{y = Xb + Zs + e}$$

Where:

$\mathbf{y}$ = the vector of observations;

$\mathbf{X}$ = the incidence matrix for fixed effects;

$\mathbf{b}$ = the vector of fixed effects;

$\mathbf{Z}$ = the incidence matrix for fixed (Z_1) and random (Z_2) sire effects;

$\mathbf{s}$ = the vector of fixed (s_1) and random (s_2) sire effects;

$\mathbf{e}$ = the vector of random residual errors.

With:

$E(y) = Xb + Z_1s_1$, $E(s_2) = 0$, $E(e) = 0$;

$V(s_2) = A_s\sigma^2_s$, $V(e) = I\sigma^2_w$, $Cov(s_2,e) = 0$

Where:

A_s = additive relationship matrix of random sire effects;

I = identity matrix;

σ^2_s = sire component of variance;

σ^2_w = within sire component of variance.

The fixed effects included were: (i) HYMT, (ii) pedigree status of the heifer (registered (pedigree) or grade and non-registered heifers), (iii) linear and quadratic regression on: (a) age at calving, (b) day of lactation for first test, and (c) proportion of Holstein in sire. Widely used sires (1-40) were fitted as fixed effects and young sires (41-747) and paternal grandsires (748-979) as

random.

3.3 Results

3.3.1 Fixed Effects

Overall phenotypic means, standard deviations (SD), coefficients of variation (CV) and residual SD for different traits are given in Table 3.3. In general SDs were similar across test days (TD) for each trait, except for fat and protein contents which were more variable at first test (TD1).

HYMT (model 3) accounted for a proportion 0.64 to 0.71 of the total sums of squares (TSS) for TD and 305-day lactation milk yield (Table 3.4). Corresponding figures for HYS (model 2) and herd (model 1) were 0.48 to 0.58 and 0.38 to 0.49, respectively. HYMC (Model 4) explained 0.63 to 0.69 of TSS. HYMT significantly reduced the residual variances as compared to the HYS model for test days as well as for 305-day lactation milk yield. HYMT (Model 3) accounted for 0.64 to 0.72 of TSS for TD and lactation yield of fat; corresponding figures for protein yield, fat and protein contents were 0.68 to 0.74, 0.59 to 0.62 and 0.64 to 0.71, respectively (Table 3.4).

All other fixed effects included in the model accounted for a much lower proportion of variance than HYMT. Age at calving explained a proportion 0.003 to 0.012 of TSS for TD and lactation yields of milk, fat and protein and a decreasing trend was observed as the lactation progressed; the corresponding figures for fat and protein contents were 0.000 to 0.004. Day of lactation for first test (interval between calving and first test) accounted for 0.001 to 0.004 of the TSS for TD2 to TD8 and total lactation but it explained a proportion 0.03 of TSS for milk yield and fat content of TD1 and 0.13 for

Table 3.3: Overall phenotypic means, standard deviations (SD), coefficients of variation (CV X100) and residual standard deviations (RSD) after fitting all the fixed effects for different traits for data set 1.

	TEST NO. 1	2	3	4	5	6	7	8	LACTATION
				MILK YIELD(kg)					
MEAN	19.44	20.42	19.43	18.31	17.32	16.45	15.82	15.00	*5098*
SD	4.02	3.95	3.94	3.85	3.86	3.75	3.79	3.85	*960*
RSD	3.12	3.06	3.01	2.91	2.84	2.80	2.80	2.81	*718*
CV	20.7	19.3	20.9	21.0	21.9	22.8	24.0	25.7	*18.8*
				FAT YIELD(kg)					
MEAN	0.780	0.773	0.747	0.718	0.695	0.674	0.654	0.628	*206.9*
SD	0.196	0.178	0.171	0.164	0.160	0.159	0.161	0.165	*40.4*
RSD	0.158	0.139	0.132	0.124	0.121	0.120	0.120	0.121	*28.9*
CV	25.2	23.0	23.0	22.8	23.0	23.6	24.5	26.2	*19.5*
				PROTEIN YIELD(kg)					
MEAN	0.623	0.621	0.604	0.582	0.558	0.534	0.521	0.502	*166.0*
SD	0.128	0.126	0.127	0.127	0.125	0.125	0.127	0.130	*30.5*
RSD	0.099	0.089	0.090	0.089	0.088	0.087	0.089	0.090	*21.3*
CV	20.6	20.2	21.1	21.8	22.5	23.4	24.4	26.0	*18.3*
				FAT CONTENT(g/100 g)					
MEAN	4.04	3.81	3.88	3.97	4.06	4.16	4.20	4.26	*4.06*
SD	0.653	0.520	0.526	0.529	0.542	0.559	0.573	0.588	*0.378*
RSD	0.550	0.458	0.461	0.456	0.464	0.472	0.482	0.500	*0.323*
CV	16.2	13.6	13.5	13.3	13.3	13.4	13.6	13.8	*9.3*
				PROTEIN CONTENT(g/100 g)					
MEAN	3.24	3.07	3.13	3.21	3.25	3.28	3.33	3.38	*3.25*
SD	0.323	0.248	0.265	0.266	0.263	0.259	0.254	0.266	*0.187*
RSD	0.215	0.185	0.193	0.195	0.198	0.196	0.196	0.204	*0.151*
CV	10.0	8.1	8.4	8.3	8.1	7.9	7.6	7.8	*5.7*

Table 3.4: Proportion (%) of total sums of squares (TSS) removed by herd (model 1),Herd-Year-Season (model 2) Herd-Year-Month of first test (model 3) and proportion of TSS left in residual after fitting all the fixed effects for data set 1.+

	DF	TEST NUMBER 1	2	3	4	5	6	7	8	LACTATION
MILK YIELD										
HERD+MONTH (MODEL 1)										
HERD	7973	37.9	43.4	44.7	45.1	45.6	45.6	44.8	44.7	*49.2*
RES	39009	53.1	54.2	52.8	52.4	50.0	51.8	51.3	50.0	*48.6*
HYS+MONTH (MODEL 2)										
HYS	13025	48.4	53.8	54.7	55.4	56.1	56.5	56.3	56.7	*57.9*
RES	33957	44.5	44.4	43.4	42.8	42.3	42.0	41.5	40.8	*40.8*
HYMT (MODEL 3)										
HYMT	21698	64.4	67.6	68.0	68.7	69.4	69.8	70.5	71.3	*69.5*
RES	25284	31.6	31.5	30.9	30.3	29.8	29.5	28.9	28.1	*29.7*
FAT YIELD										
HYMT	21698	63.9	66.6	67.6	69.0	69.2	69.4	69.7	70.4	*72.0*
RES	25284	34.5	32.3	31.3	30.3	32.2	30.1	29.8	29.1	*27.1*
PROTEIN YIELD										
HYMT	21698	67.6	72.4	72.8	73.2	73.3	73.4	73.5	74.2	*73.2*
RES	25284	31.4	26.7	26.4	26.0	26.0	25.9	25.9	25.2	*25.9*
FAT CONTENT										
HYMT	21698	59.1	58.7	59.1	60.5	61.1	62.0	62.4	61.6	*61.2*
RES	25284	37.5	41.4	40.7	39.3	38.7	37.9	37.5	38.2	*38.6*
PROTEIN CONTENT										
HYMT	21698	63.8	70.3	71.3	71.0	69.8	69.1	68.0	68.4	*65.1*
RES	25284	23.5	29.5	28.1	28.5	29.9	30.5	31.7	31.2	*34.6*

HYS = Herd-year-season; HYMT = Herd-year-month of first test; RES = Residual.
+ Results for fat and protein yield and content from model 3 only.

protein content of TD1. Although the proportion of TSS explained by the fraction of Holstein in the sire was small (ranging from 0.000 to 0.002), its effect was significant for all the traits.

In general HYMT explained less variation in TD1 for all the traits and a slightly increasing trend was observed in variability explained by HYMT as the lactation progressed (Table 3.4). When month of first test was fitted as a cross-classified effect along with herd or herd-year-season, it also explained more variation towards the end of the lactation.

3.3.2 Genetic Parameters

Heritability estimates for TD1-TD8 and lactation records from data set 1 and for TD9 and TD10 from data set 2 are presented in Figure 3.1. Genetic correlations among TD within a trait (from data set 1) and between TD and lactation records are given in Tables 3.5 and 3.6. Details of genetic parameters are given in the appendix (Appendix Tables 3.1-3.6). Results from data set 1 are discussed first.

The heritability estimate for lactation *milk yield* (LMY) was 0.49. The estimate was rather lower (0.27) for TD1 than for others (Figure 3.1). Estimates for TD2-TD8 were similar (0.33 to 0.39) in magnitude except for TD8 which had the highest heritability (0.43) of all the test days. Genetic correlations among test days were high (0.73 to 0.99); the highest correlations were between any two adjacent TD and the correlations decreased as intervals between tests increased (Table 3.5). Genetic correlations between test days and LMY were also high (0.87 to 0.99) with the highest values for TD3-TD8 (Table 3.6). Phenotypic correlations followed a similar pattern but were lower than the genetic correlations.

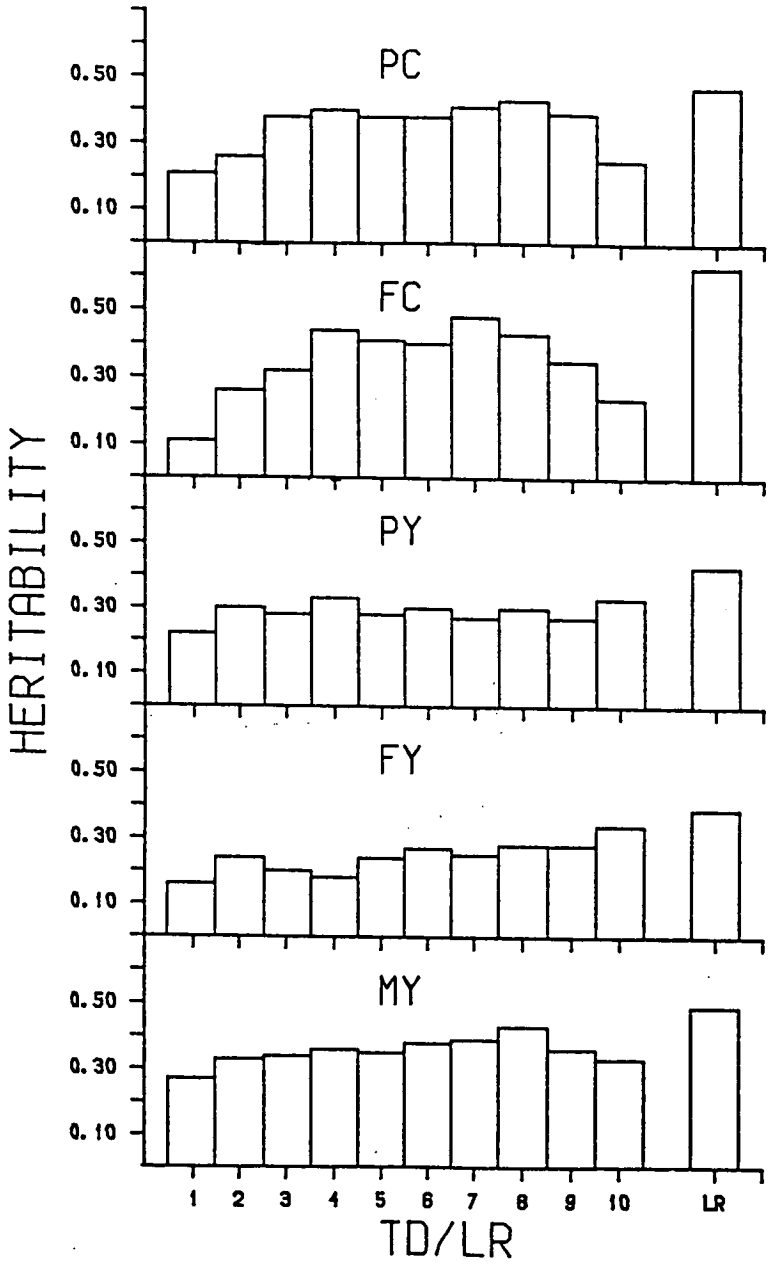

Figure 3.1: Heritability estimates of test day (TD) and lactation records (LR) of milk yield (MY), fat yield (FY), protein yield (PY), fat content (FC) and protein content (PC). (estimates for TD1-TD8 from data set 1 and for TD9-TD10 from data set 2)

Table 3.5: Average genetic correlations (X100) among TD as a function of number of tests apart (data set 1).

TESTS APART	MILK YIELD	FAT YIELD	PROTEIN YIELD	FAT CONTENT	PROTEIN CONTENT
1	97	92	93	95	94
2	95	90	92	94	90
3	91	86	90	90	85
4	89	79	87	87	78
5	83	74	84	83	70
6	79	71	82	76	65
7	73	60	79	73	56

Table 3.6: Genetic correlations (X100) of TD records with lactation records (estimates for TD1-TD8 from data set 1 and for TD9 and TD10 from data set 2)

LACTATION RECORDS	TD 1	2	3	4	5	6	7	8	9	10
MILK YIELD	87	89	97	98	99	97	98	97	95	88
FAT YIELD	77	85	95	97	97	93	97	96	97	98
PROTEIN YIELD	84	90	97	97	98	95	98	98	95	91
FAT CONTENT	81	91	98	98	99	98	96	98	94	91
PROTEIN CONTENT	76	87	95	97	96	98	92	94	89	77

The heritability estimate for 305-day lactation *fat yield* (LFY) was lower (0.39) than for milk yield. Estimates for TD fat yields were lower in the first half of the lactation than in the second half and they were substantially lower than for milk yield (Figure 3.1). Genetic and phenotypic correlations among TD fat yield and of TD with LFY revealed a similar pattern to that of milk yield but the estimates were slightly lower.

For 305-day lactation *protein yield* (LPY), the heritability estimate was 0.43. For TD protein yields, estimates of the heritability were generally higher than for fat yield (particularly during the first half of the lactation) and lower than for milk yield, but the pattern was similar to milk yield (Figure 3.1). The genetic and phenotypic correlations were similar to those for milk yield.

The heritability estimate of 305-day lactation *fat content* (LFC) was very high (0.63). The estimates for TD fat contents were lower for TD1-TD3 (0.11 to 0.32) than for TD4-TD8 (0.44 TO 0.48) (Figure 3.1). Genetic correlations among TD fat content and between TD fat content and LFC were high and revealed a similar pattern to that of milk yield, but the phenotypic correlations were substantially lower than for yield traits.

For 305-day lactation *protein content* (LPC), the heritability estimate was similar (0.47) to that for LMY. Estimates of heritability for TD protein contents were low for TD1 and TD2 (0.21 and 0.26) compared to TD3-TD8 (0.38 to 0.43) (Figure 3.1). Genetic correlations among TD protein contents and between TD protein contents and LPC were similar to those for fat yield (Table 3.5 and 3.6) and the phenotypic correlations were similar to those for milk yield.

Estimates of heritability for *TD9 and TD10 from data set 2* are also presented in Figure 3.1, and details are given in Appendix Table 3.6. Heritability estimates of lactation traits and genetic and phenotypic correlations among TD1-TD8 and between TD and lactation traits from data set 2 were similar to the estimates from data set 1. However, heritability estimates of TD1 for all the traits and of TD1-TD4 for protein yield from data set 2 were slightly lower and those for other TD slightly higher for all traits than the estimates from data set 1. As these changes may be due to sampling, results from data set 2 are discussed only for TD9 and TD10.

Heritability estimates for milk yield of TD9 (0.36) and TD10 (0.33) were slightly lower than those in mid lactation and a similar pattern was observed for fat and protein contents. Heritabilities of protein yield for TD9 and TD10 (0.27 and 0.33) were similar to those in mid lactation, but that of fat yield for TD10 was the highest (0.34) of all TD. Estimates of genetic correlations between TD9 and TD10 for all the traits ranged from 0.89 to 0.98. Genetic correlations of TD9 and TD10 with other TD and corresponding lactation traits ranged from 0.34 to 0.98 and 0.77 to 0.98, respectively. Genetic correlations between TD9 and complete lactation were generally higher than those between TD10 and complete lactation (Table 3.6).

Genetic and phenotypic *correlations among complete lactation traits* from data set 1 are presented in Table 3.7. The genetic correlations among the yields and between LFC and LPC were high and positive (0.69-0.94), the highest being between LMY and LPY. Phenotypic correlations were generally similar to genetic correlations. The genetic correlations of lactation milk yield with fat and protein contents were high and negative (-0.5) and the genetic correlations

Table 3.7: Genetic (below diagonal) and phenotypic (above diagonal) correlations (X100) among complete lactation traits from data set 1.

	LMY	LFY	LPY	LFC	LPC
LMY	-	83	94	-32	-45
LFY	72	-	86	28	-11
LPY	94	81	-	-14	-13
LFC	-56	25	-25	-	56
LPC	-53	21	-08	69	-

Range of s.e.(rg) = 0.02 for rg = 0.92 to 0.16 for rg = 0.21
Range of s.e.(rp) = 0.002-0.016

Table 3.8: Genetic correlation (X100) of TD milk yield with TD fat and protein yields and contents from data set 2.

		TD					
		1	3	5	7	9	10
		MILK YIELD					
FAT YIELD							
1	TD1	64	42	11	15	09	31
3	TD3	76	55	37	61	51	57
5	TD5	40	12	30	54	63	83
7	TD7	76	65	67	57	66	77
9	TD9	83	66	80	67	74	81
10	TD10	72	53	69	56	67	80
PROTEIN YIELD							
1	TD1	93	87	68	78	73	73
3	TD3	79	85	59	77	60	53
5	TD5	71	73	85	90	91	94
7	TD7	78	86	91	91	81	78
9	TD9	83	82	91	91	92	93
10	TD10	78	66	80	79	85	93
FAT CONTENT							
1	TD1	-56	-70	-81	-86	-81	-56
3	TD3	-24	-53	-64	-49	-44	-24
5	TD5	-31	-55	-60	-69	-60	-31
7	TD7	-09	-34	-48	-45	-37	-12
9	TD9	-14	-28	-44	-43	-49	-32
10	TD10	-07	-14	-23	-31	-32	-15
PROTEIN CONTENT							
1	TD1	-59	-63	-58	-46	-33	-13
3	TD3	-30	-30	-55	-48	-47	-21
5	TD5	-19	-37	-52	-53	-40	-11
7	TD7	-25	-38	-58	-61	-60	-48
9	TD9	-26	-22	-55	-42	-51	-39
10	TD10	-14	-11	-33	-23	-36	-31

Range of s.e.= 0.05 for rg = 0.93 to 0.22 for rg = 0.14

of LFY with LFC and LPC (0.25 and 0.21) and of LPY with LFC and LPC (-0.25 and -0.08) were low.

Genetic *correlations among TD traits*, specifically of TD milk yield with TD yields and contents of fat and protein, were similar for both the data sets, so estimates are reported from data set 2 as this was on 10 tests. For estimation of these correlations, two combinations of traits (milk yield with fat and protein yields, and milk yield with fat and protein contents) were considered and for each combination four sets of analyses including different TD were performed: (i) TD1-5 for all traits, (ii) TD6-10 for all traits, (iii) TD1-5 for milk yield and TD6-10 for yields / contents of fat and protein, and (iv) TD6-10 for milk yield and TD1-5 for yield / contents of fat and protein.

Examples of genetic correlations between TD milk yield and other traits are given in Table 3.8. The average figures for genetic correlations of TD milk yield with TD yields and contents of fat and protein were fairly similar to the figures for the complete lactation.

3.4 Discussion

Results from fitting different models for environmental effects revealed that the residual sums of squares were least for the HYMT model, indicating the importance of environmental effects specific to the time of test. Similar results were obtained by Meyer *et al.*, (1989) for TD records of Australian Black and White cows.

An increasing effect of month of first test towards the end of the lactation in the present study is in agreement with several other findings (Auran, 1973; Danell, 1982a; Wilmink, 1987a). It is perhaps

due to differences in availability of feed and fodder and management practices having a more pronounced effect on the last trimester of the lactation as body reserves can supply part of energy during early lactation. The decreasing effect of age at calving towards the end of the lactation found in the present study is in line with the results of other studies (Auran, 1973; Danell, 1982a; Wilmink, 1987a). Day of lactation for first test (interval between calving and first test) had more effect on TD1 than others as expected in view of rapid changes in milk yield during early lactation and is in agreement with other results.

The heritability estimates of LFY, LFC and LPC are more or less similar to recent U.K. estimates from pedigree Holstein-Friesian heifers using REML with a sire model (Meyer, 1987) and animal model (Visscher, 1991), but estimates for LMY and LPY are higher. Heritabilities of all these traits are higher than for other U. K. estimates obtained by least-squares using the program of Harvey (1977) (Hill *et al.*, 1983) and REML with a sire model (Meyer, 1984; Swanson and Gnanasakthy, 1990).

Estimates of genetic and phenotypic correlations among lactation traits are very similar to those reported by Visscher (1991), except that the genetic correlation between LFY and LPC is 0.21 in the present study whereas Visscher (1991) reported a small negative value in first lactation but positive in later lactations.

3.4.1 Reasons for high Heritability Estimates

The heritability estimates of 0.49 and 0.43 for LMY and LPY are higher than most in the literature. Differences between the estimates from the present study and other studies may be due to the use of

different models (i. e. HYMT vs HYS), data sets belonging to different years, different types of data (i. e. data on daughters of bulls of the MMB's Dairy Progeny Testing Scheme (DPTS) or others (NON-DPTS)) and bias due to genetic trends and non-additive effects. In order to investigate these differences, the following analyses were undertaken on different data sets (see Table 3.1). The pattern of results was similar for each trait, so results are discussed for LMY only:

(i) All the published reports are based on a HYS model with month of calving as a cross-classified effect. The present data (data set 1) were reanalysed using a HYS model. The heritability estimate of LMY was slightly reduced (0.45) due to a relatively increased within sire and reduced between sire components of variance. High heritability estimates using HYMT are expected because HYMT accounted for a greater proportion of TSS compared to HYS.

(ii) Genetic trends may bias the heritability estimate upwards. An analysis of the data by fitting year of birth of young sires as a fixed effect resulted in a small (0.02) reduction in the heritability estimate due to the reduced sire component of variance. Although the young sires were sampled from different years (1980 to 1985), the genetic progress over these years was not sufficient to bias the heritability estimates appreciably.

(iii) In a cross-bred population non-additive effects may be important and neglecting them may bias the heritability upwards (Van der Werf and De Boer, 1989a). Since there is no information on the proportion of North American Holstein in non-pedigree heifers, an analysis of the data on pedigree heifers sired by young sires but

keeping the same data on progeny of old sires as that of data set 1 (data set 3) was performed by including non-additive effects in the model. Various strategies were adopted to account for the proportion of Holstein by regression on: Holstein proportion in sire (original model); Holstein proportion in heifers; Holstein proportion in sire and dam; Holstein proportion in heifers, coefficients of heterosis and recombination loss (non-additive model). All the above strategies adopted to account for non-additive effects had little effect on the heritability estimates (Table 3.9).

(iv) The MMB selected into their DPTS about 120 young bulls each year and progeny tested them over 2000 herds (DPTS herds), whereas NON-DPTS sires either belonged to other AI organisations or were natural service sires, and may be more heterogeneous. Data on offspring of young sires were therefore divided into two parts: DPTS (data set 4) and NON-DPTS (data set 5). Analysis of these two data sets gave a substantiallly lower heritability estimate of LMY for DPTS (Table 3.10) than for NON-DPTS and the original data set. In order to get data on more DPTS sires, records on DPTS herds spanning a calving period of 16 months, rather than 12 months as in previous analyses, from progeny of 56 widely used and 141 young sires (DPTS) were analysed (data set 6). Analysis of this data set revealed that estimates of heritability of LMY, LFY and LPY were substantially lower than, but those for LFC and LPC were similar to, those from the original data set 1 (Table 3.11).

Table 3.9: Variance components and heritability (h^2) of LMY from HYMT model along with different strategies for fitting Holstein proportion (H%) from data set 3.

MODEL	VARIANCE COMPONENTS (kg^2 X100)		h^2
	BET.SIRE	WITHIN	
REGRESSION ON:			
H% IN SIRE(L+Q) (ORIGINAL)	605	4802	0.45
H% IN HEIFER(L+Q)	586	4800	0.44
H% IN HEIFER(L) +HET(L)+REC(L)	591	4799	0.44
H% IN SIRE AND DAM(L+Q)	605	4799	0.45
S.E.	134-137	44-45	0.09

L = Linear; Q = Quadratic; HET = Coefficients of heterosis; REC = Coefficients of recombination loss.

Table 3.10: Variance components and heritability (h^2) of LMY for records on progeny of DPTS (data set 4) and NON-DPTS (data set 5) sires.

DATA SET	VARIANCE COMPONENTS (kg^2 X100)		h^2
	BET.SIRE	WITHIN	
DPTS	441	4832	0.33
NON-DPTS	782	4816	0.56
S.E.	128-143	42-44	0.09

Table 3.11: Variance components and heritability (h^2) estimates of lactation production traits for data set on DPTS herds (data set 6).

TRAITS	VARIANCE COMPONENTS		h^2
	BET.SIRE	WITHIN	
LMY (kg)	**40653**	**499050**	**0.30**
	(11043)	(7787)	(0.08)
LFY (kg)	**42.74**	**815.20**	**0.20**
	(14.93)	(12.66)	(0.07)
LPY (kg)	**32.93**	**434.05**	**0.28**
	(9.38)	(6.75)	(0.08)
LFC (g/kg)	**1.901**	**9.937**	**0.64**
	(0.39)	(0.14)	(0.11)
LPC (g/kg)	**0.255**	**2.208**	**0.41**
	(0.06)	(0.03)	(0.09)

Figures within parenthesis are s.e.

Lower heritability estimates for yield traits from DPTS data sets (data sets 4 and 6) may be due to intense selection of bull sires and bull dams for production traits. The REML analysis assumes that young bulls are a random sample from the population, in fact these DPTS young bulls are progeny of highly selected bulls and dams. An alternate analysis by fitting paternal grandsires (bull sires) as a fixed effect was undertaken and it was found that heritability of LMY was increased (0.37) after multiplying the sire component of variance by 16/3 (i. e. after taking account of 1/16 of the variance removed by paternal grandsires) instead of 4. This analysis assumed that sires are not related, hence this estimate will be increased by 0.02 if relationships among sires are accounted for (it was observed that the heritability estimates of TD and lactation yield of milk are reduced by about 0.02 when relationships among sires are not included).

Assuming that bull dams were selected from the top 5% of the population and selection was for milk yield alone, for a heritability of 0.35 using the prediction given by Bulmer (1971) the additive genetic variance in bull dams is reduced by approximately 0.30 giving a further increase of 0.07 in heritability, thus making this estimate closer to the estimate (0.49) from the original data set.

All the results discussed so far are from the data sets which were edited according to the intervals between calving and first test and between subsequent tests. For more direct comparisons with previous literature estimates, two data sets on complete (305 day) first lactation yields, spanning a calving period of one year each were extracted from NMR files of MMB. The first data set was from November 1987 to October 1988 (data set 7). The second was from November 1988

to October 1989 (data set 8) and was almost identical to the original data set 1 used in this study. Results from separate analyses of data set 7 and 8 and from the combined analysis of both (data set 9) are presented in Table 3.12. Results from data set 8 were, as expected, almost the same as those from the original data set 1 and estimates from the data set 7 were similar to those from the original data set apart from differences explainable by sampling. Differences in the heritability estimates from data sets 7 and 8 were mostly due to differences in sire components of variance as there were only small changes in within sire components. Heritability estimates of LMY and LPY from the combined analysis (data set 9) were slightly lower than the estimates from the original data set but differences could be explained by sampling. From data set 9, data on DPTS herds (data set 10) were extracted and analysed. The heritability estimates for LFY (0.27) and LPY (0.32) were lower than those from data set 9 but the estimate for LMY (0.38) was similar (Table 3.12).

Genetic and phenotypic correlations either among TD of the same trait, among TD of different traits or among lactation traits did not change appreciably among the different data sets and were similar (not shown) to those from data sets 1 and 2.

It may be concluded from this analysis of different data sets that, except for those based on DPTS sires, heritabilities of LMY and LPY are higher than published values. There are several possible factors contributing to higher heritability estimates. A major difference between the data used in the present study and previous studies is that the present data came from one recent year representing about 8000 herds whereas in others data spanned many years representing 100 to 4000 herds. The data used in the present study were on offspring

Table 3.12: Heritability estimates of lactation production traits for data sets 7-10.

TRAITS	DATA SETS			
	7	8	9	10
LMY	0.43	0.48	0.41	0.38
LFY	0.37	0.41	0.35	0.27
LPY	0.36	0.46	0.37	0.32
LFC	0.71	0.57	0.64	0.62
LPC	0.65	0.47	0.50	0.51
S.E.	0.05-0.06	0.05	0.04-0.05	0.05-0.07

of DPTS as well as NON-DPTS sires while other studies used only DPTS data (Hill *et al.*, 1983; Swanson and Gnanasakthy, 1990; Meyer, 1984). Higher heritability estimates in this study are partly due to lower environmental variance due to fitting herd-year-month (of test) rather than, typically, herd-year-season (of calving). Perusal of literature estimates (Meyer, 1984, 1987; Visscher, 1991) and estimates from this study indicate that the genetic CV has increased (from 0.071 to 0.102) from 1972 to 1989. In addition to better estimation procedures used in recent years, this increasing trend in genetic variance also coincides with increases in mean performance, phenotypic variance and CV, and the proportion of Holstein genes in the population. High heritability estimates have been associated with high mean, high variance and CV (Hill *et al.*, 1983), and with a high proportion of Holstein genes in sires (Van der Werf and De Boer, 1989b).

3.4.2 Genetic Parameters of Test Day Records

Heritability estimates of TD for all the traits were higher than those previously reported (see Tables 2.3, 2.8 and 2.11 of Chapter 2) but the pattern across TD was similar to published reports (Van Vleck and Henderson, 1961a; Keown and Van Vleck, 1971; Auran, 1976a; Danell, 1982b; Wilmink, 1987b; Meyer *et al.*, 1989). Lower heritability estimates in the early part of the lactation were due to both a relatively high within sire component and a low between sire component of variance. The estimates of genetic and phenotypic correlations among TD within a trait and TD with corresponding 305-day complete lactation are in good agreement with earlier studies.

Results from data set 1 and 2 were similar for the first 8 TD except for heritabilities of TD1, suggesting that culling has little effect on estimates of these genetic parameters. Small differences (< 2%) in phenotypic SDs of TD1-8 and lactation traits from data sets 1 and 2 also suggest little influence of culling on these parameter. Meyer *et al.* (1989) also observed no change in heritability estimates from successive multivariate analysis of TD records, but Wilmink (1988) observed an increase in the heritability estimates of cumulative and 305-day lactation yields after including the extended incomplete records. In the study reported by Wilmink (1988), there was twice the culling rate (10.0% at day 240 of lactation) observed in the present study, so variances may have been inflated due to the extension of the part records.

Estimates of genetic correlations between different traits within TD (diagonal in Table 3.8) were not higher than others indicating that selection on yield traits of any TD will bring about more or less similar changes throughout lactation in another trait.

From these results it is clear that the heritability estimates for milk yield, fat and protein contents were highest in mid lactation with a similar pattern for fat and protein yields except that estimates in late lacation did not fall. In general, genetic correlations of TD with lactation traits were highest in mid lactation and genetic correlations among TD were highest for adjacent TD.

The lactation yield used for evaluation is not the actual 305-day yield but it is a predicted yield which may be biased in contrast to TD records which are actual yields. Also in theory it would be more accurate to use TD records directly in a selection index rather than

using them as at present through a simple phenotypic index. Prediction of performance in complete lactation from TD yields is a function of their heritabilities and genetic correlations with 305-day lactation. It seems not to be possible to predict the complete lactation with high accuracy from only the earliest tests (TD1 and TD2) as the heritabilities are low and their genetic correlations with complete lactation are also less than 0.90. But from TD3 onwards the genetic correlations with complete lactation are more than 0.95 and their heritabilities are also high in comparison to TD1. Although the accuracy of indirect selection on a few tests is less than direct selection based upon the complete lactation, this loss in accuracy can be compensated by increased selection intensity, so selection on early tests may become more advantageous if reduced generation interval is taken into account.

APPENDIX TO CHAPTER 3

Appendix Table 3.1: Estimates of heritability (X100) (diagonal) genetic (below diagonal) and phenotypic (above diagonal) correlations (X100) among TD and lactation milk yield (LMY).

		TD								
		1	2	3	4	5	6	7	8	LMY
1	TD1	*27*	64	57	52	50	47	45	42	66
2	TD2	92	*33*	71	66	63	59	56	53	77
3	TD3	95	95	*34*	74	70	66	63	59	81
4	TD4	86	86	97	*36*	75	71	68	64	83
5	TD5	87	86	96	99	*35*	75	72	68	84
6	TD6	75	80	91	95	96	*38*	76	71	84
7	TD7	77	83	92	97	98	99	*39*	76	84
8	TD8	73	80	90	96	97	97	99	*43*	83
LMY		87	89	97	98	99	97	98	97	*49*

Range of s.e.(h^2) = 0.05-0.07; Range of s.e.(rp) = 0.003-0.009
Range of s.e.(rg) = 0.01 for rg = 0.99 to 0.08 for rg = 0.73

Appendix Table 3.2: Estimates of heritability (X100) (diagonal) genetic (below diagonal) and phenotypic (above diagonal) correlations (X100) among TD and lactation fat yield (LFY).

		TD								
		1	2	3	4	5	6	7	8	LFY
1	TD1	*16*	47	41	39	38	35	34	32	59
2	TD2	87	*24*	52	48	47	44	42	48	67
3	TD3	86	83	*20*	55	54	51	49	44	71
4	TD4	68	79	94	*18*	60	57	54	52	74
5	TD5	67	83	93	98	*24*	62	60	56	76
6	TD6	57	64	83	92	89	*27*	64	60	77
7	TD7	60	76	88	98	98	95	*25*	65	77
8	TD8	60	77	90	98	99	91	99	*28*	76
LFY		77	85	95	97	97	93	97	96	*39*

Range of s.e.(h^2) = 0.04-0.06; Range of s.e.(rp) = 0.004-0.009
Range of s.e.(rg) = 0.01 for rg = 0.99 to 0.13 for rg = 0.57

Appendix Table 3.3: Estimates of heritability (X100) (diagonal) genetic (below diagonal) and phenotypic (above diagonal) correlations (X100) among TD and lactation protein yield (LPY).

		TD								
		1	2	3	4	5	6	7	8	LPY
1	TD1	*22*	54	47	43	41	38	37	36	60
2	TD2	83	*30*	61	56	53	49	48	46	70
3	TD3	89	83	*28*	65	60	56	54	51	74
4	TD4	81	81	99	*33*	67	62	59	56	77
5	TD5	78	85	94	96	*28*	67	64	59	79
6	TD6	69	80	89	93	96	*30*	68	63	79
7	TD7	75	88	93	96	99	98	*27*	69	79
8	TD8	79	88	96	98	97	93	98	*30*	79
LPY		84	90	97	97	98	95	98	98	*43*

Range of s.e.(h^2) = 0.04-0.06; Range of s.e.(rp) = 0.003-0.008
Range of s.e.(rg) = 0.01 for rg = 0.99 to 0.11 for rg = 0.69

Appendix Table 3.4: Estimates of heritability (X100) (diagonal) genetic (below diagonal) and phenotypic (above diagonal) correlations (X100) among TD and lactation fat content (LFC).

		TD								
		1	2	3	4	5	6	7	8	LFC
1	TD1	*11*	26	23	23	23	22	21	20	47
2	TD2	88	*26*	41	41	39	38	36	36	63
3	TD3	85	92	*32*	50	48	48	46	43	70
4	TD4	76	91	97	*44*	56	55	52	49	74
5	TD5	80	90	96	94	*41*	60	57	53	76
6	TD6	79	84	96	94	97	*40*	61	57	76
7	TD7	67	78	90	92	97	97	*48*	61	75
8	TD8	73	85	92	93	97	98	97	*43*	74
LFC		81	91	98	98	99	98	96	98	*63*

Range of s.e.(h^2) = 0.03-0.07; Range of s.e.(rp) = 0.005-0.009
Range of s.e.(rg) = 0.01 for rg = 0.99 to 0.11 for rg = 0.67

Appendix Table 3.5: Estimates of heritability (X100) (diagonal) genetic (below diagonal) and phenotypic (above diagonal) correlations (X100) among TD and lactation protein content (LPC).

		TD								
		1	2	3	4	5	6	7	8	LPC
1	TD1	*21*	48	41	38	35	34	31	31	56
2	TD2	85	*26*	65	60	56	53	49	47	75
3	TD3	76	90	*38*	71	65	62	58	55	81
4	TD4	78	89	99	*40*	72	69	65	61	83
5	TD5	67	80	91	96	*38*	73	68	63	83
6	TD6	64	78	90	94	98	*38*	73	68	83
7	TD7	58	63	80	85	92	96	*41*	74	81
8	TD8	56	71	84	86	90	97	97	*43*	79
LPC		76	87	95	97	96	98	92	94	*47*

Range of s.e.(h^2) = 0.05-0.06; Range of s.e.(rp) = 0.003-0.010
Range of s.e.(rg) = 0.01 for rg = 0.99 to 0.11 for rg = 0.56

Appendix Table 3.6: Estimates of heritability (X100) of TD9 and TD10 and their genetic and phenotypic correlations (X100) with other TD and lactation traits from data set 2.

		TD										
		LACTATION										
		1	2	3	4	5	6	7	8	9	10	
	h^2	rg/rp+										
		MILK YIELD										
TD9	*36*	65	72	77	90	98	93	96	97	-	94	95
		37	49	54	59	63	65	69	74	-	70	80
TD10	*33*	64	63	70	78	93	83	86	88	94	-	88
		30	40	44	49	53	54	58	63	70	-	69
		FAT YIELD										
TD9	*28*	79	76	70	90	95	88	97	94	-	98	97
		29	37	41	47	52	55	60	64	-	62	74
TD10	*34*	79	75	70	87	96	89	98	95	98	-	98
		25	32	37	41	47	49	53	57	62	-	67
		PROTEIN YIELD										
TD9	*27*	77	73	69	85	98	85	91	95	-	98	95
		32	41	45	50	55	56	62	68	-	65	77
TD10	*33*	75	69	64	79	92	79	84	91	98	-	91
		27	35	39	43	48	48	53	58	65	-	68
		FAT CONTENT										
TD9	*35*	85	89	76	90	85	90	91	91	-	97	94
		19	35	39	45	50	52	55	58	-	54	71
TD10	*24*	75	80	67	89	80	87	93	91	97	-	90
		17	30	34	39	43	44	47	49	54	-	61
		PROTEIN CONTENT										
TD9	*39*	71	59	74	78	90	92	87	94	-	89	89
		27	44	50	56	59	62	65	72	-	69	75
TD10	*25*	68	34	59	66	75	76	79	79	89	-	77
		23	34	39	44	46	48	51	57	69	-	61

+ rg in top row; rp in second row;
Range of s.e.(h^2) = 0.06-0.08; Range of s.e.(rp) = 0.004-0.010
Range of s.e.(rg) = 0.02 for rg = 0.98 to 0.20 for rg = 0.68

CHAPTER 4

THE EFFECT OF INCREASING THE INTERVAL BETWEEN RECORDINGS ON GENETIC PARAMETERS OF TEST DAY RECORDS

4.1 Introduction

At present in the United Kingdom and elsewhere, milk recording is usually carried out at monthly intervals and 305-day lactation yield is calculated by linear interpolation. This has a phenotypic prediction error (SD of difference between predicted and true yield) of about 0.02 to 0.03 of the mean (Anderson *et al.*, 1989). Less frequent recording, for example bi-monthly, is likely to increase the phenotypic prediction error but is cheaper. It would also reduce the number of traits to be included in any multivariate analysis. For genetic evaluation purposes, the genetic correlation of an index of test day yields with total lactation yield is the criterion of importance, and the effect of increased interval on this genetic correlation has not been investigated.

Automatic milk recording devices may become a common feature on farms so that milk yield can be recorded daily and total lactation yield estimated precisely. Lactation fat and protein yields may then be computed more precisely from the sum of daily milk yields and average fat and protein contents, with the possibility of reducing the frequency of testing.

Information is therefore needed on the influence of increased interval between recordings on the variance-covariance structure of test day (TD) records and their correlation with total yield. An investigation was undertaken into the effect of bi-monthly recording

on genetic parameters of resulting TD records and their association with predicted 305-day yields (computed by linear interpolation) from TD records taken at monthly intervals. This was done by analysis of records obtained in alternate monthly herd visits to simulate the case of bi-monthly recording.

4.2 Material and Methods

4.2.1 Data

Data on British Holstein-Friesian heifers, having their first test between November 1988 and October 1989 and at least 10 TD (TD1-TD10) records of milk yield and of fat and protein contents taken at monthly intervals, were utilised for this analysis. Details of these data have been given previously (data set 2 of chapter 3). For these monthly test day (MTD) records, herd visits for each herd were coded from one to the number of times the herd was visited. Since the data were on heifers having their first test from November 1988 onwards, if a recorder visited a herd in November 1988 for the first time, all these MTD records on heifers were assigned to visit 1 and MTD records during subsequent visits at monthly intervals were assigned to visits 2, 3 and so on. For example taking heifer 1 from Table 4.1, records of MTDs 1, 3, 5, 7 and 9 were consigned to the odd data set and MTDs 2, 4, 6, 8 and 10 to the even data set. This resulted in two data sets of odd and even visits, each having 5 bi-monthly test days (BMTD) at approximately bi-monthly intervals instead of 10 MTD records at monthly intervals.

These odd and even BMTD data sets were on 34039 and 34019 daughters, respectively, of the 40 most widely used sires and 705 "young" sires assumed to be randomly sampled and having about one-fifth of the

daughters (chapter 3).

The main differences between the 5 BMTD and 10 MTD data sets are in the interval between calving and first test, the interval between subsequent tests, and the spread of the tests over the lactation. For the 10 MTD data set, the average interval between calving and first test was 20 days and between subsequent tests about 30 days, and each test was spread over a relatively shorter time period, the first test ranging from days 4 to 45 of lactation. For the two newly constituted 5 BMTD data sets, the average interval between calving and first test increased to 34 days and that between subsequent tests to 60 days, with each test ranging over a longer period, the first from days 4 to 79 of lactation.

4.2.2 Statistical Methods

Both BMTD data sets were analysed by Restricted Maximum Likelihood (REML) procedures (Patterson and Thompson, 1971) using REML.PK programs (Meyer, 1986), with the same design and methods used for MTD in chapter 3. The mixed linear model included herd-year-month of first test and pedigree status of the heifer as fixed effects, young sires and paternal grandsires as random effects, and day of lactation for first test (interval between calving and first test), age at calving and proportion of Holstein in the sires as covariates. Preliminary analyses revealed that a quartic regression on day of lactation for first test (DLFT) fitted better than a quadratic, so a quartic regression on DLFT and quadratic on age at calving and proportion of Holstein were fitted. Proven sires were treated as fixed effects in order to improve the connectedness. Relationships among random sires were also included.

Table 4.1: Example of structure of BMTD in relation to herd visits and MTD.

YEAR	1988						1989							
MONTHS	NV.	DC.	JA.	FB.	MR.	AP.	MY.	JN.	JL.	AG.	SP.	OC.	NV.	DC.
HERD VISIT	1	2	3	4	5	6	7	8	9	10	11	12	13	14
HEIFER 1														
MTD	1	2	3	4	5	6	7	8	9	10				
BMTD:														
ODDS	1		2		3		4		5					
EVENS		1		2		3		4		5				
HEIFER 2														
MTD		1	2	3	4	5	6	7	8	9	10			
BMTD:														
ODDS			1		2		3		4		5			
EVENS		1		2		3		4		5				
HEIFER 3														
MTD			1	2	3	4	5	6	7	8	9	10		
BMTD:														
ODDS			1		2		3		4		5			
EVENS				1		2		3		4		5		
HEIFER 4														
MTD				1	2	3	4	5	6	7	8	9	10	
BMTD:														
ODDS					1		2		3		4		5	
EVENS				1		2		3		4		5		

Heritabilities, genetic and phenotypic correlations were estimated from variances and covariances averaged over the two data sets.

4.3 Results

Heritability estimates for BMTD are presented in Table 4.2. Estimates of genetic and phenotypic correlations among BMTD averaged in terms of number of tests apart, and between BMTD and lactation records (LR) are given Tables 4.3 and 4.4. Details of the parameter estimates are given in Appendix Table 4.1.

Heritability estimates of BMTD yields of milk, fat and protein and contents of fat and protein were similar to the average of the estimates for the two corresponding MTD (Table 4.2). Genetic correlations among adjacent BMTD of these traits were 2-3 units (0.02–0.03) and phenotypic correlations were 4-5 units (0.04–0.05) lower than those among adjacent MTD (not shown, but see Appendix Table 4.1 and Appendix Tables 3.1-3.6 of chapter 3), while the genetic and phenotypic correlations among BMTD one test apart were similar to those among MTD two tests apart (Table 4.3) (Note: for MTD two tests apart, for example, phenotypic correlations were computed straightforwardly, but genetic correlations were averaged over the corresponding records included in sire family means, i. e. tests 1 and 2 with 3 and 4). Genetic correlations between BMTD and the corresponding LR were similar and the phenotypic correlations were almost the same as the average of the estimates between the two corresponding MTD and LR (Table 4.4).

Table 4.2: Heritability estimates (X100) of BMTD in relation to average estimates (X100) of the two corresponding MTD.

TD	MILK YIELD (MY)	FAT YIELD (FY)	PROTEIN YIELD (PY)	FAT CONTENT (FC)	PROTEIN CONTENT (PC)
BMTD:					
1	23	15	23	20	19
2	34	23	27	33	43
3	37	25	30	40	36
4	43	30	35	45	44
5	37	29	29	33	27
MEAN	35	24	29	34	34
AVERAGE OF MTD:					
1&2	20	14	20	20	22
3&4	37	16	23	33	43
5&6	41	25	26	43	42
7&8	42	32	34	47	46
9&10	35	31	30	30	32
OVERALL MEAN	35	24	27	35	37

Table 4.3: Average genetic (rg) and phenotypic (rp) correlations (X100) among BMTD in relation to MTD as a function of number of tests apart.

	MY	FY	PY	FC	PC	MY	FY	PY	FC	PC
TESTS APART			rg					rp		
BMTD[+]:										
1	94	94	95	92	91	67	54	59	48	62
2	87	88	88	89	82	59	47	50	41	51
3	82	81	85	79	70	50	39	41	33	42
4	70	69	75	73	60	39	30	34	24	31
MTD[++]:										
2	92	86	87	92	91	67	54	58	48	62
4	88	82	84	86	81	58	46	50	41	52
6	77	76	77	80	68	50	39	42	34	39
8	66	77	74	82	58	39	31	34	25	31

[+] For example, the average genetic and phenotypic correlations among BMTD three tests apart is the mean of the correlations between BMTD 1 and 4, and 2 and 5.

[++] For example, the average genetic correlation among MTD six tests apart is the mean of the correlations between MTD 1 and 7, 2 and 8, 1 and 8, 2 and 7, 3 and 9, 4 and 10, 3 and 10, and 4 and 9; the average phenotypic correlation among MTD six tests apart is the mean of the correlations between MTD 1 and 7, 2 and 8, 3 and 9, and 4 and 10.

Table 4.4: Genetic (rg) and phenotypic (rp) correlations (X100) of BMTD with complete lactation records in relation to average estimates (X100) of the two corresponding MTD with complete lactation records.

	LACTATION RECORDS									
	MY	FY	PY	FC	PC	MY	FY	PY	FC	PC
TD			rg					rp		
BMTD:										
1	88	83	89	88	87	71	63	65	52	63
2	97	97	97	97	94	83	73	77	70	81
3	96	97	96	99	97	85	78	80	76	83
4	97	97	99	92	91	84	77	80	75	81
5	93	97	95	94	89	75	71	73	66	68
AVERAGE OF MTD:										
1&2	86	85	89	91	87	71	63	65	54	64
3&4	96	89	92	94	96	83	73	77	70	82
5&6	97	91	93	97	99	85	77	80	77	84
7&8	97	97	98	95	91	84	78	80	76	81
9&10	92	98	93	92	83	75	71	73	66	68

4.4 Discussion

Increasing the interval between recordings and spread of individual tests over the lactation period resulted in very little change in heritability estimates of BMTD compared to those of MTD for all the traits. Since the BMTD were constructed from MTD and the statistical model used for the analysis of BMTD was the same as that for MTD, a quartic regression on DLFT removed the extra variation induced by the changed data structure of BMTD. The variance-covariance structure of BMTD could therefore be predicted from that of MTD: within sire variances of BMTD were predicted as the mean of the within sire variances of the two corresponding MTD, and between sire variances as the mean of the between sire variances and covariances of the two corresponding MTD. Predicted between sire variances were similar to those estimated by REML, apart from some differences which could be attributed to sampling, and predicted within sire variances were almost the same as estimated by REML (Table 4.5).

The correlation structure of BMTD could be well predicted by noting that BMTD 1, 2, 3 and 4 tests apart correspond to MTD 2, 4, 6 and 8 tests apart (Table 4.3). The similarity of phenotypic correlations between BMTD and LR and between MTD and LR is also a consequence of the efficient correction of BMTD by a quartic regression on DLFT.

From the variance-covariance structure of BMTD, the usefulness of BMTD for the genetic prediction of complete lactation performance can be compared to that of MTD in terms of accuracy, defined as the product of the square root of the heritability of the selection criterion (e. g. sum of BMTD or MTD records) and the genetic correlation between the selection criterion and complete lactation performance. The accuracy of prediction of breeding value for 305-day

lactation yields of milk, fat and protein using the sum of 5 BMTD would be 0.68, 0.61 and 0.66 compared to 0.70, 0.62 and 0.67 using the sum of 10 MTD (Table 4.6). The accuracies are high because the BMTD and MTD records had high heritability estimates in these data (chapter 3), but these would not affect the comparisons. The accuracies using the sum of the first 5 MTD would be substantially lower than the sum of 5 BMTD (Table 4.6) as heritabilities are lower in the first half of the lactation, particularly for fat and protein yields, and the genetic correlation is less than unity between the first half and the complete lactation.

When records on daily milk yield are available, an obvious way of predicting the phenotype for 305-day lactation fat yield (LFY) is from the product of the sum of daily milk yields and the mean fat content for the days tested (corrected for stage of lactation). The accuracy of this estimator can be evaluated in terms of the phenotypic correlation with LFY. Since records on daily milk yield and true LFY were not available, as an example the correlation was computed between the product of the sum of 10 MTD milk yields and the mean of 5 alternate MTD fat contents ($\Sigma MY_i(\Sigma FC_j)/5$; i=1,10 and j=1,3,..,9 or 2,4,..,10) and the sum of 10 MTD fat yields (ΣMY_iFC_i; i=1,10). The phenotypic correlation was calculated using a Taylor series approximation (Kendall and Stuart, 1963) ignoring terms of higher than second order. The estimate was 0.980 which is higher than the phenotypic correlation between the sum of 5 alternate MTD fat yields and the sum of 10 MTD fat yield (0.956) and near to the phenotypic correlation between the sum of 10 MTD fat yields and 305-day lactation fat yield (0.996). This suggests that this method

Table 4.5: Estimated (ES) and predicted (PR) components of variances for BMTD of different traits+.

VARIANCE COMPONENTS		BMTD 1	2	3	4	5
		MILK YIELD (kg)				
BET. SIRE	ES	0.559	0.764	0.752	0.852	0.769
	PR	0.563	0.797	0.799	0.807	0.698
WITHIN	ES	9.040	8.279	7.344	7.055	7.612
	PR	8.931	8.249	7.291	7.003	7.633
		FAT YIELD (kgX10)				
BET. SIRE	ES	0.0876	0.0966	0.0921	0.1095	0.1119
	PR	0.0692	0.0614	0.0771	0.1145	0.1186
WITHIN	ES	2.193	1.618	1.354	1.350	1.443
	PR	2.145	1.611	1.349	1.339	1.447
		PROTEIN YIELD (kgX10)				
BET. SIRE	ϵs	0.0526	0.0543	0.0575	0.0679	0.0646
	PR	0.0416	0.0451	0.0436	0.0652	0.0663
WITHIN	ϵs	0.8590	0.7631	0.7109	0.7187	0.8207
	PR	0.8481	0.7626	0.7056	0.7153	0.8237
		FAT CONTENT (g/kg)				
BET. SIRE	ES	0.1322	0.1803	0.2269	0.2697	0.2476
	PR	0.1055	0.1637	0.2337	0.2742	0.2125
WITHIN	ES	2.521	2.000	2.027	2.139	2.769
	PR	2.479	1.999	2.012	2.129	2.786
		PROTEIN CONTENT (g/kg)				
BET. SIRE	ES	0.0197	0.0420	0.0357	0.0440	0.0401
	PR	0.0183	0.0414	0.0418	0.0443	0.0440
WITHIN	ES	0.3868	0.3481	0.3605	0.3562	0.5593
	PR	0.3856	0.3478	0.3610	0.3535	0.5606

+ Between sire variances were predicted as the mean of the between sire variances of the two corresponding MTD and the covariance between them, and within sire variances were predicted as the mean of the within sire variances of the two corresponding MTD.

Table 4.6: Accuracy of prediction of breeding value for 305-day lactation production from TD records.

SELECTION CRITERION	ACCURACY		
	MY	FY	PY
Sum of 10 MTD or phenotype for predicted 305-day yields:	0.70	0.62	0.67
Sum of 5 BMTD:	0.68	0.61	0.66
Sum of first 5 MTD:	0.63	0.46	0.55

of prediction of lactation fat yield or protein yield could utilize extra information on daily milk yield records. The genetic prediction of 305-day lactation fat yield or protein yield by the proposed method would be more accurate than the phenotypic because genetic correlations between TD and LR are higher than phenotypic correlations.

The slightly lower accuracy for the sum of 5 BMTD records of milk yield compared to 10 MTD records is due to the lower heritability estimate for the sum of 5 BMTD (0.47 Vs. 0.49) and a slightly lower genetic correlation of the sum of 5 BMTD with lactation production (0.990 Vs. 0.999). The accuracy of using the sum of 7 tests with an approximate interval of 6 weeks, estimated from the predicted variance-covariance structure of 7 tests from that of 10 MTD, would be at least 0.99 of the accuracy of the sum of 10 MTD. From these results it is clear that reducing the frequency of testing to less than monthly would cut costs without proportional loss in accuracy.

APPENDIX TO CHAPTER 4

Appendix Table 4.1: Estimates of heritability (X100) (diagonal), genetic correlations (below diagonal) and phenotypic correlations (above diagonal) (X100) among BMTD and lactation records (LR).

	BMTD					
	1	2	3	4	5	LR
			MILK YIELD			
1 BMTD1	*23*	60	53	48	39	71
2 BMTD2	94	*34*	69	63	52	83
3 BMTD3	77	91	*37*	72	60	85
4 BMTD4	78	92	97	*43*	66	84
5 BMTD5	70	85	93	93	*37*	75
LR	88	97	96	97	93	*48*
			FAT YIELD			
1 BMTD1	*15*	43	40	36	30	63
2 BMTD2	88	*23*	54	50	42	73
3 BMTD3	73	92	*25*	60	51	78
4 BMTD4	72	91	97	*30*	58	77
5 BMTD5	69	90	99	98	*29*	71
LR	83	97	97	97	97	*39*
			PROTEIN YIELD			
1 BMTD1	*23*	50	44	40	34	65
2 BMTD2	96	*27*	59	54	42	77
3 BMTD3	76	88	*30*	64	52	80
4 BMTD4	82	93	98	*35*	61	80
5 BMTD5	75	87	96	97	*29*	73
LR	89	97	96	99	95	*45*
			FAT CONTENT			
1 BMTD1	*20*	30	29	27	24	52
2 BMTD2	88	*33*	51	47	39	70
3 BMTD3	84	97	*40*	58	47	76
4 BMTD4	72	89	93	*45*	52	75
5 BMTD5	73	86	95	89	*33*	66
LR	88	97	99	92	94	*62*
			PROTEIN CONTENT			
1 BMTD1	*19*	49	42	38	31	63
2 BMTD2	89	*43*	67	59	46	81
3 BMTD3	78	90	*36*	69	53	83
4 BMTD4	65	77	92	*44*	61	81
5 BMTD5	60	74	91	92	*27*	68
LR	87	94	97	91	89	*50*

CHAPTER 5

PHENOTYPIC VARIANCES AND CORRELATIONS OF DAILY YIELDS

5.1 Introduction

305-day yield is usually predicted from test day (TD) yields. Therefore the phenotypic correlations between TD records and predicted lactation milk yield (LMY) may be different from the phenotypic correlations between TD records and actual LMY. Further, the first TD record is taken, on average, at day 20 of lactation. The phenotypic variance of daily yields in the early part of lactation is not precisely known. The phenotypic correlations among daily yields on successive days of lactation less than 30 days apart over the whole lactation have not been found in the literature.

In the present study the extent of bias in phenotypic correlations between TD milk yields and predicted LMY, variances of daily milk yields and phenotypic correlations between daily milk yields on successive days of lactation are investigated.

5.2 Material and Methods

5.2.1 Data

Daily milk yield records on British Holstein-Friesian heifers that calved between September 1989 and March 1991 and were maintained at Langhill farm of the University of Edinburgh and Scottish Agricultural College were used for this analysis. The number of heifers in this data set varied between 14 and 39, being least at the beginning and the end of the lactation.

5.2.2 Statistical Methods

The traits analysed were daily milk yields for each day of lactation on all available heifers. These daily milk yields on successive days of lactation were analysed together in blocks of 35 days of lactation at a time except at the beginning and the end of lactation where numbers of heifers were small. So as to avoid missing records and utilise data on all animals, smaller blocks were used at the beginning and the end of lactation. Least squares analyses (Harvey, 1977) were carried out using the following fixed effects model:

$$Y_{jkl} = \mu + R_j + bA_k + e_{jkl}$$

Where,

Y_{jkl} = daily milk yield;

μ = overall mean;

R_j = year-month of recording for the first daily yield in a block;

b = linear regression on age at calving;

A_k = age at calving in months;

e_{jkl} = random error.

Phenotypic correlations among daily yields i days apart (i = 1, ..., 34) were averaged within a block. In order to increase the number of pairs for correlations apart by 10 to 30 days, averaged phenotypic correlations within blocks of 35 and 18 days, avoiding the pairs

already included, were also calculated.

Phenotypic correlations between TD yields and predicted LMY may be biased upward. If actual LMY (sum of daily yields) is known, the phenotypic correlations between TD yields and actual LMY are free from this bias, but the phenotypic correlations between TD milk yields and actual LMY could not be computed directly because daily yields for the whole of the lactation were known only for a few heifers. In order to calculate the approximate bias in these correlations three indirect methods were considered: (1) The phenotypic correlations between TD yields and actual LMY (sum of daily yields) were predicted using estimated within block variances and covariances from Langhill data and predicted covariances between blocks (using phenotypic correlations among TD records from chapter 3 and variances from Langhill data). The bias was estimated as the difference between phenotypic correlations of TD yields with predicted and actual LMY. (2) The phenotypic correlations between daily milk yields and the sum of 30 daily yields within blocks taking 10 blocks of 30 days each were computed from Langhill data and averaged. Considering predicted LMY as the sum of 10 blocks of approximately 30 days with each having 30 replicates of each TD yield. The phenotypic correlation of each TD yield with the sum of 30 days yield within each block is then 1.0. The approximate upper limit for bias in the phenotypic correlations between TD yields and predicted LMY was estimated as the deviation of phenotypic correlation between daily yields and the sum of 30 days yield from 1.0. (3) The phenotypic correlations between TD yields and LMY (assuming LMY as the sum of 10 TD yields) with and without including the TD yield in question were computed using phenotypic variances and

covariances from chapter 3. An approximate upper limit for bias in phenotypic correlations between TD yields and predicted LMY was calculated as the difference between phenotypic correlations of TD yields with the sum of 10 TD yields with and without including the TD yield in question.

5.3 Results

Residual standard devations (RSDs) of daily yields from day 6 to 305 of lactation ranged from 3.41 to 7.44 kg with an average value of 4.40 kg (see Table 5.1). These RSDs were similar throughout the lactation with a tendency to decrease slightly at the end.

Average phenotypic correlations (weighted average, weighted by the product of the number of heifers and number of pairs) between daily yields apart by 1 to 30 days ranged from 0.84 to 0.75 (Table 5.2). Correlations were lower at the beginning and the end of lactation than in mid lactation.

5.4 Discussion

RSDs of daily yields were similar throughout the lactation with a tendency to decrease at the end, but they were higher than the estimates for TD yields presented in chapter 3. Higher RSDs in this data set could be due to a higher level of production because high variances have been found to be associated with high mean (Hill *et al.*, 1983).

Estimates of phenotypic correlations between adjacent daily yields were high (0.84) and the correlations decreased as the interval between daily yields increased. The phenotypic correlations between daily yields were slightly lower at the beginning and the end of

Table 5.1: Examples of least-squares mean, residual standard deviation (RSD) and coefficient of variation (CV) for daily milk yields.

DAY OF LACTATION	MEAN (kg)	RSD (kg)	CV
6	18.7	6.7	0.37
10	22.6	4.3	0.20
15	24.1	3.9	0.17
20	25.0	3.1	0.13
25	26.1	4.9	0.19
30	25.7	4.8	0.19
40	23.3	6.0	0.25
50	25.6	4.6	0.18
60	25.4	5.7	0.23
80	26.7	5.9	0.22
100	24.3	6.3	0.26
120	23.4	6.5	0.28
140	22.9	3.6	0.17
160	20.1	3.9	0.19
180	18.4	4.9	0.26
200	17.1	6.1	0.34
220	17.3	4.6	0.28
240	17.0	5.0	0.31
260	20.0	4.9	0.29
270	17.9	4.6	0.29
280	15.4	4.1	0.27
285	15.8	4.1	0.27
290	15.9	4.0	0.26
295	15.3	3.2	0.21
300	15.7	3.8	0.25

Table 5.2: Average phenotypic correlations (X100) between daily milk yields as a function of number of days apart.

DAYS APART		1	2	3	5	10	15	20	25	30
BLOCKS OF LACTATION DAYS	NUMBER OF HEIFERS	PHENOTYPIC CORRELATIONS								
SET 1										
6-13	18	57	50	53	55	-	-	-	-	-
+		(07)	(06)	(05)	(03)					
13-30	16	78	79	73	75	67	65	-	-	-
		(17)	(16)	(15)	(13)	(08)	(03)			
30-64	14	87	87	88	86	84	85	86	87	86
64-98	23	84	83	82	82	80	82	77	81	84
98-132	27	83	83	83	82	79	80	79	80	75
132-166	29	83	82	80	80	80	76	83	80	80
166-200	28	86	86	85	84	81	80	81	81	80
200-234	32	87	88	87	85	81	80	77	80	78
234-268	22	88	88	85	82	78	72	72	70	73
268-302	14	82	79	73	68	61	50	53	43	29
		(34)	(33)	(32)	(30)	(25)	(20)	(15)	(10)	(05)
SET 2										
45-79	18	-	-	-	-	-	-	76	73	80
79-113	27	-	-	-	-	-	-	81	81	88
113-147	25	-	-	-	-	-	-	81	76	72
147-181	30	-	-	-	-	-	-	77	75	74
181-215	34	-	-	-	-	-	-	76	73	67
215-249	29	-	-	-	-	-	-	74	72	77
249-283	16	-	-	-	-	-	-	66	71	65
								(15)	(10)	(05)
SET 3										
21-39	22	-	-	-	-	82	80	-	-	-
55-73	27	-	-	-	-	79	73	-	-	-
89-107	39	-	-	-	-	83	82	-	-	-
123-141	32	-	-	-	-	88	90	-	-	-
157-175	37	-	-	-	-	82	84	-	-	-
191-209	39	-	-	-	-	87	86	-	-	-
225-243	37	-	-	-	-	77	76	-	-	-
259-277	26	-	-	-	-	69	75	-	-	-
						(09)	(04)			
WEIGHTED AVERAGE*		84	84	82	82	79	78	77	76	75

Figures within parentheses are number of pairs.

\+ Number of pairs are same above the block unless specified.

* Weights are product of number of heifers and number of pairs.

lactation as were the phenotypic correlations between TD yields reported in chapter 3. It could be due to differential rates of increase in daily yields in the early part of the lactation and differential rates of decrease at the end of the lactation.

It is not easy to estimate algebraically the SE of these correlation coefficients which have been averaged over pairs of records on the same set of heifers. SEs were therefore estimated by simulation. Replicate sets of thirty multivariate normal random deviates were simulated using estimated variances and covariances from the Langhill data. The number of heifers were taken as 9 (minimum number of heifers in this data less degrees of freedom for fixed effects) and 29 (maximum number of heifers in this data less degrees of freedom for fixed effects). The phenotypic correlation coefficients i days apart (i = 1, ..., 30) were estimated and averaged over the number of pairs. The standard deviation of 1000 replicates was taken as the SE of correlation coefficients. The SE of a correlation coefficient of 0.84 based on 29 pairs was 0.038 and 0.084 for 29 and 9 heifers, respectively. Corresponding figures of SE for a correlation coefficient of 0.75 based on single pairs were 0.089 and 0.191 for 29 and 9 heifers, respectively.

In the special case where all the correlations (r) are equal, the variance of correlation coefficients (v(r)) based on d degrees of freedom and m pairs can be calculated approximately using statistical differentiation (W. G. Hill, personal communication) to give:

$$V(r) = [\ (1 - r)^2/d\ ]\ [\ r^2(2 - 2/m + 1/m^2) + 2r(2/m - 1/m^2) + 1/m\ ]$$

The SEs of correlation coefficients using this equation were almost

equal to those obtained by simulation (Appendix Table 5.1).

The approximate upper limit for bias in phenotypic correlations between TD yields and predicted LMY, calculated as the deviation of the average phenotypic correlation between daily yields and the sum of 30 days yield from 1.0 (method 2), ranged from 0.07 to 0.13 (Table 5.3). These average correlations between daily yields and the sum of 30 days yield also indicate that the accuracy of phenotypic prediction of total yield in a month from any one daily yield would be only 87 to 93 percent. The genetic correlation between daily yields and the sum of 30 days yield would be almost 100 percent, because genetic correlations between adjacent TD yields are usually more than 0.95, indicating that the genetic correlation is essentially 1.0 between adjacent daily yields. Method 3 gave slightly lower figures for the upper limit of bias compared to method 2 (Table 5.4). Predicted phenotypic correlations between TD yields and actual LMY (method 1) indicated much lower values of bias than methods 2 and 3. The phenotypic correlation between the sum of 10 TD yields and actual LMY was 0.98, compared to a phenotypic correlation of 0.996 between the sum of 10 TD yields and predicted LMY, indicating very little bias in phenotypic correlations between TD yields and LMY (Table 5.4).

These phenotypic correlations between TD yields and actual LMY were predicted indirectly. More research is needed to estimate these correlations using actual LMY (sum of daily yields) on a large data set.

Table 5.3: Average phenotypic correlations (X100) of daily milk yield with the sum of 30 days yield within a block, taking 10 blocks of 30 days each.+

DAYS OF LACTATION (BLOCKS)	AVERAGE PHENOTYPIC CORRELATIONS
6-35	88
36-65	89
66-95	89
96-125	93
126-155	90
156-185	88
186-215	91
216-245	90
246-275	87
276-305	93

+ Method 2, see Page 86.

Table 5.4: Phenotypic correlations (X100) of TD milk yields with LMY1 (sum of 10 TD yields), LMY2 (sum of 10 TD yields minus TD yield in question) and LMY3 (sum of daily yields).†

TDs	PHENOTYPIC CORRELATIONS WITH		
	LMY1	LMY2	LMY3
1	65	56	63
2	77	71	77
3	82	77	81
4	84	80	85
5	85	82	85
6	85	81	82
7	84	81	84
8	83	79	81
9	80	75	78
10	69	63	68
SUM OF 10 TDs	100	-	98

† Method 1 (LMY3) and Method 3 (LMY1 and LMY2), see Page 86.

APPENDIX TO CHAPTER 5

Appendix Table 5.1: The SE of correlation coefficients using the simulation and the algebraic expression (all the correlations are equal; r = 0.8 and N = 29)

NUMBER OF PAIRS	SE	
	SIMULATION	ALGEBRAIC
1	0.07010	0.07190
5	0.05329	0.05400
10	0.04869	0.04996
15	0.04843	0.04847
20	0.04733	0.04769
25	0.04689	0.04721
29	0.04722	0.04695

CHAPTER 6

PREDICTION OF LACTATION PERFORMANCE FROM INCOMPLETE RECORDS

6.1 Introduction

Current genetic evaluation for lactation production traits in the U. K. and elsewhere, are based on phenotypes for lactation records predicted from test day records. Thus prediction of breeding value for lactation production traits is a two step procedure. Test day (TD) records are the actual measurements on the cows and hence the most accurate method for genetic evaluation would be multivariate BLUP including TD records, but it is computationally not (yet) feasible for national evaluation with a large data set. TD records in progress could easily be handled in this framework without any projection.

In the U. S. A., records of less than 305 days are included in genetic evaluation after projecting them to 305 days. Records in progress (RIP) / part records (records of less than 305 days) are projected by a method based on the regression of yield in the remaining part of the lactation on the last test day yield (Wiggans and Dickinson, 1985). Herd average is also included in the projection for records of less than 155 days. The projected records have less variance than the completed records simply because predictors have less variance than the variables they predict. In order to include the projected part records along with the completed records in animal model evaluations, the genetic variance of projected records should be equal to the genetic variance of the completed records or the mixed model equations have to be modified as proposed by Weller

(1988). VanRaden *et al.* (1991) argued that the techniques of Weller (1988) are computationally demanding and can not be applied directly. They proposed expansion of projected records to equate the genetic variance of projected records to the genetic variance of completed records. In addition to equality of variances, genetic correlations between projected and completed records should be unity. The expansion factors were the ratio of the sire SD of completed records to the sire SD of projected records.

As the expanded records have higher error variance than the completed records, less weight is given to expanded records in animal model evaluations. Since permanent environmental effects are included in the USDA animal model, weights (lactation length weights) for expanded records are the ratio of temporary environmental variance of completed records to the temporary environmental variance of expanded records.

In the U. K., the animal model has recently been introduced for the prediction of breeding values of lactation records. At present, records of less than 200 days are not eligible for inclusion in genetic evaluations. The sire proofs from such records may be biased. The MMB are using a software package developed in U. S. A. for animal model evaluations. Therefore, there is a need to develop the procedures to include the records in progress in U. K. animal model evaluations.

In this chapter accuracy of genetic prediction of lactation performance from the indices of TD records and loss in accuracy assuming a repeatability model are investigated. A second objective is to investigate the procedures to include records in progress in genetic evaluations.

6.2 Material and Methods

6.2.1 Data

Estimates of genetic and phenotypic variances and covariances from data set 2 (chapter 3) were used for subsequent calculations. Details of these estimates for milk, fat and protein yields are presented in Appendix Tables 6.1, 6.2 and 6.3, respectively. Although the 10 X 10 genetic covariance matrices of TD records of milk, fat and protein were negative definite, the smallest eigen values were only -0.0002, -0.00001 and -0.00001, respectively, so no attempt was made to make these matrices positive semi-definite.

6.2.2 Statistical Methods

6.2.2.1 Genetic Prediction of Lactation Performance

For genetic prediction of lactation performance from TD records, genetic selection indices (assuming fixed effects are known) were constructed in the following way:

$$I = b'X$$

$$b = P^{-1} Cov_g$$

$$r_{IH} = \sigma_I / \sigma_H$$

Where,

X = n x 1 vector of deviations from the mean of phenotypic values of TD records;

I = index value to predict H;

H = aggregate breeding value for lactation yields;

b = n x 1 vector of index weights;

P = n x n phenotypic covariance matrix among TD records;

Cov_g = n x 1 genetic covariance matrix between TD and lactation records;

r_{IH} = accuracy of the index;

σ_I = phenotypic SD of the index;

σ_H = genetic SD of lactation records.

Genetic prediction of lactation performance was also attempted assuming a repeatability model (assuming TD records as repeat records with equal variance and a genetic correlation of unity) and the loss in accuracy compared to a genetic index was calculated.

6.2.2.2 Inclusion of Records in Progress in Genetic Evaluation

Inclusion of records in progress (RIP) / part records in genetic evaluation was addressed on a TD basis assuming a repeatability model (model 1) without projection and also by predicting the phenotype (deviation) of part and completed TD records using the phenotypic index of part or completed TD records (model 2).

Assuming a repeatability model, the complete records (CR) were assumed to be the average phenotype of 10 successive TD records and part records (PR) were assumed to be the mean phenotype of the first i TD records for each value of i from 1 to 9. The current animal model in the U. S. A., subsequently adopted in the U. K., is set up such that all the records (part or complete) should have equal genetic variance, but it can handle records of unequal error variances by differential weighting of records according to their lactation lengths. Solutions for animal, permanent environmental and herd-sire

interaction effects are obtained as the weighted sum of the differences between lactation records of a cow and the other effects in the model over all lactations divided by the sum of weights plus appropriate variance ratio. Weights (lactation length weights) depend on the lactation length of the record and parity.

Genetic variances of PR are not equal to the genetic variance of CR, partly because of the rather low genetic variance of TD1. Deviations (from the mean) of PR are multiplied by the actual expansion factors (X) in order to equate the genetic variance of PR to the genetic variance of CR under the assumption that variances of all other random effects except error are also equalised. Equal genetic variance does not mean that the genetic effects in PR and CR are similar unless the genetic correlation between PR and CR is unity. If the genetic correlation between PR and CR is less than unity, PR should be considered as a separate trait or may be excluded from the data.

As expanded part records (EPR) have higher error variance than the completed records, less weight is given to EPR in animal model evaluations. Since the permanent environmental effects are included in the animal model, weights are calculated as the ratio of temporary environmental variance (VES) of CR to the temporary environmental variance of EPR.

Following VanRaden *et al.* (1991), VES of CR was calculated as:

$$VES(CR) = (1 - r) [VP(CR)]$$

assuming a repeatability model.

VES of EPR was calculated as:

$$VES(EPR) = X^2 [VP(PR)] - r [VP(CR)]$$

under the assumption of equality of genetic, permanent environmental and herd-sire interaction effects for CR and PR.

Where,

X = actual expansion factor (ratio of genetic SD of completed records to the genetic SD of part records;

r = repeatability of test day records.

The repeatability (r) (average of phenotypic correlations among 10 TD records) estimates for milk, fat and protein yields were 0.59, 0.47 and 0.51, respectively (from data set 2 of chapter 3).

Inclusion of RIP in genetic evaluation was also attempted by predicting the phenotype (deviation) from part and completed TD records using phenotypic indices (model 2).

Records in progress are projected to 305 day lactation using the regression of yield in the remaining part of the lactation on the last TD yield (Wiggans and Dickinson, 1985) for sire evaluation in U. S. A. A phenotypic index of part or completed TD records would theoretically result in the most accurate prediction of LR. The phenotypic selection indices were constructed in a similar way to the genetic indices but using the phenotypic covariance matrix between TD and LR and phenotypic SD of LR instead of the genetic covariance matrix and genetic SD. The phenotypic selection indices actually predict the deviations from the mean lactation yield.

The predicted phenotype (phenotypic deviation) of part TD records (less than 10 TD records) (PPP) would have lower variances than the

predicted phenotype (phenotypic deviation) of completed TD records (10 TD records) (PPC). These deviations of PPP from the mean are multiplied by the actual expansion factors (ratio of the genetic SD of PPC to the genetic SD of PPP) in order to equate the genetic variance of PPP to genetic variance of PPC.

The theoretical expansion factors (ratio of phenotypic variance of PPC to PPP or reciprocal of the squared phenotypic correlation between PPP and PPC) were also calculated for comparison.

Since the expanded PPP (EPPP) have higher error variance than PPC, EPPP are given less weight in animal model evaluation. The weights (lactation length weights) were calculated in a similar way to those for model 1.

6.3 Results

6.3.1 Genetic Prediction of Lactation Performance

Genetic selection indices were constructed by including successive TD records. Index weights and accuracies of selection for genetic prediction of LR are presented in Table 6.1.

The accuracy of the index I10 (index of 10 TD records) for *milk yield* was 0.71 (Table 6.1). The accuracy of the index I5 (index of first 5 TD records) was 5.63% lower than the accuracy of I10. There was very little increase (0.53%) in the accuracy of the index when TD9 and TD10 were added to the index of the first 8 TD records (I8).

The accuracy of the index I10 for *fat yield* was 0.66 (Table 6.1). The index I5 has much lower accuracy (only 74.24% of index I10) than the index I10 and there was a gradual increase in the accuracy, in contrast to milk yield, when later tests were added to the index.

Table 6.1: Index weights (X10) and accuracy of genetic indices of TD records of milk yield (MY), fat yield (FY) and protein yield (PY) for prediction of breeding value for lactation records.

INDICES	INDEX WEIGHTS FOR TD RECORDS										ACCURACY	
	1	2	3	4	5	6	7	8	9	10	1	2
I1	587										0.36	-
	280										0.26	-
	518										0.39	-
I2	164	699									0.49	0.49
	136	366									0.37	0.35
	349	363									0.44	0.39
I3	461	320	647								0.57	0.57
	92	271	248								0.40	0.40
	282	176	378								0.49	0.46
I4	7	153	309	680							0.63	0.63
	62	210	143	303							0.44	0.44
	250	91	207	383							0.53	0.50
I5	-16	78	175	384	635						0.67	0.67
	38	162	69	168	384						0.49	0.49
	230	47	141	238	347						0.55	0.53
I6	-23	61	123	298	460	399					0.68	0.68
	26	141	23	89	250	374					0.52	0.52
	221	36	110	188	247	247					0.56	0.55
I7	-23	50	101	252	373	244	371				0.69	0.69
	16	119	-10	27	131	187	512				0.58	0.58
	216	121	76	133	140	66	466				0.60	0.59
I8	-24	41	96	213	314	157	185	440			0.71	0.71
	13	107	-24	-67	78	91	356	421			0.61	0.61
	203	-7	59	85	75	-35	250	537			0.65	0.64
I9	-24	40	95	208	304	147	162	389	118		0.71	0.71
	12	100	-28	-18	58	66	308	342	231		0.63	0.63
	202	-10	57	81	66	-42	231	494	104		0.65	0.64
I10	-23	42	94	210	297	143	152	363	40	147	0.71	0.71
	13	96	-35	-24	36	38	266	269	84	386	0.66	0.66
	203	-8	31	87	50	-52	209	437	-34	293	0.67	0.66

Top row: MY; middle row: FY; bottom row: PY
I1: TD1; I2: TD1 and 2; I3: TD1-3, and so on.
Accuracy 1: indices with all successive TD records.
Accuracy 2: indices excluding TD1.

The accuracy of the index I10 for *protein yield* was similar (0.67) to that for fat yield but the accuracy of index I5 for protein yield was higher than for fat yield (Table 6.1). A gradual increase in the accuracy for protein yield was also observed when later tests were added to the index. Invariably the last test recieved the largest weight in the index, particularly when first 8 or less TD records were included in the index. The index weights were different for various TD records and they were even negative (for TD1 of milk yield; TD3 and TD4 of fat yield; TD 2, 6 and 9 of protein yield) when later tests (TD5-TD10) were added to the index (Table 6.1).

The first test is more variable due to rapid changes in milk yield in the initial part of the lactation. The genetic variance of the first TD record and its genetic correlation with LR are rather lower than the genetic variances of the remaining TD records and their genetic correlations with LR. When the first test was dropped from the indices there was very little change in the accuracy of all the indices for milk yield and fat yield, but the accuracy for protein yield was slightly reduced for the indices incorporating TD2-TD5 (Table 6.1). Some of the index weights for TD records of fat and protein were still negative (not shown).

Genetic prediction of lactation performance was also attempted assuming a repeatability model including successive TD records, i. e. TD1, average of TD1 and TD2, average of TD1-3, and so on. Estimates of variance, heritability and accuracy for these successive TD records and their genetic and phenotypic correlations with LR for milk, fat and protein yields are given in Table 6.2.

Table 6.2: Estimates of additive variance (VA), phenotypic variance (VP), heritability and accuracy of successive TD records and their genetic (rg) and phenotypic (rp) correlations with lactation records assuming a repeatability model.

SUCCESSIVE TD RECORDS	VA	VP	h^2	rg	rp	ACCURACY	
						1	2
R1	1.888	9.710	19	83	65	36	-
	0.2147	2.5001	09	88	59	26	-
	0.1895	0.9913	19	90	60	39	-
R2	2.250	7.707	29	88	79	48	49
	0.2768	1.6013	17	85	74	35	35
	0.1662	0.6761	25	90	75	44	39
R3	2.496	7.079	35	91	86	54	56
	0.2578	1.2950	20	88	82	39	40
	0.1678	0.5880	29	91	83	49	45
R4	2.586	6.600	39	95	91	59	61
	0.2394	1.1294	21	93	88	43	43
	0.1632	0.5427	30	95	88	52	50
R5	2.644	6.359	42	98	94	63	65
	0.2381	1.0322	23	96	92	46	47
	0.1578	0.5135	31	98	92	54	53
R6	2.675	6.113	44	98	96	65	67
	0.2400	0.9684	25	98	95	49	50
	0.1568	0.4924	32	99	95	56	54
R7	2.695	5.924	45	99	97	67	68
	0.2605	0.9305	28	99	97	52	54
	0.1651	0.4810	34	99	96	58	57
R8	2.731	5.766	47	100	98	69	70
	0.2797	0.9046	31	100	98	55	57
	0.1772	0.4740	37	99	98	61	60
R9	2.693	5.597	48	100	99	69	70
	0.2897	0.8793	33	100	99	57	58
	0.1782	0.4662	38	100	99	62	61
R10	2.643	5.420	49	100	100	70	71
	0.3100	0.8594	36	100	100	60	61
	0.1853	0.4609	40	100	100	63	63

Estimates of h^2, rg, rp and accuracy are multilpied by 100 ; those of VA and VP for fat and protein yields by 10^4.
Top row: MY; middle row: FY; bottom row: PY.

R1: TD1; R2: average of TD1 and 2; R3: average of TD1-3, and so on.
Accuracy 1: including all successive TD records.
Accuracy 2: excluding TD1.

An important issue is to investigate the loss in accuracy when repeatability model is assumed instead of a genetic index. The maximum loss in accuracy was about 5-7% for milk yield and protein yield and about 9-12% for fat yield (Table 6.2). When first test was excluded the maximum loss in accuracy was about only 3% for milk yield and 5-8% for fat and protein yields (Table 6.2).

6.3.2 Inclusion of Records in Progress in Genetic Evaluation

6.3.2.1 Model 1

Inclusion of RIP in genetic evaluation was investigated on a TD basis assuming a repeatability model.

The accuracy (phenotypic correlation between PR / CR and predicted 305-day records(LR)) of average phenotype of TD records for milk, fat and protein yields are presented in Table 6.2. The accuracy of CR (average phenotype of 10 TD records) for milk, fat and protein yields was almost 100%. These accuracies were more than 90% for PR having at least the first 5 TD records.

The actual expansion factors for PR of milk, fat and protein yields were close to 1.0 (Table 6.3). The actual expansion factors for PR of milk, fat and protein yields ranged from 0.98 to 1.18, 1.03 to 1.20 and 0.99 to 1.09, respectively. The weights given to EPR of milk yield ranged from 0.21 (for EPR having only TD1) to 0.97 (for EPR having the first 9 TD records) (Table 6.3). Corresponding weights for EPR of fat and protein yields ranged from 0.14 to 0.85 and 0.31 to 0.90, respectively.

Table 6.3: Actual expansion factors and lactation length weights for part and completed TD records assuming a repeatability model.

TD RECORDS	ACTUAL EXPANSION FACTORS			LACTATION LENGTH WEIGHTS		
	MY	FY	PY	MY	FY	PY
R1	1.18	1.20	0.99	0.21	0.14	0.31
R2	1.08	1.06	1.06	0.38	0.33	0.44
R3	1.03	1.10	1.05	0.52	0.40	0.54
R4	1.01	1.14	1.07	0.62	0.43	0.59
R5	1.00	1.14	1.08	0.70	0.48	0.61
R6	0.99	1.14	1.09	0.78	0.54	0.65
R7	0.99	1.09	1.06	0.85	0.65	0.74
R8	0.98	1.05	1.02	0.93	0.76	0.87
R9	0.99	1.03	1.02	0.97	0.85	0.90
R10	1.00	1.00	1.00	1.00	1.00	1.00

R1: TD1; R2: average of TD1 and 2; R3: average TD1-3, and so on.

Table 6.4: Index weights (X10) and accuracy of phenotypic indices of TD records of milk yield (MY), fat yield (FY) and protein yield (PY) for prediction of phenotype for lactation records.

	INDEX WEIGHTS FOR TD RECORDS										
INDICES	1	2	3	4	5	6	7	8	9	10	ACCURACY
PI1	1490										0.65
	1066										0.59
	1259										0.60
PI2	640	1404									0.80
	659	1031									0.75
	682	1244									0.76
PI3	429	724	1159								0.88
	495	679	926								0.83
	495	721	1052								0.84
PI4	370	470	645	1034							0.92
	406	502	620	881							0.89
	415	512	631	946							0.89
PI5	337	363	454	611	908						0.95
	356	400	463	594	817						0.93
	368	402	468	587	856						0.93
PI6	325	333	359	453	589	727					0.97
	339	361	376	446	564	703					0.95
	343	372	380	446	576	690					0.95
PI7	325	312	322	372	435	455	653				0.98
	321	334	336	372	422	481	607				0.97
	335	340	334	372	432	447	629				0.97
PI8	323	301	315	323	360	344	417	560			0.99
	318	320	320	332	360	367	420	505			0.98
	322	321	318	323	365	344	409	547			0.98
PI9	325	298	309	304	320	306	321	351	479		0.99
	316	307	313	311	322	320	331	355	434		0.99
	317	306	309	303	325	311	320	356	462		0.99
PI10	327	301	307	307	306	297	302	301	332	277	1.00
	317	304	308	307	306	300	302	305	333	267	1.00
	319	308	286	309	311	303	301	306	343	259	1.00

Top row: MY; middle row: FY; bottom row: PY;
PI1: TD1; PI2: TD1 and 2: PI3: TD1-3, and so on.

6.3.2.2 Model 2

The records having 10 tests were assumed to be completed records and those having less than 10 were assumed to be RIP / part records. Phenotypes (deviations) of part and completed TD records were predicted by phenotypic selection indices incorporating sequential TD records. The accuracies and index weights for these phenotypic indices are presented in Table 6.4.

The accuracies (phenotypic correlation of index values with LR) of the index PI10 (having 10 TD records) for milk, fat and protein yields were almost 100%. These accuracies were more than 90% for the indices having at least the first 4 TD records. The index weights for the index PI10 were equal to about 30 for all the TD records as expected. The last TD in indices except the index PI10 received the largest weight (Table 6.4).

The actual and theoretical expansion factors to equate the genetic variance of PPP to the genetic variance of PPC and lactation length weights (LLW) for milk, fat and protein yields are presented in Table 6.5.

The actual expansion factors for *milk yield* ranged from 1.01 (for PPP from the first 9 TD records) to 2.42 (for PPP from the first TD record). The actual expansion factors were similar to the theoretical expansion factors except a few. The LLW ranged from 0.21 (for expanded PPP from the first TD record) to 0.99 (for expanded PPP from the first 8 or 9 TD records).

Table 6.5: Additive variance (VA), phenotypic variance (VP), expansion factors and lactation length weights for predicted (by phenotypic index) phenotypes of part and completed lactation records of milk, fat and protein yields.

RECORDS HAVING TD	VP	VA	EXPANSION FACTORS			LACTATION LENGTH WEIGHTS
			ACTUAL	THEORETICAL		
				1	2	
1	215634	41918	2.42	2.35	2.37	0.21
	283.89	24.38	3.42	2.82	2.87	0.14
	157.20	30.05	2.38	2.71	2.78	0.31
1-2	329600	102670	1.55	1.54	1.56	0.42
	448.26	84.00	1.84	1.78	1.78	0.37
	246.83	60.52	1.68	1.73	1.73	0.44
1-3	391049	145580	1.30	1.29	1.30	0.57
	559.61	115.04	1.58	1.43	1.45	0.42
	304.09	86.69	1.40	1.40	1.41	0.55
1-4	431980	176416	1.18	1.17	1.18	0.68
	639.12	137.33	1.44	1.26	1.26	0.44
	343.68	102.72	1.29	1.24	1.26	0.59
1-5	459004	200698	1.11	1.10	1.11	0.79
	696.34	166.26	1.31	1.15	1.16	0.52
	372.43	113.65	1.23	1.15	1.16	0.61
1-6	474762	215748	1.07	1.07	1.07	0.85
	734.49	191.67	1.22	1.09	1.11	0.59
	390.03	124.01	1.17	1.09	1.11	0.65
1-7	486734	227555	1.04	1.04	1.05	0.91
	760.55	227.27	1.12	1.05	1.06	0.73
	403.94	142.24	1.10	1.06	1.06	0.78
1-8	495768	240353	1.01	1.02	1.03	0.99
	778.82	254.07	1.06	1.03	1.04	0.85
	414.80	161.11	1.03	1.03	1.04	0.94
1-9	502854	243091	1.01	1.01	1.01	0.99
	792.99	267.85	1.03	1.01	1.02	0.90
	422.64	163.34	1.02	1.01	1.01	0.93
1-10	506081	246087	1.00	1.00	1.00	1.00
	799.52	285.66	1.00	1.00	1.00	1.00
	426.64	170.77	1.00	1.00	1.00	1.00

Top row: MY; middle row: FY; bottom row: PY.
Theoretical expansion factor 1: VP(PPC) / VP(PPP);
Theoretical expansion factor 2: reciprocal of squared phenotypic correlation between predicted phenotypes of part (PPP) and completed (PPC) records.

The actual expansion factors for *fat yield* were higher than those for milk yield (Table 6.5). The actual expansion factors were generally higher than the theoretical expansion factors. The LLW ranged from 0.14 (for expanded PPP from the first TD record) to 0.90 (for expanded PPP from the first 9 TD records).

For *protein yield*, the actual expansion factors were slightly higher than for milk yield and lower than for fat yield (Table 6.5). On average, the actual expansion factors were similar to the theortical expansion factors except for PPP from the first TD record. The LLW ranged from 0.31 to 0.93.

6.4 Discussion

Genetic prediction of lactation performance from TD records was attempted only in terms of accuracy assuming fixed effects were known. There was a small increase in the accuracy when later tests (TD 9 and 10) were added to the index for milk yield and protein yield, but the accuracy for fat yield increased substantially when TD10 was added to the index, obviously due to the high heritability of TD10 for fat yield. Van Vleck and Henderson (1961e) and Keown and Van Vleck (1971) also reported that there is very little increase in accuracy for milk yield when last 2-3 tests are added in the selection criterion. On the basis of these results it is confirmed that 6 or 7 equally spaced tests over the entire lactation would be sufficient to get an accuracy equal to the predicted phenotype for LR as concluded in chapter 4. The index weights were different (even negative for a few TD records) for various TD records. The index weights depend on the estimates of heritability and genetic correlation among TD records and between TD records and LR. Sampling variation in these parameter

estimates may be responsible for much of the unequality in index weights. Danell (1979) observed a wide variation in index weights by changing heritability estimate of individual TD records.

A repeatability model had little effect on the accuracy for milk yield and protein yield, but accuracy for fat yield was reduced by upto 12%. When first TD record of fat yield was dropped the maximum loss in accuracy was only about 8%. The loss in accuracy is due to a genetic correlation of less than unity among TD records, because the repeatability model assumes a genetic correlation of unity. Most of the reduction in accuracy of fat and protein yields under the repeatability model may be explained by some unexpected genetic correlations among TD records; for example, the estimates of genetic correlation between TD3 and TD5 of fat and protein yields were only 0.57 and 0.66, respectively. Thus accuracy of repeatability model is not much lower than a genetic index.

Inclusion of RIP / part records in genetic evaluation was considered on a TD basis assuming a repeatability model (model 1) as well as by predicting (by phenotypic index of TD records) phenotype for part and completed TD records (model 2). The accuracy of predicting the phenotype for LR from part and completed TD records was similar for these models.

For repeatability model (model 1), the actual expansion factors were close to 1.0. The PR are expanded by multiplying the deviation of PR from management group mean (**m**) by the actual expansion factors and adding back to **m**. In practice **m** is not known and an estimate of **m** is used. If the expansion factors are close to 1.0, the estimate of **m** will have little effect. Thus repeatability model may be useful when **m** can not be precisely estimated (probably due to a small number of

cows in a management group). The weights given to expanded PR under the repeatability model were generally smaller than those under model 2. There is not much need to expand the part records because the expansion factors are close to 1.0. Thus inclusion of part records assuming a repeatability model is straightforward and only a little expansion of part records and differential weighting of part and completed records is required in animal model evaluation.

For model 2, the actual expansion factors, particularly for fat yield, were slightly higher than the theoretical expansion factors. The actual expansion factors in the present study could not be compared directly to those reported by VanRaden *et al.* (1991) because of inclusion of part records on a TD basis rather than on length (of records) basis. A phenotypic index of all the available TD records for predicting the phenotypes (deviations) of part records in this study, would be better than a single regression on the last test used by VanRaden *et al.* (1991). On average, the actual expansion factors in the present study were slightly higher and the lactation length weights were slightly lower than those reported by VanRaden *et al.* (1991).

The proposed model (model 2) could be directly implemented in current genetic evaluation in the U. K. It would increase the timeliness and accuracy of prediction and would overcome the bias due to exclusion of records of less than 200 days. For example, a heifer with the first two TD milk yields of 25 and 22 kg in a management group with mean of 20 and 18 kg could be included in genetic evaluation by calculating her expanded record in the following way: this heifer would have a phenotypic index value (deviation from mean lactation milk yield (LMY))

64.0 (25 - 20) + 140.4 (22 - 18) = 881.6 kg

where 64.0 and 140.4 are the index weights for TD1 and TD2 from Table 6.4.

Assuming the mean LMY in her management group is 5000 kg, her predicted phenotype from first two TD records would be

881.6 + 5000 = 5881.6 kg

and her expanded predicted phenotype would be

(881.6 x 1.55) + 5000 = 6366.5 kg

where 1.55 is the actual expansion factor from Table 6.5. This expanded record would receive a weight of 0.42 in animal model evaluation.

Prediction of breeding values for LR from TD records assuming a repeatability model seems to be an alternative to multivariate BLUP under the present limited computational facilities. RIP /part records could easily be handled assuming a repeatability model without any need for projection and even expansion. The genetic correlation of predicted phenotype / linear function of first 3 TD records with complete lactation may not be unity (see Table 6.3) indicating that part records having only first 3 TD records may be excluded from the data.

Current animal model evaluations include first five LR of milk, fat and protein yields. Therefore more research is needed to extend these results for later lactations.

APPENDIX TO CHAPTER 6

Appendix Table 6.1: Estimates of genetic variance (bottom diagonal) and covariance (below diagonal), phenotypic variance (top diagonal) and covariance (above diagonal) among TD and lactation records (LR) of milk yield.

	TD										
	1	2	3	4	5	6	7	8	9	10	LR
1	*9710* *1888*	5884	5213	4590	4285	3964	3726	3539	3216	2788	1447
2	2108	*9348* *2897*	6550	5885	5441	5008	4779	4534	4144	3605	1689
3	2286	2820	*9360* *3254*	6558	5993	5635	5325	4999	4592	4058	1783
4	1971	2733	3070	*8790* *3352*	6342	5869	5601	5354	4907	4299	1779
5	1967	2656	2818	3172	*8297* *3498*	6096	5805	5515	5092	4575	1755
6	1933	2431	2938	3152	3072	*7944* *3136*	5966	5597	5100	4563	1700
7	1762	2458	2792	3162	3153	3029	*7811* *3070*	5910	5410	4835	1683
8	1644	2442	2749	3238	3314	3064	3202	*7860* *3432*	5811	5246	1658
9	1500	2054	2310	2768	3060	2750	2815	2989	*7796* *2795*	5871	1589
10	1518	1856	2173	2444	2997	2534	2584	2808	2708	*8909* *2964*	1477
LR	570	750	834	894	911	863	862	888	790	756	*509097* *249592*

All estimates are multiplied by 1000 except genetic and phenotypic covariance of TD records with LR and genetic and phenotypic variance of LR.

Appendix Table 6.2: Estimates of genetic variance (bottom diagonal) and covariance (below diagonal), phenotypic variance (top diagonal) and covariance (above diagonal) among TD and lactation records (LR) of fat yield.

	TD										
	1	2	3	4	5	6	7	8	9	10	LR
1	*2500* *2147*	9851	8189	7337	6921	6431	6284	5957	5513	5071	2664
2	2636	*1935* *3653*	9102	8045	7597	6971	6912	6632	6269	5814	2644
3	2222	2407	*1792* *2870*	8870	8205	7830	7558	7192	6614	6300	2682
4	1908	2162	2211	*1566* *2536*	8664	8223	7976	7718	7102	6616	2632
5	2093	2776	1785	2256	*1457* *3394*	8745	8632	8212	7624	7233	2635
6	1861	1868	2442	2814	2629	*1418* *3674*	9139	8842	7979	7578	2616
7	2498	2954	2771	3079	3452	3727	*1430* *4267*	9560	8719	8167	2656
8	2324	2555	2829	3292	3450	4214	4514	*1480* *5021*	9454	8855	2657
9	2346	2930	2380	2907	3519	3401	4060	4241	*1476* *4075*	9659	2560
10	2731	3388	2806	3279	4164	4039	4799	5014	4665	*1659* *5568*	2464
LR	701	842	767	816	912	954	1121	1165	1068	1256	*80604* *29776*

Estimates of genetic variance and genetic and phenotypic covariance among TD records are multiplied by 10^6, phenotypic variance of TD records by 10^5, covariance between TD and LR by 10^4 and variance of LR by 100.

Appendix Table 6.3: Estimates of genetic variance (bottom diagonal) and covariance (below diagonal), phenotypic variance (top diagonal) and covariance (above diagonal) among TD and lactation records (LR) of protein yield.

	TD										
	1	2	3	4	5	6	7	8	9	10	LR
1	*9913* *1895*	4600	4051	3655	3442	3220	3118	3128	2897	2661	1248
2	1563	*7931* *1628*	4759	4260	3963	3594	3609	3526	3321	3062	1300
3	1766	1489	*8257* *1938*	5072	4528	4280	4186	4006	3725	3769	1412
4	1520	1351	1722	*7938* *1833*	5039	4631	4512	4397	4063	3721	1440
5	1458	1352	1291	1584	*7590* *1967*	4982	4836	4615	4314	4066	1444
6	1520	1107	1630	1730	1576	*7475* *1849*	5114	4816	4365	4082	1416
7	1720	1621	1786	1998	2029	1901	*7666* *2373*	5365	4891	4533	1455
8	1906	1696	1994	2251	2223	2162	2559	*7967* *2932*	5471	5105	1473
9	1584	1376	1426	1712	2038	1725	2091	2414	*8187* *2206*	5818	1439
10	1834	1570	1603	1924	2320	1918	2301	2790	2604	*9638* *3198*	1394
LR	513	449	510	539	545	525	622	702	587	676	*43084* *17332*

Estimates of variance and covariance among TD records are multiplied by 10^6, covariance between TD and LR by 10^4 and variance of LR by 100.

CHAPTER 7

GENERAL DISCUSSION AND CONCLUSIONS

The main objectives of this study were to estimate the genetic parameters of test day (TD) records of British Holstein-Friesian heifers and to determine how best to use these parameters for the prediction of breeding values for lactation records (LR).

The cost of recording is very high and as a consequence only a proportion of cows are recorded in the national recording scheme in the U. K. A further objective was therefore to investigate the effect of reducing the cost of recording (less frequent than monthly recording) on the accuracy of selection.

Parameter estimates for TD1-8 and for LR discussed in chapter 3 were from the data (data set 1) on heifers having at least 8 TD records. This data set was specifically considered because culling was low (4.5%) at this stage, whereas almost 32% heifers did not have a 10^{th} test. The culling had little effect on the estimates of genetic parameters of TD records and LR records, however, there being little difference in genetic parameters between data set 1 and data set 2 (a subset of data set 1 on heifers having at least 10 tests). Estimates of heritability for TD records of milk, fat and protein yields were high in mid and late lactation. Heritability estimates of first TD records of milk, fat and protein yield were lower than for the other TD records because of lower sire variances and higher within sire variances. The genetic correlation of first TD record with corresponding LR was also lower than the genetic correlations of other TD records with LR. The accuracy of selection depends on the

heritability estimate of TD records and their genetic correlation with LR. One of the options might be to exclude the first test for genetic evaluation or the first test may be postponed until daily yields are stabilised. More research is needed to estimate the heritability of daily yields and their genetic correlations with LR in the early part of the lactation in order to decide the earliest day for the first test. Genetic correlations among TD records and between TD and LR were also high in mid lactation.

High heritability estimates of TD records in mid lactation and high genetic correlations among TD records and between TD and LR in this study are in line with the results of previous studies. However, estimates of heritability for TD and LR, particularly of milk yield and protein yield, were higher than the previously reported estimates. After a detailed dissection of reasons for the high heritability estimates by analysing alternative data sets by different models (section 3.4.1), it was concluded that the high heritability estimates in this study may be due to lower within sire variances due to fitting herd-year-month (of test) rather than herd-year-season (of calving) and using a mixture of data on daughters of DPTS and non-DPTS sires (usually heritability estimates were higher for non-DPTS data). There seems to be time trend in heritability estimates of LR as discussed in chapter 3. Further research is needed to investigate the reasons for the time trend in heritability estimates by analysing data spanning many years using REML.

In chapter 4 the consequences of reduction in cost of recording on loss in accuracy of selection were investigated using the parameter estimates from 5 bi-monthly test day records (BMTD) obtained in

alternate herd visits to simulate the case of bi-monthly recording. Estimates of heritability of BMTD records and genetic correlations among BMTD and between BMTD and LR were similar to the average estimates for corresponding monthly TD records. The additional variance induced by the changed data structure (wide range of days to first test) of BMTD records can therefore be removed by fitting an appropriate statistical model, i. e. quartic regression on day of lactation for the first test. The loss in accuracy of prediction of breeding value for LR using 5 bi-monthly tests compared to 10 monthly tests was only 3%. This loss in accuracy seems negligible compared to the potential reduction in cost of recording and possible broadening of the selection base for national genetic evaluations. The number of traits for any multivariate analysis is also halved by bi-monthly recording. Current genetic evaluations are carried out using predicted phenotypes (from monthly TD records) for LR. The predicted phenotypes for LR from BMTD records are less accurate than those predicted from monthly TD records. But for genetic evaluation for LR from the genetic indices of TD records, the interval between tests does not matter as long as the error structure of TD records can be corrected by fitting an appropriate statistical model. The results obtained in chapter 4 suggest that the cost of recording can be reduced with very little loss in accuracy by genetic evaluation on an index (with equal or unequal weights) of 6-7 TD records.

The LR used for parameter estimation in this study were predicted (by linear interpolation) from the TD records and actual LR (sum of daily yields) were not known. The phenotypic correlations between TD records and actual LR may be different from the phenotypic correlations between TD records and predicted LR. Phenotypic

correlations between daily yields less than 30 days apart were estimated from daily milk yield records from Langhill data. Approximate phenotypic correlations between TD records and actual LR were predicted using the information on correlations between daily yields and between TD records (chapter 5). There was not much difference between estimated phenotypic correlations of TD records with predicted LR and predicted phenotypic correlations of TD records with actual LR. The correlations between TD records and actual LR were indirectly predicted. More research is needed to estimate the phenotypic correlations between TD records and actual LR using a large data set.

Current genetic evaluation is a two step procedure, i. e. prediction of breeding values from phenotypically predicted LR. Theoretically the most accurate method to predict breeding values for LR would be a multivariate BLUP specifying the full variance-covariance structure of TD records. In chapter 6 the accuracy of genetic prediction of lactation performance from TD records using genetic indices (assuming fixed effects are known) and a repeatability model was investigated. Using genetic indices of successive TD records it was shown that there was very little increase in accuracy, particularly for milk yield and protein yield, when later tests (TD 9 and 10) were added to the index. It indicates that the first 8 tests would be sufficient to achieve an accuracy equal to 10 tests. These findings support the results obtained in chapter 4, where it was concluded that the accuracy of 6-7 TD records over the entire lactation was similar to the accuracy of 10 monthly TD records.

There was not much reduction in accuracy for milk and protein yields assuming a simplified (repeatability) model, but accuracy for fat

yield was reduced by 12%, because of an unexpected genetic correlation between TD3 and TD5 (section 6.4). This loss in accuracy could be avoided by excluding the first TD record. The repeatability model was found to have some additional advantages when part records are included in genetic evaluations. There is no need to project the part TD records under the repeatability model. The values obtained for expansion factors (to equate the genetic variances of part records to the genetic variance of completed records) and for the weights (lactation length weights) given to part records in animal model evaluations assuming a repeatability model, were lower than in other methods (phenotypic index used in this study and a method proposed by VanRaden *et al.*, 1991). Lower weights for part records would result in more stable evaluations. Although the genetic correlation of the average phenotype for the first three TD records with LR or the average phenotype of 10 TD records is less than unity (rg = 0.90), these records are useful for making preliminary selection decisions in a sequential selection scheme.

Inclusion of part records in current genetic evaluations in the U. K. was investigated by predicting the phenotypic deviations (from the mean) of part and completed TD records using the phenotypic indices of successive TD records. This method for inclusion of part TD records in genetic evaluations could be directly implemented in current genetic evaluations in the U. K. Inclusion of records of less than 200 days in genetic evaluations in the U. K. would result in increased accuracy (simply by including daughters of a sire with records of less than 200 days in addition to daughters with 305-day records) and reduced bias.

For genetic prediction of lactation performance from TD records, a repeatability model could be used as an alternative to multivariate BLUP under the present limited computational facilities for national evaluation with a large data set. Further research is needed for estimation of parameters in later lactations and to optimise the breeding value prediction for lactation performance using TD records of the first five lactations.

BIBLIOGRAPHY

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

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

Amir, S., Halevi, A., Eldelman, Z. and Kali, J. 1978. The effect of age at first calving on subsequent milk yield of dairy cows. *Special Publication no.* **119**. *Dir. Sci. Publ. The volcani centre, Bet. Dagan, Israel* (cited by Danell, 1982a).

Anderson, S. M., Mao, I. L. and Gill, J. L. 1989. Effect of frequency and spacing of sampling on accuracy and precision of estimating total lactation yield and characteristics of lactation curve. *J. Dairy Sci.* **72**: 2387-2394.

Auran, T. 1973. Studies on monthly and cumulative monthly milk yield records. I. The effect of age, month of calving, herd and month of first test period. *Acta Agric. Scand.* **23**: 189-199.

Auran, T. 1976a. Studies on monthly and cumulative monthly milk yield records. III. Estimates of genetic and phenotypic parameters. *Acta Agric. Scand.* **26**: 3-9.

Auran, T. 1976b. Studies on monthly and cumulative monthly milk yield records. IV. Estimating total lactation from part lactation. *Acta Agric. Scand.* **26**: 10-17.

Bar-Anan, R., Weller, J. I., Wiggans, G. R. and Osterkorn, K. 1986. Prediction of annualized lactation yield from partial lactations. *J. Dairy Sci.* **69**: 1897-1903.

British Standards Institution. 1972. Recommendations for the methods of milk recording of cow. *British Standards Institution* **4866**, pp. 12-17.

Bulmer, M. G. 1971. The effect of selection on genetic variability. *American Naturalist* **105**: 201-211.

Chauhan, V. P. S. and Hill, W. G. 1986. Seasonal grouping in a herd-year-season model of sire evaluation. *Anim. Prod.* **43**: 63-71.

Congleton, Jr. W. R. and Everett, R. W. 1980. Application of incomplete gamma function to predict cumulative milk production. *J. Dairy Sci.* **63**: 109-119.

Danell, B. 1979. Calculation of selection indices that include highly correlated traits. *Acta Agric. Scand.* **29**: 103-107.

Danell, B. 1982a. Studies on lactation yield and individual test day yields of Swedish dairy cows. I. Environmental influences 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. II. Estimates of genetic and phenotypic parameters. *Acta Agric. Scand.* **32:** 83-92.

Danell, B. 1982c. 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.

Danell, B. 1990. Genetic aspects of different parts of lactation. *In*: *Proc. 4th World Congr. Genet. Appl. Livest. Prod.*, **Edinburgh, Vol. XIV:** 114-117.

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

Gnanasakthy, A. 1989. *Comparison between the current method and a proposed method of estimating total yield of cows on the basis of production records collected by national recording scheme.* **M. Sc. thesis, Sheffield City Polytechnic, U. K.**

Goddard, M. E. 1987. Genetic evaluation of dairy cattle for profitability. *In*: *Proc. 6th Conf. Aust. Assoc. Anim. Breed. Genet.*, Perth. pp. 192-198.

Harman, H. H. 1970. *Modern factor analysis.* University of Chicago Press, Chicago and London.

Harvey, W. R. 1977. *Users guide for LSML76. Mixed model least-squares and maximum likelihood computer program.* Ohio State Univ., Columbus. (Mimeograph).

Hickman, C. G. 1971. Changes in the shape of lactation curve and other correlated responses from selection for short period lactation yield. *In*: *Proc. 22nd Ann. Meet. Eur. Assoc. Anim. Prod.*, Paris (cited by Danell, 1982c).

Hill, W. G., Edwards, M. R., Ahmed, M.-K. A. and Thompson, R. 1983. Heritability of milk yield and composition at different levels and variability of production. *Anim. Prod.* **36:** 59-68.

Kendall, M. G. and Stuart A. 1963. *The Advance Theory of Statistics. Vol. I. Distribution Theory. 2nd ed.* Griffin, London.

Keown, J. F. and Van Vleck, L. D. 1971. Selection on test day fat percentage and milk production. *J. Dairy Sci.* **54:** 199-203.

Lamb, R. C. and McGilliard, L. D. 1960a. Variables affecting ratio factors for estimating 305-days production from part lactations. *J. Dairy Sci.* **43:** 519-528.

Lamb, R. C. and McGilliard, L. D. 1960b. Comparison of ratio factor for extending part time milk and fat records. *J. Dairy Sci.* **43**: 879 (Abstr).

Lee, A. J. and Hickman, C. G. 1972. Age and herd adjustment of first lactation milk yield. *J. Dairy Sci.* **55**: 433-438.

Lindgreen, N., Philipsson, J. and Berstedt, A. E. 1980. Studies on monthly protein records of individual cows. *Acta Agric. Scand.* **30**: 437-444.

Madden, D. E., McGilliard, L. D. and Ralston, M. P. 1959. Relations between test day milk production of Holstein cows. *J. Dairy Sci.* **42**: 319-326.

McDaniel, B. T. 1969. Accuracy of sampling procedures for estimating lactation yields: a review. *J. Dairy Sci.* **52**: 1742-1761.

Meinert, T. R., Korver, S. and Van Arendonk, J. A. M. 1989. Parameter estimation of milk yield and composition for 305 days and peak production. *J. Dairy Sci.* **72**: 1534-1539.

Meyer, K. 1984. Estimates of genetic parameters for milk and fat yield for the first three lactations in British Friesian cows. *Anim. Prod.* **38**: 313-322.

Meyer, K. 1986. Restricted maximum likelihood to estimate genetic parameters-in practice. *In*: *Proc. 3rd World Congr. Genet. Appl. Livest. Prod.*, **Nebraska, Vol. XII**: 454-459.

Meyer, K. 1987. Estimates of variances due to sire x herd interactions and environmental covariances between paternal half-sibs of first lactation dairy production. *Livest. Prod. Sci.* **17**: 95-115.

Meyer, K., Graser, H. U. and Hammond, k. 1989. Estimates of genetic parameters for first lactation test day production of Australian black and white cows. *Livest. Prod. Sci.* **21**: 177-199.

Milk Marketing Board. 1979. The improved contemporary comparison. *Report of the Breeding and Production Organisation, Milk Marketing Board* No. 29, pp. 91-95.

Milk Marketing Board. 1989. *Report of the Breeding and Production Organisation, Milk Marketing Board* No. 39

Miller, R. H., McDaniel, B. T. and Corley, E. L. 1967. Variation in ratio factors for age adjusting part lactation records. *J. Dairy Sci.* **50**: 1819-1823.

Nelder, J. A. 1966. Inverse polynomials, a useful group of multifactor response functions. *Biometrics* **22**: 128-141.

Patterson, H. D. and Thompson, R. 1971. Recovery of inter block information when block sizes are unequal. *Biometrika* **58**: 545-554.

Quaas, R. L., Everett, R. W. and McClintock, A. C. 1979. Maternal grandsire model for dairy sire evaluation. *J. Dairy Sci.* **62**: 1648-1654.

Rao, M. K. and Sundaresan, D. 1980. Comparison of methods of extending part lactation records for sire evaluation. *J. Agric. Sci.* **94**: 57-62.

Ronningen, K. 1976. Phenotypic and genetic parameters of characters related to milk production in cattle. *Acta Agric. Scand.* **17**: 83-100.

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

Searle, S. R. 1961. Part lactations. II. Genetic and phenotypic studies of milk fat yields. *J. Dairy Sci.* **44**: 282-295.

Shook, G. E., Johnson, L. P. and Dickinson, F. N. 1980. Factors for improving accuracy of estimates of test-interval yield. *DHI Lett.* **56(4)** (cited by Anderson *et al.*, 1989).

Spike, P. W. and Freeman, A. E. 1967. Environmental influences on monthly variation in milk constituents. *J. Dairy Sci.* **50**: 1897-1904.

Swanson, G. J. T. and Gnanasakthy, A. 1990. Estimation of genetic parameters for milk cell count and correlation with production traits in Friesian / Holstein heifers. *Anim. Prod.* **52**: 611(Abstr).

Syrstad, O. 1964. Studies on dairy herd records. I. Evaluation of incomplete records. *Acta Agric. Scand.* **14**: 129-149.

Syrstad, O. 1965. Studies on dairy herd records. II. Effect of age and season at calving. *Acta Agric. Scand.* **15**: 31-61.

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

Van der Werf, J. H. J. and De Boer, W. 1989a. Influence of nonadditive effects on estimation of genetic parameters in dairy cattle. *J. Dairy Sci.* **72**: 2606-2614.

Van der Werf, J. H. J. and De Boer, W. 1989b. Estimation of genetic parameters in a crossbred population of Black and White dairy cattle. *J. Dairy Sci.* **72**: 2615-2623.

VanRaden, P. M., Wiggans, G. R. and Ernst, C. A. 1991. Expansion of projected lactation yield to stabilize genetic variance. *J. Dairy Sci.* **74**: 4344-4349.

Van Vleck, L. D. 1975. What to expect from 'part records' sire proofs. *Hoard's Dairyman.* **10**: 291-292. (cited by Danell, 1982c).

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

Van Vleck, L. D. and Henderson, C. R. 1961b. Regression factors for extending part lactation milk records. *J. Dairy Sci.* **44**: 1085-1092.

Van Vleck, L. D. and Henderson, C. R. 1961c. Ratio factors for adjusting monthly test day data for age and season of calving and ratio factors for extending part lactation records. *J. Dairy Sci.* **44**: 1093-1102.

Van Vleck, L. D. and Henderson, C. R. 1961d. Regression factors for predicting a succeeding complete lactation milk record from part lactation record. *J. Dairy Sci.* **44**: 1322-1334.

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

Van Vleck, L. D. and Henderson, C. R. 1961f. Extending part lactation milk records by regression ignoring herd effects. *J. Dairy Sci.* **44**: 1519-1528.

Van Vleck, L. D. and Henderson, C. R. 1961g. Utilizing both part and complete daughter records in sire evaluation. *J. Dairy Sci.* **44**: 2068-2076.

Visscher, P. M. 1991. *Estimation of genetic parameters in dairy cattle using an animal model and implications for genetic improvement.* **Ph. D. thesis, University of Edinburgh, U. K.**

Weller, J. I. 1988. Inclusion of partial lactations in genetic analysis of yield traits by differential weighting of records. *J. Dairy Sci.* **71**: 1873-1879.

Wiggans, G. R. and Van Vleck, L. D. 1978. Effect of proportion of concentrates in ration on lactation curves. *J. Dairy Sci.* **61**: 135-140.

Wiggans, G. R. and Dickinson, F. N. 1985. Standardisation of NCDHIP dairy cattle lactation records. *In: Report of the National Cooperaive Dairy Herd Improvement Program.*

Wilmink, J. B. M. 1987a. Adjustment of test day milk, fat and protein yield for age, season and stage of lactation. *Livest. Prod. Sci.* **16**: 335-348.

Wilmink, J. B. M. 1987b. Comparison of different methods of predicting 305 day milk yield using means calculated from within herd lactation curves. *Livest. Prod. Sci.* **17**: 1-17.

Wilmink, J. B. M. 1987c. Efficiency of selection for different cumulative milk, fat and protein yields in first lactation. *Livest. Prod. Sci.* **17**: 211-224.

Wilmink, J. B. M. 1988. Effects of incomplete records on genetic parameters among cumulative yields in first lactation and on extension of part lactations. *Livest. Prod. Sci.* **18**: 19-34.

Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. *Nature* **216**: 164.

Wood, P. D. P. 1969. Factors affecting the shape of lactation curve in cattle. *Anim. Prod.* **11**: 307-316.