The Importance of Different On-Farm Feeding and Management Practices on Milk Protein Concentration and Yield

Final Report - Project 4353

[1]J.J. Murphy, [2]D. Harrington, [1]D. Cliffe and [1]J. Crilly,

[1]Dairy Production Department, Moorepark Research Centre, Fermoy, Co. Cork. and [2]Data Systems Department, Teagasc Headquarters, 19 Sandymount Avenue, Ballsbridge, Dublin 4

Summary

The objective of this project was to identify the most important factors which influence milk protein concentrations and yields on farms. Data collected on approximately 300 DairyMis farms were subjected to statistical evaluation by regression analysis, analysis of variance and factor analysis. The range in protein concentrations was 12% of the mean value compared to a variation in protein yield of 86% of the mean. The range in milk yield was 90% of the mean value.

Initially data from 1995 were analysed. The regression model used only accounted for 4% of the variation between farms in milk protein concentration but for over 97% of the variation in milk protein yield. Milk protein yield was highly correlated with milk yield.

Confining the analysis to Spring-calving herds only in 1995 (n=128) gave similar results but including genetic data in the regression model, for those spring-calving herds for which it was available (n = 36), explained approximately 25% of the variation in protein concentration.

Analysis of variance on the spring-calving herd data (for 1994, 1995 and 1996) divided into quartiles on the basis of protein concentration indicated that higher protein concentration was associated with later calving, a higher proportion of milk produced on pasture and lower milk yield per cow.

Factor analysis on all herds in 1994, 1995 and 1996 indicated that higher protein concentration was associated with Spring-calving herds, larger herds and with grass based milk production. These three factors accounted for 8% of the observed variation in milk protein concentration.

Introduction

Protein is now regarded as being the most valuable component of milk due to the ongoing change in market demand influenced by consumer diet preferences and the absence of quota restrictions on its production. Therefore, the objective of many producers and processors in the 1990's has been to produce and process greater quantities of protein. In general, producers have attempted to produce more protein by increasing the protein concentration in the milk rather than by increasing the milk output. This strategy was forced on them because of a quota limit on milk volume production. While much work has been done in controlled experiments to investigate the nutritional factors that affect milk protein concentration it is necessarily more complex in a farm environment where multiple influences may affect the outcome. The objective of this project was to determine the most important on-farm factors influencing both milk protein concentration and yield.

Teagasc
Agriculture and Food Development Authority

Methods

The data currently being collected from DairyMIS (dairy management information system) farms were used in the study. The factors for which information was available are shown in Table 1.

Table 1 : Data on farms available for inclusion in the Model

Fat (g/kg)	April to March by month
Protein (g/kg)	April to March by month
Calving %	April to March by month
% Meal fed by month April to March	
% Milk produced by month April to March	
% Milk produced between April and October (on pasture)	
No. of cows per calving day farm	
% in milk	
Gal/ac	
Gal/cow	
Average Fat (g/kg)	
Average Protein (g/kg)	
Fat yields per cow (kg)	
Protein yield per cow (kg)	
Fat per ha (kg)	
Protein per ha (kg)	
Concentrate fed (no. of 50 kg)	

Concentrate cost (£ / t)	
Concentrate cost (£ / gal of milk produced)	
Stocking Rate (lu/ha)	
Percentage of farm for 1st cut silage	
% of farm for 2nd cut silage	
Utilisable ME (GJ/ha)	
Utilisable ME (GJ/lu)	
Nitrogen fertilizer used (kg/ha)	
Phosphorus fertilizer used (kg/ha)	
Potassium fertilizer used (kg/ha)	
Area of farm (ha)	

Based on existing knowledge as to factors in Table 1 that were most likely to influence average annual protein percent and yield the following six were used in the initial regression model ; percentage of milk produced between April and October (on pasture), milk yield per cow, the quantity of concentrates fed per cow, stocking rate, utilisable metabolisable energy (UME) in GJ/lu (energy got from forage) and calving day.

Data from a total of 314 herds were available in 1995. This total data set, a subset consisting of Spring-calving herds only i.e. those herds having greater than 80% of their cows calving in the January to March period and a subset of the Spring-calving herds for which genetic data were available were subjected to analysis. Genetic data for the 314 herds were requested from the IDRC (Irish Dairy Records Co-op) and was available on 180 herds. The data supplied were the number of cows in the herd having genetic data, the average RB1 and the average predicted differences for milk yield, fat yield, protein yield, fat % and protein %. Only herds that had genetic data on at least 50% of the cows in the herd were used in the analysis. This reduced the number of useable herds to 56, 36 of which were principally Spring calving.

Silage analysis was requested from the farmers via a posted letter and analysis sheet which they had to complete. Ninety three replies were obtained of which 44 had an analysis of their silage. Only 4 of the 36 Spring calving herds and 7 of the 20 Autumn calving herds for which genetic data were available had silage analysis and therefore it was excluded from the model.

The analysis applied consisted of (1) determining simple correlation coefficients between protein percent and yield and other factors, (2) regression analysis using a backwards elimination procedure, (3) forming quartiles based on protein percent and performing an analysis of variance on quantitative variables and (4) factor analysis applied to the mean of variables related to farm management practices over the years 1994, 1995 and 1996. Initially analyses (1), (2) and (3) were applied to the 1995 data set and subsequently analysis (3) only was applied to the data from the Spring-calving herds in 1994 and 1996 of which there were 100 and 112 respectively.

Results and Discussion

The characteristics of the herds in the total data set are shown in Table 2. Herd size ranged from 18 to 432 cows with milk yields, protein yields and protein percents ranging from 3136 to 8143 kg per cow, 103.8 to 258.1 kg per cow and 3.03 to 3.43 % respectively. The range of protein concentration was 12% of the mean value compared to a variation in protein yield of 86% of the mean. The range in milk yield was 90% of the mean value.

Table 2 : Characteristics of the herds used in the initial model

	Minimum	Maximum	Mean
Milk April - October (%)	52.1	99.7	70.2
Yield (kg/cow)	3136	8143	5546
Concentrate fed (kg/cow)	209	2253	820
Stocking Rate (lu/ha)	1.56	4.02	2.71
U.M.E. (GJ/lu)	33.3	51.9	42.5
Calving day (days from Jan 1)	-69	119	42
Calving January - March (%)	0	100	68
Number of cows	18	432	79
Milk fat (g/kg)	33.3	42.6	36.6
Milk fat per cow (kg)	118.3	283.1	202.9
Milk fat per ha (kg)	258.1	965.0	552.0
Milk protein (g/kg)	30.3	34.3	32.2
Milk protein per cow (kg)	103.8	258.1	178.7
Milk protein per ha (kg)	229.7	853.9	485.6

Figures 1 and 2 show the relationship of milk yield with protein yield and protein concentration respectively on the study farms, together with fitted least squares regression lines and table 3 gives the correlation coefficients between protein concentration, protein yield and other variables.

Table 3: The correlation coefficients between protein concentration (g/kg), protein yield (kg/cow) and a number of other variables.		
	Protein Conc.	**Protein Yield**
Fat (g/kg)	0.53	0.08
Protein (g/kg)	1.00	-0.01
Milk April - October (%)	0.15	-0.56
Yield per cow (kg)	-0.17	0.97
Concentrate fed (kg)	-0.18	0.50
Stocking Rate (lu/ha)	0.02	0.09
U.M.E. (GJ/lu)	0.01	0.52
Calving day	0.08	-0.35

Protein concentration was most highly correlated with fat concentration (r = 0.53) which is probably a reflection of the genetic link between them. There was a weaker positive correlation with the percentage of milk produced in April to October and weak negative correlations with milk yield per cow and level of concentrates fed per cow. Protein yield was very strongly correlated with milk yield per cow (r = 0.97).

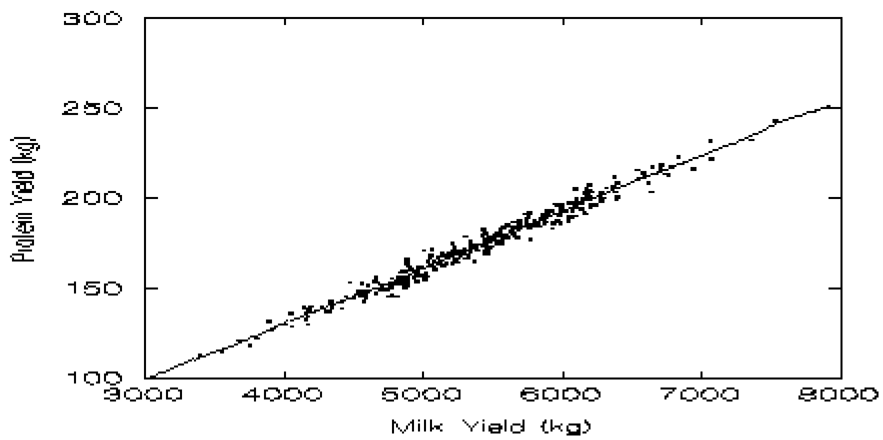

Figure 1 : The relationship between protein concentration and milk yield.

Milk yield is clearly the main determinant of protein yield (P<0.001, r = 0.97). Therefore any attempt to raise protein yield should, in the absence of quota, focus on factors which increase milk yield. There is a weak but significant negative correlation between milk yield and protein concentration (P = 0.005, r = -0.17) (Figure 2 and Table 3), indicating that any such increase may be accompanied by a small reduction in protein concentration.

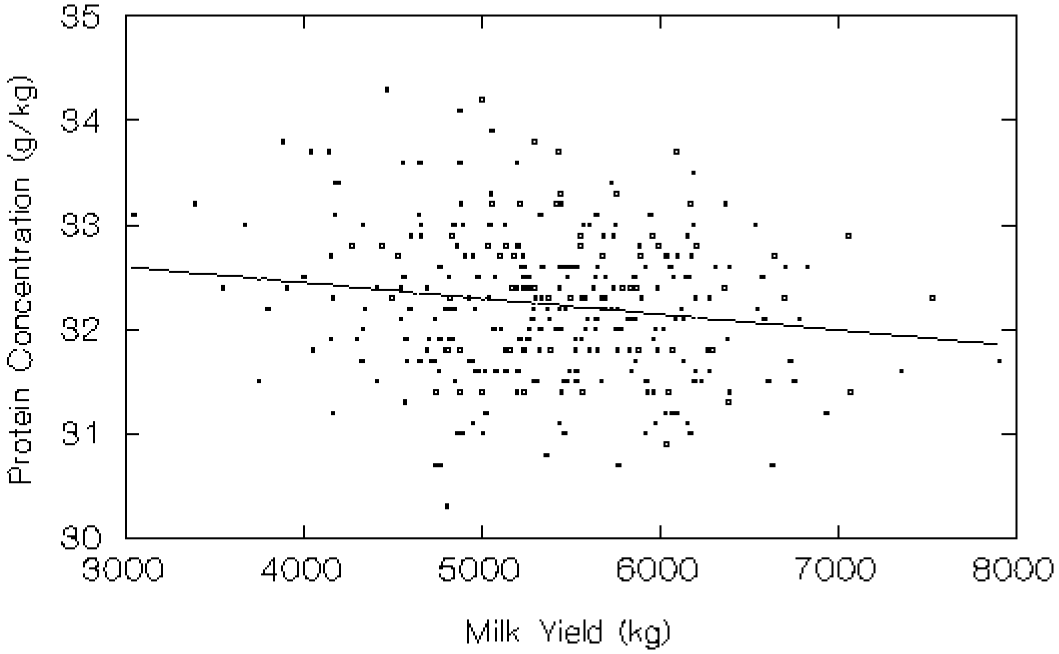

Figure 2 : The relationship between protein concentration and milk yield

There was no linear relationship between protein yield and protein concentration as can be seen graphically in figure 3 and as evidenced by the correlation coefficient of -0.01 (Table 3). However, the data suggested some association at either end of the protein concentration range which appeared to be related to milk volume. Protein concentration in the 46 lowest yielding herds, (milk yield <=4700 kg), averaged 32.57g/kg (sem=0.10) compared to 32.18 g/kg (sem=0.04) for the remaining 263 herds. This difference of 0.39 g/kg was statistically significant (P<0.001). There was evidence of a corresponding dilution effect on protein concentration in the highest yielding herds. Protein concentration averaged 32.05 g/kg (sem=0.12) in 34 herds with milk yields => 6200 kg compared to 32.26 g/kg (sem=0.04) in the remaining 275 herds. However this difference did not reach statistical significance (P=0.103).

There was no significant correlation between protein yield and protein concentration (Table 3) suggesting that the potential for raising protein yield by increasing protein concentration is extremely limited particularly when compared to the benefits which can be achieved by increasing milk yield.

There are two separate mechanisms whereby on-farm practices might influence milk protein levels. They could either have a direct effect on protein yield or they could act indirectly through their effect on milk yield. On-farm factors which impact on protein yield through increased milk yield are best examined by looking at their effects on milk yield itself. Direct effects on protein yield are equivalent to changes in protein concentration and can be investigated by examining the effects of on-farm factors on residual protein yields after the milk yield effect has been removed. These residuals are effectively the dispersion of the scatterpoints around the regression line in Figure 1.

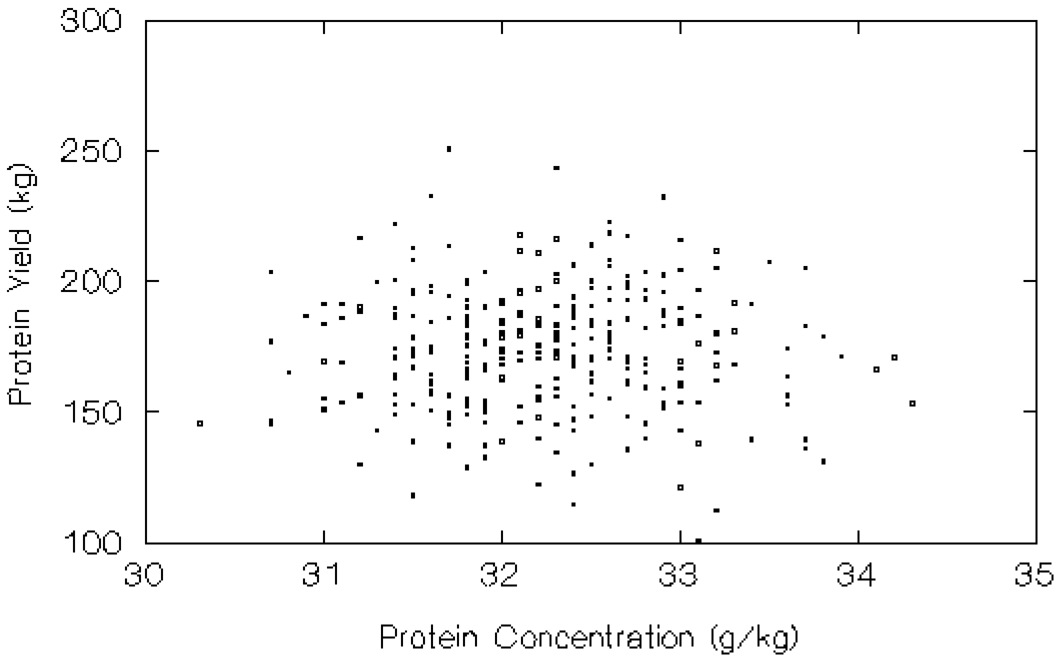

Figure 3 : The relationship between protein yield and protein concentration

When the residuals themselves are plotted against milk yield (figure 4), it can be seen that for any given milk yield, the 95% confidence interval for the residuals is no more than ± 10 kg of protein. The fitted ellipse shows the 95% confidence interval for the sample means. The major axes are determined by the standard deviations of X and Y and the orientation by the sample covariance between X and Y.

A typical herd with a protein yield of 174kg could therefore expect to get no more than 10kg (6%) extra protein by increasing protein concentration. This is consistent with the observed range of protein concentration on the study farms of 4 g/kg or 12% of the mean. Again the maximum improvement which could be achieved by an average herd would be 6%.

Alternatively, to achieve a 10 kg increase in protein yield by increasing milk production would require only an extra 310 kg of milk at 32.3 g/kg protein.

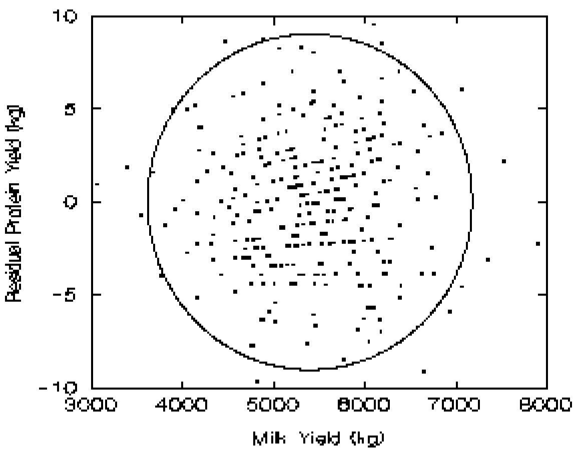

Figure 4 : Plot of residual protein yield (after removing the effect of milk yield) against milk yield. The ellipse shows 95% confidence interval for sample means.

The proportion of the variation in fat concentration, fat yield per cow (kg) fat yield per ha (kg), protein concentration, protein yield per cow (kg) and protein yield per ha (kg) accounted for by the regression model using the 6 variables outlined in the methods section (percentage of milk produced between April and October (on pasture), milk yield per cow, the quantity of concentrates fed per cow, stocking rate, utilisable metabolisable energy (UME) in GJ/lu (energy got from forage) and calving day). is shown in Table 4.

Table 4 : The percentage of the variation accounted for in the regression model when applied to all the herds in 1995

Response Variate	Percentage of Variation Accounted for
Fat (g/kg)	7.5
Fat/cow (kg)	24.7
Fat/ha (kg)	97.0
Protein (g/kg)	3.8
Protein/cow (kg)	97.5
Protein/ha (kg)	98.1

Little of the variation in either protein or fat concentration was accounted for by the regression model. Only .038 of the variation in protein percent was accounted for. However .975 of the variation in protein yield per cow was accounted for by the regression model.

Similar regression analysis was done on the subset of Spring-calving cows (n = 128) and the results of the analysis are shown in Tables 5 and 6. Similar trends in the correlation coefficients and in the proportion of variation accounted for by the regression model were observed as for the larger data set.

Table 5 : The correlation coefficients for protein concentration, protein yield (kg/cow) and a number of other variables (n=128)

	Protein %	Protein Yld
Fat (g/kg)	0.54	0.02
Protein (g/kg)	1.00	-0.06
Milk April - October (%)	0.21	-0.54
Yield per cow (kg)	-0.22	0.99
Concentrate fed (kg)	-0.14	0.27
Stocking Rate (lu/ha)	0.04	0.14
UME (GJ/hA)	-0.13	0.78
Calving day	0.13	-0.22

Table 6 : The percentage of the variation accounted for in the regression model applied to the Spring-calving herds (n=128)

Response Variate	Percentage of Variation Accounted for
Fat (g/kg)	8.2
Fat/cow (kg)	94.0
Fat/ha (kg)	97.6
Protein (g/kg)	3.1
Protein/cow (kg)	97.2
Protein/ha (kg)	98.3

Genetic variability in protein concentration.

As reliable genetic assessment was unavailable for most cows on the study farms, it was not possible to remove this source of variation from the observed variability in milk protein concentration between herds. However genetic merit had been calculated for at least 50 per cent of the cows in 56 of the herds. These herds were therefore used as a subset to investigate the genetic component of between herd differences. As milk yield and composition were used to calculate cow genetic merit, a degree of auto-correlation would be expected between observed differences in milk protein yield and estimated genetic potential.

Assuming that the calculated genetic merit was representative of the whole herd, the predicted difference in milk yield for the 56 herds ranged from -205 to +174 kg with a mean of -51.4 kg. For protein yield, the predicted difference ranged from -5.7 to +5.5 kg (mean -1.0 kg) while for protein concentration the genetic component varied from -1.0 to +4.0 g/kg with a mean of +1.4 g/kg. The fact that the predicted genetic component was negative for milk yield and positive for protein concentration indicated that the latter was, at least in part, a concentration effect. After using a linear regression model of predicted difference in protein yield on predicted difference in milk yield to remove the milk yield effect, the residual protein yield for the 56 herds attributable to differences in genetic merit ranged from -1.4 to +1.3 kg of protein with a mean value of 0.0 kg. The observed differences in protein yield in these herds after correction for milk yield ranged from -0.9 to +0.8 kg, again with a mean value of 0.0 kg.

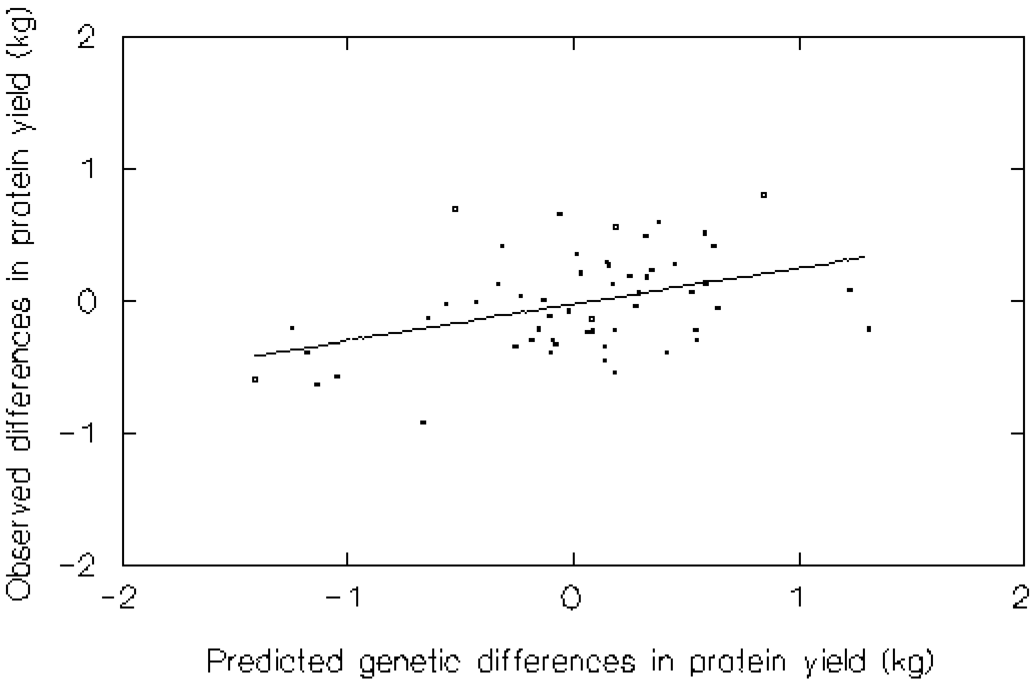

Figure 5 : Relationship between observed and predicted differences in protein yield.

A scatterplot of the relationship between the observed differences in protein yield and those predicted from genetic assessment is shown in figure 5. While there is a significant linear association between the observed and predicted differences (P =0.001, r = 0.42) , the magnitude of the effect is less than expected. The slope of the regression line was 0.275.

The characteristics of the 36 Spring-calving herds for which genetic data was available (in greater than 50 % of the cows in the herds) are shown in Table 7. Herd size ranged from 37 to 214 cows with milk yields, protein yields and protein concentrations ranging from 4165 to 6772 kg per cow, 140.3 to 208 kg per cow and 31.2 to 33.7 respectively. Unfortunately the range in RBI95 and predicted differences for protein concentration and protein yield were not large.

The correlation coefficients between protein concentration, protein yield and other variables are shown in Table 8. Protein concentration was positively correlated with fat concentration though not as strongly as in the previous data sets. Protein concentration and the percentage of milk produced between April and October was more strongly correlated (r = 0.39) than in the previous data sets and again there were negative correlations between protein concentration and milk yield per cow and level of concentrates fed. The positive correlation between protein concentration and milk produced between April to October would suggest that as more milk is produced on pasture the annual average protein concentration is higher.

Table 7 : The characteristics of the Spring - calving herds with genetic data (n=36)

	Minimum	Maximum	Mean
Milk April - October (%)	61.1	97.6	73.2
Yield (kg/cow)	4165	6772	5719
Concentrate fed (kg/cow)	397	1029	692
Stocking Rate (lu/ha)	1.95	3.43	2.69
U.M.E. (G.J./ha)	38.5	48.1	44.8
Calving day (day from Jan.1)	36	119	58
Calving January - April (%)	55.0	100	94.3
Number of cows	37	214	70
Milk fat (g/kg)	3.43	3.89	3.65
Milk fat per cow (kg)	149.5	259.6	209.1
Milk fat per ha (kg)	325.1	710.4	565.4
Milk protein (g/kg)	3.12	3.37	3.23
Milk protein per cow (kg)	140.3	208.0	184.4
Milk protein per ha (kg)	300.4	665.7	498.7
RB195	93	106	100
Predicted Diff. Milk	-205	25	-61.5
Predicted Diff Fat (g/kg)	-0.1	0.8	0.3
Predicted Diff. Fat (kg)	-4.38	4.16	-0.85
Predicted Diff. Protein (g/kg)	-0.1	0.4	0.2
Predicted Diff. Protein (kg)	-5.72	1.83	-1.09

Table 8 : The correlation coefficients between protein concentration (g/kg), protein yield (kg/cow) and a number of other variables in the herds for which genetic data was available (n=36)		
	Protein %	**Protein Yld**
Fat (g/kg)	0.25	0.36
Protein (g/kg)	1.00	-0.12
Milk April - October	0.39	-0.61
Yield per cow (kg)	-0.29	0.99
Concentrate fed (kg)	-0.16	0.60
Stocking Rate (lu/ha)	0.18	0.17
UME (G.J./lu)	-0.25	0.84
Calving day	0.45	-0.59
Predicted Diff. Milk (kg)	-0.15	0.11
Predicted Diff. Fat (g/kg)	0.16	-0.11
Predicted Diff. Fat Yield (kg)	-0.03	0.05
Predicted Diff. Protein (%)	0.35	-0.04
Predicted Diff. Protein Yield (kg)	-0.04	0.08
RB195	0.07	0.06

This would agree with previous research findings as would the positive correlation with calving day (r = 0.45). The negative correlation with milk yield would also be expected due to a dilution effect but the negative correlation with the level of concentrates fed is more difficult to explain. In controlled experiments there was a consistent positive relationship between milk protein concentration and level of concentrate fed with grass silage. However, on farms the increased concentrate may be used to supplement pasture or it may be compensating for poorer quality silage. As would be expected protein concentration was positively correlated with the predicted difference for protein concentration; the size of the correlation (r = 0.35) would suggest that the genetic effects are modulated by environmental factors. Protein yield was highly positively correlated with milk yield (r = 0.99). This was observed in all the data sets analysed which suggests that by increasing milk yield protein yield will also be increased. In the absence of quota restrictions this would be the obvious approach to take in order to increase protein production on the farm.

The proportion of the variation in fat concentration, fat yield per cow (kg), fat yield per ha (kg), protein concentration, protein yield per cow (kg) and protein yield per ha (kg) accounted for in the regression model when genetic data was included (using the 6 independent variables used previously plus RBI95 and the [1]appropriate predicted difference) is shown in Table 9.

Table 9 : The proportion of the variation accounted for in the regression model applied to the Spring - calving herds for which genetic data was available (n=36)	
Response Variate	**Proportion of Variation Accounted for**
Fat (g/kg)	.243
Fat/cow (kg)	.958
Fat/ha (kg)	.988
Protein (g/kg)	.255
Protein/cow (kg)	.974
Protein/ha (kg)	.987

Approximately 25% of the variation in protein concentration is now accounted for as opposed to less than 4% when genetic data was not included. Again over 97% of the variation in protein yield per cow or per ha was accounted for by the model. The proportion of variation in protein concentration and yield accounted for in the 20 Autumn-calving herds for which genetic data was available was similar (results not shown)

Quartile Analysis

The analysis of variance of variables in the 4 quartiles based on protein concentration for the 1995 data is given in Table 10.

Table 10 : The 128 Spring - calving herds in 1995 divided into 4 quartiles with approximately equal ranges of protein concentration.

	Q1	*Q2*	*Q3*	*Q4*	*SED*	*F-test*
N	37	40	26	25		
Protein (g/kg)	31.6	32.2	32.6	33.3	0.006	***
Milk A.0 (%)	73.3	72.5	72.3	76.6	1.33	*
Conc.Fed (kg/cow)	705	740	709	618	46.2	P=0.07
Stocking Rate (lu/ha)	2.63	2.63	2.72	2.63	0.11	NS
UME (G.J./lu)	43.9	44.4	44.4	43.5	0.84	NS
Calving Day	52	48	50	59	2.7	***
Fat (g/kg)	35.8	36.4	37.0	37.3	0.03	***

	Q1	*Q2*	*Q3*	*Q4*	*SED*	*F-test*
Fat Yield (kg/cow)	199.3	209.3	209.6	196.2	6.1	P=0.06
Fat/ha (kg)	527.1	550.8	573.0	516.7	29.7	NS
Protein Yield (kg/cow)	175.8	185.1	185.1	175.4	5.2	P=0.08
Protein/ha (kg)	464.1	486.8	506.1	460.9	25.7	NS
Milk Yield (kg/cow)	5565	5747	5672	5270	163.3	*

The average protein concentration in the 4 quartiles were 31.6, 32.2, 32.6 and 33.3. The percentage of milk produced between April and October was higher, the calving day was higher (later in the year) and the fat concentration was higher in the highest protein concentration quartile compared to the lowest. Milk yield was lowest in the highest protein concentration quartile. When a similar analysis was applied to the 1994 and 1996 data (Tables 11 and 12 respectively) the numerical trends were similar for the percentage of milk produced in April to October and calving day but the differences were not significant. Again fat concentration was significantly higher in the highest protein quartile compared to the lowest. In 1994 and 1996 protein yield per cow was significantly higher in the highest protein concentration quartile compared to the lowest. These analyses would again indicate that higher protein concentration in herds is associated with later calving, and a greater proportion of milk being produced on pasture (April to October)

Table 11 : The 100 Spring - calving herds in 1994 divided into quartiles on the basis of protein concentration

	Q1	Q2	Q3	Q4	SED	F-test
N	21	29	23	27		
Protein (g/kg)	31.4	32.1	32.6	33.2	0.07	***
Milk A.0 (%)	76.3	77.6	76.3	77.6	1.53	NS
Conc. Fed (kg/cow)	595	634	577	608	52.9	NS
SR (la/ha)	2.64	2.71	2.79	2.81	0.12	NS
UME (G.J./lu)	43.4	42.1	43.0	43.6	0.73	NS
Calving day	49	47	47	52	3.4	NS
Fat (g/kg)	35.1	36.3	36.7	36.8	0.28	***
Fat Yield (kg/cow)	182.8	182.4	186.1	193.6	6.46	NS
Fat/ha (kg)	486	493	520	545	29.6	NS

Table 11 : The 100 Spring - calving herds in 1994 divided into quartiles on the basis of protein concentration						
	Q1	Q2	Q3	Q4	SED	F-test
Protein Yield (kg/cow)	163.4	161.5	165.4	174.7	5.4	*
Protein/ha (kg)	433.7	436.7	481.9	492.3	25.7	P=0.07
Milk Yield (kg/cow)	5204	5022	5068	5265	167	NS

Table 12 : The 112 Spring - calving cows in 1996 divided into 4 quartiles on the basis of protein concentration						
	Q1	Q2	Q3	Q4	SED	F
N	31	28	27	32		
Protein (g/kg)	31.2	31.9	32.2	32.8	0.011	***
Milk A - O (%)	78.9	77.7	79.5	77.0	1.36	NS
Conc. fed (kg/cow)	513	656	628	624	52.9	*
SR (lu/ha)	2.57	2.48	2.46	2.69	0.10	P=0.07
UME (GJ/LU)	44.2	43.5	44.1	44.1	0.68	NS
Calving day	50	50	55	51.4	2.50	NS
Fat (g/kg)	35.7	36.6	36.7	37.3	0.028	***
Fat Yield (kg/cow)	184.4	196.8	199.2	202.3	5.58	**
Fat/ha (kg)	477.7	488.0	490.6	545.9	25.5	*
Protein Yield (kg/cow)	161.3	171.1	174.8	177.6	4.48	**
Protein/ha (kg)	417.2	424.5	430.4	479.2	21.5	*
Milk Yield (kg/cow)	5171	5368	5419	5415	142	NS

Factor Analysis

From the database of approximately 300 herds per year, information on farm management practices was available from 170 herds for each of the years 1994, 1995 and 1996. A total of 18 variables relating to farm practices were used in the analysis, a list of which is given below in table 13. For each variable, a mean value for the three year period was calculated. A common factor analysis model using Systat computer software, (Wilkinson and Hill, 1994) was then used to condense the information contained in the 18 original variables into a smaller set of dimensions (factors).

The underlying assumption of factor analysis is that there exists a number of unobserved latent variables or factors which account for the correlations among observed variables. In other words, each observed variable can be expressed as a weighted composite of a set of latent variables. Apart from being a data reduction technique, factor analysis can provide an insight to any conceptual structure among the variables by an examination of the number of factors extracted and their relationships with the observed variables.

Table 13 : Names and descriptions of on-farm variables used in factor analysis

Variable Name	Description
CALVOCTD	Percent of herd calving from October to December
CALVJANA	Percent of herd calving from January to April
FDCOSTGA	Feed costs per gallon
CONC	Amount of concentrates fed per cow
CALVDAY	Days elapsed from 1st January to mean calving date
MILKAO	Percent of milk produced from April to October
UMEGJLU	Proportion of milk produced from grass or silage per livestock unit
UMEGJHA	Proportion of milk produced from grass or silage per hectare
YIELD	Average milk yield per cow
PAC	Phosphorus fertilizer usage
KAC	Potassium fertilizer usage
NAC	Nitrogen fertilizer usage
AREA	Area of farm devoted to dairy enterprise
NOCOWS	Herd size
CNCCOST	Concentrate cost per tonne
GALAC	Volume of milk produced per acre

Table 13 : Names and descriptions of on-farm variables used in factor analysis	
Variable Name	**Description**
SRLUAC	Stocking rate in livestock units per acre
INMILK	Cows in milk as a percent of the total herd (a proxy for lactation length)

The correlation matrix of each variable with all other variables was the basis for factor extraction. The variance accounted for by each factor , its eigenvalue, was used to determine its significance and whether it would be retained for further analysis. A total of 6 factors with eigenvalues greater than 1, indicating that the factor explained more of the variance in the original data than would a single variable by itself, were deemed to be significant. Eigenvalues for these six latent variables are shown below in table 14. Together they accounted for 82.4% of the total variance of all 18 variables.

Table 14 : Extraction of components/factors

Factor	Eigenvalue	% of variance	Cumulative % of variance
1	5.36	29.8	29.8
2	3.04	16.9	46.7
3	2.10	11.6	58.3
4	1.87	10.3	68.6
5	1.44	8.0	76.6
6	1.04	5.8	82.4

After extraction of the initial factor solution, the matrix was rotated using the varimax method in order to distribute the variable loadings between the factors thereby making the results more interpretable. Each of the study farms was then assigned scores for the six factors, based on a linear weighted combination of the variables associated with that factor.

Table 15 shows the loadings of the observed variables into the six factors after varimax rotation. The loadings represent the correlation between a given variable and the overall factor. The factors were orthogonal or mutually independent of each other. Factor loadings - equivalent to correlation coefficients - of 0.40 and above are highlighted to indicate the main attributes of the different factors. Due to the relatively large sample size, factor loadings as low as 0.20 were statistically significant.

Factor 1 (Calving Pattern)

Factor 1 was the most clearly defined, with 6 of the 18 original variables loading significantly. Farms with high scores for factor 1 had a high proportion of cows calving in the autumn (CALVOCTD, r = 0.922) and a low proportion between January and April (CALVJANA, r = -0.859). As a consequence, the mean calving date was early (CALVDAY, r = -0.842) with a low percentage of milk produced between April and October (MILKAO, r = -0.707). Concentrate feeding was high (CONC, r = 0.419) with correspondingly high feed costs per gallon (FDCOSTGA, r = 0.405).

Table 15 : Loadings of observed variables in rotated factor matrix

Variable Name	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
CALVOCTD	**0.922**	0.086	0.107	-0.010	-0.168	0.095
CALVJANA	**-0.859**	0.030	-0.119	0.060	0.242	-0.150
CALVDAY	**-0.842**	-0.180	-0.145	-0.007	-0.071	-0.216
MILKAO	**-0.707**	-0.212	-0.057	-0.021	-0.284	-0.371
SRLUAC	0.139	**0.949**	0.109	0.052	-0.148	0.048
UMEGJHA	0.020	**0.939**	0.044	0.032	0.292	-0.026
GALAC	0.225	**0.805**	0.031	0.084	0.274	**0.427**
NAC	0.074	**0.512**	0.031	**0.494**	-0.089	0.085
NOCOWS	0.247	0.114	**0.875**	0.072	-0.018	0.089
AREA	0.225	-0.146	**0.875**	0.056	0.090	0.039
CNCCOST	0.097	-0.195	**-0.644**	0.004	0.205	0.004
PAC	-0.018	0.046	0.010	**0.929**	-0.024	-0.001
KAC	-0.045	0.053	0.081	**0.915**	0.045	0.034
UMEGJLU	-0.204	0.122	-0.133	-0.033	**0.937**	-0.165
YIELD	0.215	0.176	-0.079	0.085	**0.653**	**0.642**
FDCOSTGA	**0.405**	-0.020	-0.121	0.100	**-0.509**	**0.669**
CONC	**0.419**	0.058	0.050	0.118	-0.262	**0.814**
INMILK	0.077	0.098	0.112	-0.045	0.050	**0.628**

Factor 2 (Stocking Density)

The main attributes loading into factor 2 were a high stocking rate (SRLUAC, r = 0.949) and high milk production per acre (GALAC, r = 0.805). The proportion of milk per hectare produced from grass or silage was very high (UMEGJHA, r = 0.939) with a high rate of nitrogen usage (NAC, r = 0.512).

Factor 3 (Herd Size)

Factor 3 related to cow numbers (NOCOWS, r = 0.875) and farm size (AREA, r = 0.875). A negative correlation between herd size and concentrate cost meant that the latter variable also loaded significantly into this factor (CNCCOST, r = -0.644).

Factor 4 (Fertilizer Usage)

The variables loading highly into factor 4 were high use of phosphorus fertilizer (PAC, r = 0.929), potassium fertilizer (KAC, r = 0.915) and to a lesser extent nitrogen fertilizer (NAC, r = 0.494). As this factor was not associated with high stocking rate, high milk yields, low concentrate input or any other indicator of intensivism, the high fertilizer usage may have been either remedial or inappropriate.

Factor 5 (Grass-based Milk Production)

The variable with the highest loading in factor 5 was the proportion of milk produced from grass or silage per livestock unit (UMEGJLU, r = 0.937) which gave correspondingly low feed costs (FDCOSTGA, r = -0.509). Milk yield in these herds tended to be high (YIELD, r = 0.653).

Factor 6 (Concentrate-based Milk Production)

The most distinguishing feature of factor 6 was the high level of concentrate feeding (CONC, r = 0.814). This was reflected in increased feed costs per gallon (FDCOSTGA, r = 0.669) and milk yield (YIELD, r = 0.642). Although stocking rate was not particularly high, milk production per acre was high due to the yield effect (GALAC, r = 0.427). The proportion of cows in milk also loaded significantly in this factor (INMILK, r = 0.628), indicating that these herds had significantly longer lactation lengths.

Linear Regression

Factor scores were used as determinants in a linear regression model in order to identify which if any of the factors relating to farm management practices correlated with milk protein concentration. As milk yield had been included as a variable in the factor analysis and milk yield is the main determinant of protein yield, the relationship between protein yield and management factors could not be examined.

The dilution/concentration effect of high and low milk yields on protein concentration was first removed by regressing protein yield on milk yield. The residual protein yield - equivalent to a yield corrected protein concentration - was then used as the dependent variable in a stepwise linear regression model which contained the scores for each of the six significant factors as determinants. Three of the six factors (1,3,and 5) were included in the final regression model at an α-to-enter and α-to-remove of 0.15. There was no constant term. The results of the regression model are shown below in table 16.

Dependent variable : Residual Protein Yield
Adjusted Squared Multiple R: 0.081
Standard Error of Estimate: 3.512

Table 16 : Results of stepwise regression model of milk protein concentration on six factors representing farm management practices.

Variable	Coefficient	Std. error	Std. Coef.	P(2 tail)
Factor(1)	-0.491	0.200	-0.134	0.015
Factor(3)	0.897	0.200	0.244	0.000

Table 16 : Results of stepwise regression model of milk protein concentration on six factors representing farm management practices.				
Variable	**Coefficient**	**Std. error**	**Std. Coef.**	**P(2 tail)**
Factor(5)	0.402	0.200	0.110	0.046

The factor showing the strongest correlation with protein concentration was factor 3, indicating a significant herd size effect (P<0.001). This was followed by factor 1, representing seasonality of calving (P=0.015) and finally factor 5 which was associated with grass-based milk production (P=0.046). Both factors 3 and 5 were positively correlated with milk protein while the negative correlation for factor 1 indicated that autumn calving herds tend to have lower protein levels. Together these three factors accounted for only 8% of the observed variation in milk protein concentration, most of which 5.7% was attributable to factor 3. Therefore not only is the variability in milk protein concentration small, but the proportion of this which can be explained by management factors is minimal. The correlation between herd size and milk protein appeared to have a nutritional basis. Of the 18 original variables, the unit cost of concentrates was the best predictor of milk protein levels, with higher protein associated with cheaper rations. In the factor analysis model, this association appeared as a herd size effect because large herds had lower mean concentrate costs, presumably due to a higher proportion of these herds feeding straights rather than compound rations. However, the rationale for such an effect is unclear.

Conclusions

1. The proportion of the observed variation in milk protein concentration between herds that could be accounted for by analysis of the DairyMis data was less than 10%. Including genetic data on RBI95 and predicted differences for protein concentration and yield increased the proportion of variation in protein concentration that could be explained to 25%.
2. In Spring-calving herds milk protein concentration was associated with later calving more milk produced on pasture (between April and October) and lower milk yields.
3. The proportion of variation in milk protein yield that could be accounted for was approximately 97%. This was highly correlated with milk yield which accounted for most of the variation.
4. Based on the herd variation in the data analysed here (4 g/kg in protein concentration or 12% of the mean) a typical herd could expect to get no more than a 10kg (6%) increase in protein yield by increasing protein concentration. This 10 kg increase in protein yield could be achieved by increasing milk yield by 310 kg (at 32.3 g/kg protein).
5. In the absence of quota limits on milk production increasing milk protein yield per farm would be best achieved by increasing milk yield.

---

[1] *If protein yield was the dependent variable then predicted protein yield was the independent variable etc.*