Dairy Herd Management Types Assessed from Indicators of Health, Reproduction, Replacement, and Milk Production

CARSTEN ENEVOLDSEN,[1] JENS HINDHEDE, and TROELS KRISTENSEN
Danish Institute of Animal Science, Research Centre Foulum,
DK-8830 Tjele, Denmark

ABSTRACT

Variables related to health, reproduction, replacement, and milk production in 111 Danish dairy herds were studied with factor analysis. The objectives were to identify management types and to assess the relevance of those types for herd milk production. Median herd size and total milk production were 59 cows and 7100 kg of energy-corrected milk, respectively. Based on cow data, 22 herd variables were defined. A factor analysis identified 10 first-order factors and 5 second-order factors. The latter factors were valid indicators of replacement intensity, variability of milk production, potential for peak milk production, disease, and a complex pattern related to herd size and age, cow size, and live cattle sales. The potential for peak milk production, replacement intensity, and variability of milk production were strong predictors of herd milk production. Interactions with herd size were important. The derived factor scoring coefficients allowed assessment of the management type of a given herd.
(**Key words**: factor analysis, systems analysis, production management, health management)

INTRODUCTION

Evaluation of the efficiency of production in a dairy herd is important for producers and production consultants who want to identify economically optimal farm management improvements. Valid and precise indications of promising areas for management improvement may also serve to motivate the producer to change production routines, tactics, and strategies. Farm organization leaders, politicians, or scientists may also want to analyze the production efficiency of a broad spectrum of herds prior to making decisions in the areas of animal welfare legislation, allocation of research grants and extension efforts, and determination of breeding goals.

The identification of weak or strong points of herd management is usually based on relatively systematic comparisons of management indicators, for instance, those indicators related to milk production and reproduction, for herds that have similar opportunities for production. Such herds then serve as reference, or "target", herds. Comparison is relatively simple if it concerns a minor part of herd management, such as mastitis control. An evaluation of management across a large number of herds, however, is difficult for two reasons. First, data may be scarce and not comparable. For instance, pregnancy checks may only be available for a few herds. Consequently, the most precise indicators of reproduction efficiency may not be available on a larger scale, and this lack of data may be related to management. Second, a dairy herd is a complex and dynamic system in which input and output are related to management strategies and animal status in a complex manner (3). Consequently, many management indicators are strongly interrelated, which makes it difficult to specify explicitly those variables that are dependent and those that are independent in a traditional statistical analysis. Therefore, evaluation of herd management indicators is a complicated task.

This study was designed to provide support to the development of management indicators that validly and precisely identify dairy herds that differ markedly from the typical or expected pattern of comparable herds. The study was aimed at utilizing herd data that could be available in any dairy herd enrolled in a typical DHI scheme.

The specific objectives of this study were twofold: 1) to describe the structure of dairy herd management indicators related to health, reproduction, replacement, and milk production, that is, to consolidate related variables into fewer, interpretable organizing concepts or labels to allow a simple, valid, and precise indication of production characteristics; and 2) to assess the relevance of these indicators for milk production, which is the most important source of income in a dairy herd.

Received June 14, 1995.
Accepted February 8, 1996.
[1]Address correspondence to Carsten Enevoldsen, Danish Institute of Animal Science, Research Centre Foulum, PO Box 39, DK-8830 Tjele, Denmark.

MATERIALS AND METHODS

Data

Data from 86 dairy herds constituting a broad spectrum of herd types were made available through collaboration with four Danish practicing veterinarians. Data from 34 herds participating in other studies of farming systems were also available (7). Data from this total of 120 herds were collected through the Danish milk recording scheme. Monthly milk production and SCC were collected for individual cows. Dates of calvings, cullings, and disease treatments were available, and diagnoses at disease treatment were recorded. Breeding values for milk production were also available. Nine herds were excluded because some of the variables described could not meaningfully be derived because the herds were too small (<25 cow yr).

Definition of Herd Variables

Twenty-two variables were calculated for the 111 herds with suitable records, and these variables are described in Table 1 by means of 5, 25, 50, 75, and 95 percentiles. Cut-off dates for calculations varied from April 30, 1993 to December 31, 1994. Variable definitions, measurement characteristics, and the purposes of including the variables in the study follow.

Herd size. Dates of first calvings, purchases, and cullings of individual dairy cows were used to determine herd size. The numbers of heifers and cows that were present at any day in each herd during the last year prior to the cut-off date were summed and then divided by 365. Herd size was thus measured as the mean number of cow years during the last year of observation. Herd size was included because housing system and management strategies and routines often differed with herd size.

Mature BW. Breed information in the cow files was used to estimate mature BW. Jersey, crossbreeds, Danish Red Breed, and Danish Friesians were assumed to have mature BW of 430, 550, 625, and 625 kg, respectively. Calculations were based on all cows that calved during the last 2 yr prior to the cut-off day minus 50 d. Basically, the mature BW variable measured breed. The variable was included because cow size or breed probably influenced several biologic responses, such as feed conversion efficiency and health (e.g., less dystocia and better claw health among Jerseys).

Age at first calving. The age at first calving was estimated from the same cows as were used to measure BW. Dates of birth were available for >86% of all cows in all herds and could, consequently, be calculated very precisely. Age at first calving most likely was a reflection of management strategy and was expected to influence milk production in some way, depending on management of replacement heifers.

Calving interval. Calving interval was based on all calvings in each herd from 3.5 to 1 yr prior to the cut-off day. For these calvings, the interval to the subsequent calving or culling was identified. This interval was analyzed in an event time analysis with censoring (16) by means of PROC LIFETEST (12). Culling was used as a censoring variable. This analysis identified the number of days postpartum by which half of the cows had calved again. This median calving interval was less influenced by the very long calving intervals and by the nonpregnant cows than by the mean calving interval based only on cows calving again (9). Calving interval is generally accepted as a very important herd variable. Pregnancy checks were not performed uniformly in all herds, and some herds used bulls according to unknown strategies (e.g., problem cows or during pasture grazing). Consequently, more precise and actual indicators of reproductive efficiency were not included in the study. A long observation period was chosen because calving interval might have long-term effects, especially when replacements are home reared and because variances of means increased as numbers of observations decreased.

Calf survival. Data used to determine calf survival included all calvings during the 2 yr prior to the cut-off day. Time of death or, alternatively, sale for slaughter or fattening (bull calves) was identified. Bull calves were included because parturition management was expected to be the same for bull and heifer calves. Mortality during the first 6 mo of life for heifers and for bull calves was estimated for each herd by means of an event time analysis; data concerning slaughter, sale for fattening, and study cut-off date were used as censor variables. Typically, files for young stock are kept much less accurately than are cow files. Consequently, the mortality rates in this study were probably underestimated. However, because individual animal identification is mandatory in Denmark and because the largest proportion of mortality occurs peripartum, lack of calf identification after calving was regarded as a valid indicator of death. The long period of observation was chosen to reduce random error and to consider the long-term effects of calf mortality on herd production efficiency. Perinatal mortality was expected to be strongly correlated with postpartum reproduction disorders in cows and heifers. Such disorders appear to decrease production markedly (8).

Cow survival. The same data files as were used to determine BW and age at first calving were used to determine cow survival (2-yr period). The proportion of cows being culled for any reason before 12 wk postpartum was calculated. This relatively short period was chosen because culling later in lactation probably was heavily influenced by numerous factors, such as reproductive efficiency, milk production, and cash flow. Early cullings, before 12 wk postpartum, were most likely to be involuntary and probably of major biological and managerial relevance.

Calving rate. Using the same data file as for cow survival, calving rate was estimated. The number of calvings was divided by the number of cow years during the same period. This variable measured the overall intensity of calvings. Little measurement error was expected for the reasons already stated. Calving rate was primarily expected to be a function of calf and cow survival (health), reproductive efficiency, and replacement policy.

Sales of live cattle. Records of sales and purchases of cows and heifers for reasons other than slaughter during the last year prior to the cut-off date were used to determine the variable sales of live cattle, which was calculated as sales minus purchases of replacement heifers, heifers, or cows. Values of zero indicated a closed herd. Very little measurement error related to purchases was expected because the national animal identification system uniquely identifies animal and herd of birth. Sales were more difficult to separate precisely between slaughter and livestock sales. The variable was expected to be primarily a function of calf and cow survival (health), reproduction efficiency, changes in herd size or milk quota, and replacement policy. Open versus closed herds might have health effects.

Cow age. The mean lactation number during the last 2 yr was cow age. Virtually no measurement error was expected. The variable was expected to be primarily a function of calf and cow survival (health), reproduction efficiency, changes in herd size or milk quota, and replacement and production policy.

Breeding value for milk production. Calculations for breeding value for milk production were

TABLE 1. Five, 25, 50, 75, and 95 percentiles[1] of 22 variables describing 111 Danish dairy herds.

Variable	Percentile 5	25	50	75	95
1 Herd size, cow years	34	46	59	71	118
2 Mature BW, kg[2]	431	605	623	625	625
3 Age at first calving, d	754	798	823	862	939
4 Calving interval, median d to calving or culling	412	393	375	361	346
5 Calf survival, % alive at 6 mo postpartum	77	83	88	92	95
6 Cow survival, % in herd at 12 wk postpartum	88	93	95	96	98
7 Calving rate, calves born per 100 cow years	96	105	111	119	128
8 Livestock sales, no.[3]	−23	0	2	10	25
9 Cow age, mean lactation number	1.9	2.2	2.3	2.5	2.9
10 Breeding value for milk, index[4]	−290	−124	−20	79	241
11 First parity peak milk, kg of ECM[5]	18	20	23	25	27
12 Cow peak milk, kg of ECM[5] (parity ≥3)	25	28	31	33	37
13 First parity persistency, kg of ECM[6]	0.5	1.3	1.9	2.4	3.6
14 Cow persistency, kg of ECM[6] (parity ≥3)	2.7	4.2	4.9	5.7	6.9
15 First parity peak variability (CV), %	6	10	12	16	23
16 Cow peak variability (CV), % (parity ≥3)	6	8	10	13	19
17 First parity persistency variability, 75 to 25 percentile	0.5	1.0	1.0	2.0	2.5
18 Cow persistency variability, 75 to 25 percentile	0.0	1.0	2.0	2.5	4.0
19 Disease-related milk loss, %[7]	4	5	7	9	11
20 Mastitis level, mean annual SCC	163	236	304	354	414
21 Drug use, administrations per 100 calvings	4	29	63	99	145
22 Milk quota constraints, composite score[8]	−0.9	−0.4	0.0	0.4	1.0

[1]The 50th percentile (the median), for instance, is the number at which 50% of the observations have either lower or higher values.

[2]Estimated from breed means.

[3]Sales minus purchases per 100 cow years.

[4]Deviation from sire breed means in the entire data file.

[5]Kilograms of ECM (energy-corrected milk) at 4 wk postpartum.

[6]Kilograms of ECM decline per 100 d postpartum.

[7]Disease-related milk loss as a percentage of mean herd production capacity.

[8]Score derived from data related to changes in policies for drying off and culling over the quota year.

made from the breeding value for milk production for the sire of each lactating cow present in the herd during the last year prior to the cut-off date. The variable was calculated as the mean deviation from the sire breed means in the entire data file.

Eight variables related to lactation curve characteristics were derived. All of these variables were based on an approach that identified cows without evidence of disease symptoms and estimated variables within herd and within test day (2). The method allowed description of seasonal variation in lactation curve parameters and estimation of the milk production loss associated with records of disease symptoms. All of the milk production variables were based on individual cow test day records from the last 13 monthly test days prior to the cut-off day. Variables were estimated separately for first lactation heifers and cows in third or later lactation. The variables were expected to be functions of feeding and management strategies in particular. Variables might also be risk factors for disease. The potential effects of variability in these variables was of particular relevance.

Peak production of heifers and cows. Two variables, peak milk production of heifers and peak milk production of cows, were estimates of milk production, given that no detrimental effect of disease was revealed. The two variables were calculated as the means of the expected milk production (kilograms of energy-corrected milk) at each test day at wk 4 postpartum for a heifer and a cow, respectively.

Production persistency. The two variables, persistency of first parity production and persistency of cow production, were estimates of slopes of lactation curves, given that no detrimental effect of disease was revealed. Persistency of milk production was calculated as the means of the slopes of the lactation curves for first parity heifers and cows, respectively, at each test day. The units were kilograms of decline in energy-corrected milk per 100 d during the entire lactation.

Peak variability. Two variables, peak variability of first parity heifers and peak variability of cows, were defined as the coefficients of variation of the 13 peak values that were estimated for both first parity heifers and cows, respectively.

Persistency variability. The two variables, persistency variability of first parity heifers and persistency variability of cows, were defined as the differences between the 75 and the 25 percentiles of the 13 values for slopes of the lactation curves for both first parity heifers and cows.

Disease-related milk loss. The sum of differences between expected and observed milk production for all cows as a percentage of the total expected herd milk production defined disease-related milk loss. The calculation methods (2) primarily aimed at identifying acute depressions of milk production with a low rate of false positives. Consequently, minor depressions of milk production, for instance, those that were due to diseases such as subclinical mastitis, probably would not be quantified.

Mastitis level. Mastitis level was defined as the sum of somatic cells in the milk of the individual cows and heifers at all test days during the last year (SCC per milliliter times kilograms of milk produced per cow) divided by kilograms of milk produced per herd during the same period. This measure combined the frequency and severity of udder inflammation of all cows in the herds. Frequency and severity were regarded as the major determinants of milk production loss from mastitis.

Drug use. Drug use was defined as the total number of drug administrations to cows and heifers that were recorded during the last year prior to the cut-off date minus 50 d divided by the number of calvings times 100 for the same period. Drug use was measured in relation to calvings because calving and early life are high risk periods for most diseases. The involved veterinarians consistently recorded their own drug administrations to first parity heifers and cows with animal identification. Records for treatments of young stock were kept less accurately. Danish legislation does not permit dairy producers to apply injections of antibiotics, hormone-like substances (e.g., prostaglandins and corticosteroids), and some other drugs to cattle. Producers are not allowed to administer intramammary and intrauterine drugs either. Consequently, most drug administration was probably measured. However, no information was available on the criteria applied for treatment. Some producers might request treatment for minor symptoms, but others might want to treat only very severe symptoms of disease. Such differences should, however, be reflected in variables that pertain to milk loss from disease or mastitis.

Milk quota constraints. Since 1984, Danish milk producers have been producing milk under quota constraints, which might have affected management strategies. Whether the milk quota actually affected herd management was difficult to assess. A score was created from the data to assess whether the producers tended to dry-off or cull higher producing cows earlier toward the end of the quota year.

The 22 variables were all normalized with PROC RANK (12). The Blom option for normalization was applied.

Statistical Analyses

Interrelations among the 22 herd variables were described by a second-order common factor analysis using iterative maximum likelihood procedures. If no other references were given, the following description of factor analysis was based on Rummel (11). All observed variables are regarded as dependent variables, and the aim of factor analysisis is to identify a number of unobserved common factors that explain the observed variables. The structural model for factor analysis would be a simultaneous equation. In comparison, structural models for regression and ANOVA used single equations (4).

A major result of factor analysis is the separation of the total variance of each variable into its common and unique components, which are also termed the communality and the uniqueness of a variable, respectively. The unique variance again consists of specific variance and random error that usually cannot be separated. The common and specific variance components together constitute the reliable variance.

The derived common factors are the largest, statistically independent (uncorrelated or orthogonal) patterns of relationships among the variables. The factors are defined by loadings that measure which variables are involved in what factor and to what degree. Loadings are correlation coefficients between variables and factors. The square of a loading is the proportion of variance that a variable has in common with a common factor. For instance, if the loading is 0.50, then the factor explains 0.50 times 0.50, which is equal to 0.25 of the variance of the variable. Although such decision criteria are inherently arbitrary, loadings <0.25 are regarded as insignificant in this study because such a loading explains only about 6% of the variance. The sum of the squared loadings over the factors for a variable is the communality that enables interpretation of the fit of each variable to the factor space. Communalities are usually denoted as heritabilities. Thus, $1 - h^2$ is the uniqueness of the variable. The sum of the column of the squared loadings for a given factor is the eigen value or the amount of variance accounted for by that factor (the factor variance). This value measures the strength of the relationships among variables and factors. The sum of the communalities measures the amount of variance accounted for by all the derived common factors. These different measures of variance must be evaluated in relation to the total variance, which is equal to the number of variables.

Factors may be rotated by several techniques to simplify the factor structure and facilitate interpretation. The goal was to obtain a simple structure, which is defined by the Thurstone criteria (11): 1) each variable should have at least one loading in the factor matrix that is near zero; 2) for a factor matrix of p factors, each column of factor loadings should have at least p variables with loadings near zero; 3) for each pair of columns of loadings (factors), several variables should have loadings that are near zero in one column but not in the other; 4) for each pair of columns of loadings (factors), a large proportion of the variables should have loadings that are near zero in both columns; and 5) for each pair of columns of loadings (factors), only a small proportion of variables should have loadings that are not zero in both columns.

Several rotation techniques allow orthogonality to be maintained, but the oblique rotation techniques produce correlations among common factors. Simple structure is usually better obtained with oblique rotations.

One output from a factor analysis is a set of scoring coefficients that allows calculation of factor scores. That set is a compilation of score values for each observation (in this case each herd) on each factor. If, for instance, several reproduction variables are strongly correlated and comprise a common factor that is labeled as reproductive efficiency, each herd can be assigned a score for this unobserved new variable. Factor scores are normally distributed with a mean of zero and usually have unit variance. The multiple correlation coefficients of estimates of regression factor scores of the original data indicate the indeterminancy of the factor scores for the data. Multiple correlations below approximately 0.80 would denote, for instance, a high degree of indeterminancy of factor scores. Alternatively, factor scores can be estimated from the values of one or more variables with high loadings on the factor.

The labeling of factors is an important aspect of the interpretation of factor analysis results. The criteria for naming factors usually are descriptive, causal, or symbolic. A meaningful label facilitates visual representation as an aid to understand and communicate the results.

If >1 common factor is identified in a factor analysis and an oblique rotation has produced correlations between the factors, the interrelations between these first-order factors can be described by a second-factor analysis based on factor scores or interfactor correlations from the first-order factor analysis. This second-order factor analysis then identifies a second set of common factors that explains the variation in the obliquely rotated first-order factors.

The PROC FACTOR of SAS (12) was applied to the factor analyses. The varimax and promax options were applied to produce orthogonal and oblique rotations, respectively. Factor scores were estimated from standardized regression coefficients.

The number of factors to extract from the data is a critical and somewhat subjective decision in factor analysis. In this study, the decision was made from evaluations of the scree plots of the preliminary eigenvalues; chi-square tests of the null hypothesis that the number of factors was sufficient versus the alternative, that more factors were needed; and residual analyses (12). Goodness of fit of the final factor models was evaluated from a visual inspection of the residuals and was evaluated numerically from the guideline stating that fit is regarded as acceptable when the standard deviation of the residuals is less than the standard error of a zero correlation for a given n (11). In the actual study, the value of this reference point was 1 divided by the square root of 111 minus 1 equal to 0.095.

Factor analysis efficiently identifies and describes structure in data, but the technique does not readily provide quantitative estimates to assess the importance of the factors. Such estimates were provided by ordinary least squares regression analysis (12) of the relationship between the total milk production (energy-corrected milk) in the herds and the factor scores from the second-order factor analysis (orthogonal solution). This analysis compared otherwise incomparable aspects of management, such as replacement and health, on one scale. Because the herd milk production was the economically most important output and was readily available, this measure was chosen to evaluate the importance of the management types that were identified by the factor analysis.

All two-factor interactions and number of cows squared were included as independent variables in an initial, full regression model. Terms were then eliminated individually through a backward elimination strategy. Terms with probability values >5% were eliminated from the models. Goodness of fit was evaluated from residual plots, and predictive ability was assessed by adjusted R^2 values.

RESULTS

Table 1 shows that several variables were markedly skewed (e.g., herd size, mature BW, and variation in persistency of first parity heifers).

Table 2 shows the correlation matrix of the 22 study variables. The squared multiple correlations of each variable with all other variables are shown on the diagonal. These values were employed as prior communality estimates in the subsequent factor analysis. The highest bivariate correlation was 0.79 between peak production and production persistency of cows (steep slope). Peak production of cows also showed the largest squared multiple correlation with all of the other variables (0.88). Milk quota constraints showed the poorest squared multiple correlation (0.18).

Ten factors were extracted from the data in the first-order factor analysis after comparison with solutions comprising 7 to 12 factors. A probability value of 0.51 from the chi-square test of 10 factors indicated an acceptable solution. The standard deviation of the residuals was 0.028, which was substantially less than the reference value of 0.095. Table 3 shows the unrotated pattern of first-order factors with loadings, final communalities, sum of squared loadings (eigen values), and percentages of common variance. The overall common variance of the 10-factor model was 62%. Mature BW, calving interval, calving rate, breeding value for milk, peak production of cows, and production persistency of first parity heifers all had 100% communality and were representative of Heywood cases (12), which occur often in maximum likelihood factor analysis (10). Iterative principal component analysis and unweighted least squares factor analysis with the same number of factors yielded similar results, however. Therefore, the Heywood cases were not regarded as serious. Milk quota constraint had the poorest communality (15%).

Table 4 shows the pattern for obliquely rotated first-order factor (promax rotation) and the interfactor correlation matrix with squared multiple correlations on the diagonal. Factor 1 had the highest sum of squares with highest loadings on variables for peak production of first parity heifers and cows and lack of persistency of cow production. The variables for peak variability and persistency variability of both cows and first parity heifers explained the majority of the variance of factor 2. When loadings of <0.25 were regarded as insignificant, 14 variables were loading on one factor only. The oblique rotation clearly produced a simpler structure. Because several variables were loading significantly on one factor only, the third Thurstone criterion for simple structure was not met acceptably. The 10 first-order factors were labeled as shown in Table 5 and in Figure 1. A few examples are described in detail to aid interpretation.

Peak production of cows, production persistency of cows (steep slope of the lactation curve), and peak production of first parity heifers were strongly and positively related. Together these three variables accounted for the majority of the variance of first-order factor 1, and they all had high communalities. Mature BW was also positively related to this factor. Because

TABLE 2. Product moment coefficients of correlation among 22 herd level variables from 111 dairy herds.[1]

Variable	Variable number																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 Herd size	**42**																					
2 Mature BW	−5	**57**																				
3 Age at first calving	−17	18	**48**																			
4 Calving interval	17	−35	−40	**75**																		
5 Calf survival	−27	24	8	−21	**37**																	
6 Cow survival	−10	3	−8	−7	6	**35**																
7 Calving rate	15	−22	−25	65	−24	−41	**66**															
8 Sales of live cattle	−25	−6	−18	34	28	−3	9	**42**														
9 Cow age	−16	−20	−22	31	8	6	0	31	**40**													
10 Breeding value	8	1	−23	−9	11	7	−7	10	−15	**41**												
11 First parity peak	−18	47	15	−29	36	−5	−28	−1	2	18	**79**											
12 Cow peak	−21	52	2	−13	34	−1	−15	11	7	31	73	**88**										
13 First parity persistency	−13	15	−14	−5	14	2	−1	4	13	5	54	24	**57**									
14 Cow persistency	−16	49	11	−2	14	−3	1	4	−1	25	56	79	31	**75**								
15 CV of Parity peak	−23	−10	28	−25	−2	−1	−14	−8	−5	9	−3	−4	−16	0	**44**							
16 CV of Cow peak	−30	−7	27	−19	−3	12	−23	−12	−9	0	−4	−9	−21	−6	49	**54**						
17 Persistency IQ, first parity[2]	−8	−17	11	13	−8	6	−1	2	17	4	0	14	−17	14	30	22	**32**					
18 Persistency IQ, cows	−9	10	15	4	8	−3	5	−7	−13	19	11	27	−7	30	32	46	27	**47**				
19 Disease milk loss	7	−8	1	9	−4	−4	4	−15	−2	−11	3	18	−12	11	17	17	8	21	**39**			
20 Mastitis level	15	10	4	−12	−17	14	−9	−21	2	−6	−14	−5	−21	−2	14	13	−1	0	26	**28**		
21 Drug use	3	22	−22	−5	−2	4	−6	−4	17	15	20	46	9	33	5	−10	12	4	33	22	**47**	
22 Milk quota	−9	10	12	3	17	−9	5	11	0	−1	0	2	−11	−6	12	14	−6	4	−1	4	−7	**18**

[1]Correlation coefficients multiplied by 100. Elements in the principal diagonal are squared multiple correlation coefficients of the variable with all of the others.

[2]IQ = Difference between 75th and 25th percentiles.

TABLE 3. Unrotated, first-order factor pattern, communalities (h^2), sum of squared loadings (SSL), and percentage of common variance accounted for by the factors (F).

Variable[1]	Factor loading										h^2
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	
	(×100)										
1 Herd size	26	11	–11	–9	–11	–10	–34	38	28	–31	55
2 Mature BW	5	1	–29	96	5	1	0	0	0	0	100
3 Age at first calving	–37	–19	–27	13	–2	–5	34	–15	30	0	50
4 Calving interval	51	27	55	–20	–57	7	0	0	0	0	100
5 Calf survival	–8	–13	2	25	20	18	–4	–43	–17	–6	38
6 Cow survival	21	–41	5	4	14	–8	3	10	–7	11	28
7 Calving rate	7	85	16	–17	–46	6	0	0	0	0	100
8 Sales of live cattle	29	–2	26	1	–9	13	–7	–45	–25	26	51
9 Cow age	5	–17	46	–7	–11	4	–9	–5	–29	20	40
10 Breeding value	53	19	–22	–12	62	48	0	0	0	0	100
11 First parity peak	–19	–4	29	56	44	29	4	–12	9	–20	78
12 Cow peak	–10	–2	22	60	21	73	0	0	0	0	100
13 First parity persistency	–10	28	62	32	59	–29	0	0	0	0	100
14 Cow persistency	–1	14	22	57	17	50	21	12	27	26	87
15 CV of First parity peak	–15	–9	–23	–17	12	9	54	9	–17	7	47
16 CV of Cow peak	–6	–24	–19	–13	2	2	74	2	–18	–13	71
17 Persistency IQ, first parity[1]	–1	–14	10	–14	–11	30	30	10	1	10	26
18 Persistency IQ, cows	7	5	–3	9	0	32	59	8	2	–17	50
19 Disease milk loss	–17	–8	12	–3	–17	26	15	42	–13	–19	40
20 Mastitis level	–1	–11	–23	5	–10	–2	4	43	–13	3	29
21 Drug use	–5	–1	10	25	9	36	–15	52	–36	13	66
22 Milk quota	5	3	–10	8	–12	4	13	–22	–18	–12	15
SSL	1.0	1.3	1.6	2.3	1.7	1.7	1.6	1.3	0.7	0.7	13.7
SSL, % of Common variance	5	6	7	11	8	8	7	6	3	2	62

[1]IQ = Difference between 75th and 25th percentiles. Cows were in second and later parities.

of the logical relationships of these variables with high peak production and the lack of strong associations with the health variables, this first-order factor was labeled potential peak production.

Variability in peak production and persistency among both cows and first parity heifers were all positively related to first-order factor 2. Consequently, the factor name production variability was an obvious choice. Communalities were relatively low for most of the variables involved.

In general, the factor pattern allowed relatively simple labels to be assigned. Factor 8 was an exception. The correlations showed that, as herd size increased, fewer cattle proportionately were sold live (or more were bought), and herd age decreased (cows were younger). The correlation with calf survival was 0.24; that is, fewer calves survived. This complex set of relations made sense biologically but was difficult to describe with one simple label.

The strongest interfactor correlation was –0.46 between first-order factors 4 and 10. The squared multiple correlations of each first-order factor with all the other factors (on the principal diagonal) varied between 5% for factor 5 and 36% for factors 4 and 10. These interfactor correlations produced correlated factor scores that served as input to the second-order factor analysis.

Five factors were extracted from the factor scores that were produced from the first-order factor analysis. Table 5 shows the second-order pattern for unrotated factors with loadings, communalities, percentage of common variance, and eigenvalues. A probability value of 0.82 from the chi-square test of five factors was sufficient to indicate an acceptable solution. The standard deviation of the residuals was 0.015, which was substantially less than the reference value of 0.095. The overall common variance of the five-factor model was 55%. The first-order factors 1, 2, and 6 all had 100% communality. The first-order factor that was labeled genetic production potential had the poorest communality (5%). The second-order factor analysis also produced Heywood cases.

Table 6 shows the pattern for orthogonally rotated, second-order factor (varimax rotation). This solution served to produce factor scores for the five second-order factors. The third simple structure criterion was not met satisfactorily.

In addition, an oblique rotation was performed in the second-order factor analysis (Promax). Figure 1 shows the labels assigned to the five factors that were extracted in the second-order factor analysis and the overall data structure revealed by the first- and second-order factor analyses together. Except for the third criterion, the five Thurstone criteria for simple structure appeared to be reasonably well satisfied. However, second-order factor number 5 contained no really strong loadings and had relatively low total sums of squares, which could indicate low validity of the factor. The relationships showed that larger cows (larger breeds) in the herds occurred together with poorer reproductive efficiency and the complex of less sales, younger cows, and larger herds. Despite the complexity, the relationships were sufficient to make retention of the factor worthwhile.

The matrices of regression coefficients to estimate factor scores for first- and second-order factors are shown in Tables 7 and 8, respectively. These tables also show the multiple correlation coefficients of estimates of regression factor scores of the original data. Multiple correlations below the critical limit of 0.80 were found for first-order factors 6, 8, and 10 and for second-order factor 5. Factor scores for these common factors have, consequently, the highest degree of indeterminancy.

Table 9 shows the final regression model of the relationships among total herd milk production per year and those terms that were retained in the model. Five two-factor interactions were significant. The adjusted R^2 value was 97.5%, and the coefficient of variation was 6.5%. Herd size was clearly the strongest determinant of herd milk production. For comparison, the R^2 of a model with herd size as the only variable was 93.4%, and the coefficient of variation was 10.6% ($y = 40621 + 5997x$). The overall mean for milk production per cow in each herd was 7100 kg of energy-corrected milk. The interaction terms showed

TABLE 4. First-order factor pattern from oblique rotation (Promax), sum of squared loadings (SSL), and correlations among factors with squared multiple correlations of the factors with all of the other factors on the diagonal.

Variable	Factor loading									
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
	(×100)									
1 Herd size	−8	−24	12	−7	12	3	−6	62	3	−10
2 Mature BW	28	−4	−4	−7	−10	9	0	3	78	21
3 Age at first calving	17	16	−32	0	−26	−28	−13	1	2	−9
4 Calving interval	1	5	83	0	−7	−10	−1	3	−2	1
5 Calf survival	21	−4	−7	2	4	−10	−4	−24	3	34
6 Cow survival	−8	7	14	−49	6	4	4	−2	8	−10
7 Calving rate	−7	0	24	70	−1	3	4	−3	1	6
8 Sales of live cattle	2	−7	35	−6	10	−13	−9	−47	4	18
9 Cow age	0	−5	30	−15	−17	16	7	−33	−14	1
10 Breeding value	10	10	−5	−4	97	−4	−2	9	−6	0
11 First parity peak	61	1	−8	−3	−1	−6	28	7	0	13
12 Cow peak	83	−5	−1	5	6	16	−10	−7	1	8
13 First parity persistency	10	0	3	−1	−1	−4	88	−4	0	−12
14 Cow persistency	73	4	4	−1	5	−1	5	−7	14	−29
15 CV of First parity peak	−14	53	−20	4	10	13	2	−17	−6	0
16 CV of Cow peak	−16	76	0	−12	0	1	4	−7	2	17
17 Persistency IQ, first parity[1]	19	25	11	−7	−1	4	−15	−8	−19	−12
18 Persistency IQ, cows	23	58	14	4	11	−2	1	9	5	11
19 Disease milk loss	16	22	7	4	−18	38	−5	18	−17	2
20 Mastitis level	−14	10	−7	−8	−4	37	−9	11	16	−7
21 Drug use	18	−4	−4	1	6	70	1	−6	5	−5
22 Milk quota	−6	17	9	10	−1	−2	−8	−10	15	32
SSL	2.0	1.5	1.2	0.8	1.1	1.0	1.0	0.9	0.8	0.5
	Correlation among first-order factors									
Factor loading										
F1	**31**									
F2	4	**27**								
F3	−11	−18	**34**							
F4	−12	4	42	**36**						
F5	17	−10	8	2	**5**					
F6	25	−5	10	−10	7	**17**				
F7	26	−33	−7	−9	7	−2	**22**			
F8	−21	4	−3	18	−5	6	15	**14**		
F9	19	−1	−17	12	3	−7	−1	12	**18**	
F10	22	−24	−31	−46	3	−6	21	−18	−11	**36**

[1]IQ = Difference between 75th and 25th percentiles. Cows were in second and later parities.

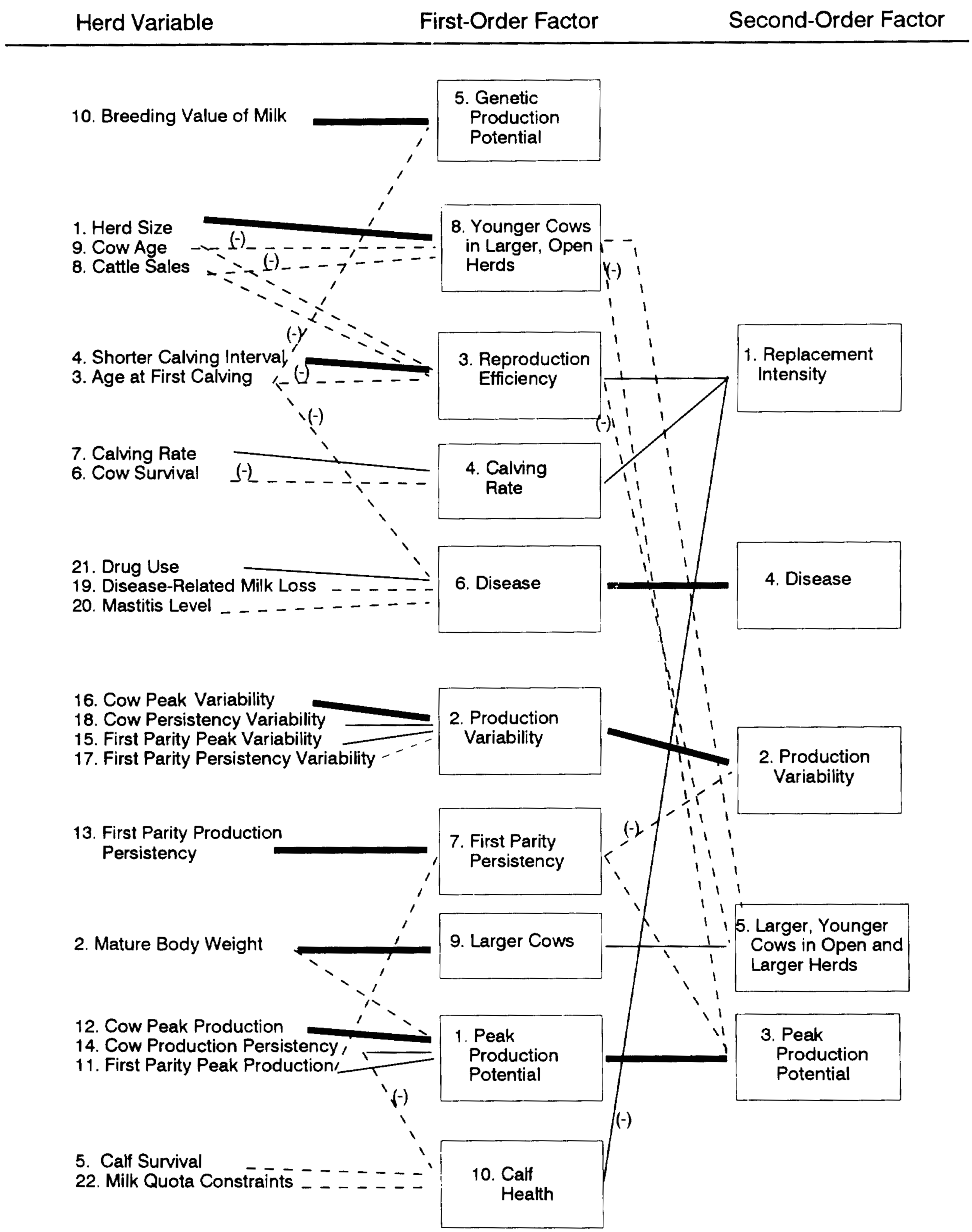

Figure 1. Data structure revealed by first- and second-order factors. Oblique rotations (Promax). Absolute factor loadings 0.75 to 0.100 (▬), 0.50 to 0.74 (——), and 0.25 to 0.49 (- - -).

TABLE 5. Unrotated, second-order factor pattern, communalities (h^2), sums of squared loadings (SSL), and SSL as a percentage of common variance.

First-order factor label	Second-order factor loading					
	F1	F2	F3	F4	F5	h^2
	(×100)					
1 Peak potential	0	100	0	0	0	100
2 Variability of milk	100	4	0	0	0	100
3 Reproduction	–17	–11	12	82	–22	77
4 Calving rate	5	–12	–7	62	35	53
5 Genetic potential	–10	17	3	9	–2	5
6 Disease	–6	25	97	0	0	100
7 First parity persistency	–34	26	–11	–11	–6	21
8 Open, young herds	4	–21	12	2	35	19
9 Larger cows	–2	19	–12	–4	50	30
10 Calf health	–25	22	–13	–48	–37	49
SSL	1.2	1.3	1.0	1.3	0.7	5.5
SSL, % of Common variance	12	13	10	13	7	

that the relationships between replacement intensity and larger, younger cows in open and larger herds varied significantly with herd size.

Because of the complexity of the regression model, Table 10 was provided to show predicted herd milk production, given the combinations of values of explanatory variables that described the effects of the interaction terms in Table 9. The table shows that a large production variability (plus two standard deviation units) was associated with less predicted milk (614,000 vs. 663,000 kg of energy-corrected milk) per 100 cows. There was virtually no predicted difference at 40 cow yr. Table 10 also shows that the highest milk production in a 40-cow herd (302,000 kg, equal to 7550 kg per cow) was predicted with high peak production, high replacement intensity, and a low factor score for larger cows in younger, open, and larger herds. In a 100-cow herd, the highest predicted production was 739,000 kg at a high potential for peak production and low values for the other variables.

DISCUSSION

Data and Indicator Variables

Herd size was approximately 10 cows above the Danish national mean (1). This difference is partly explained by the aforementioned problems in deriving the variables from the smallest herds. The veterinarians also deliberately left out herds that they expected to cease milk production within a few years; virtually all of the excluded herds were smaller herds. Breed distribution in the data (not shown) was very similar to the national distribution (1). The data thus appeared to represent virtually the entire spectrum of

TABLE 6. Orthogonally rotated (varimax) second-order factor pattern and sums of squared loadings (SSL).

First-order factor label	Second-order factor loading				
	F1	F2	F3	F4	F5
	(×100)				
1 Peak potential	–17	5	96	12	17
2 Variability of milk	–1	100	–1	–4	1
3 Reproduction	76	–16	9	7	–39
4 Calving rate	69	5	–4	–9	21
5 Genetic potential	6	–9	18	5	–1
6 Disease	0	0	14	99	–5
7 First parity persistency	–16	–33	26	–6	1
8 Open, young herds	14	4	–27	12	28
9 Larger cows	3	–2	12	–6	53
10 Calf health	–58	–25	19	–10	–22
SSL	1.5	1.2	1.2	1.0	0.6

TABLE 7. Standardized factor scoring coefficients for first-order factors and squared multiple correlations of the variables with the first-order factors.

Variable[1]	First-order factor regression coefficient									
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
1 Herd size	0.02	−0.08	0.00	0.02	−0.03	0.07	0.02	0.42	−0.04	−0.01
2 Mature BW	0.02	−0.06	0.03	−0.04	0.07	−0.02	0.03	0.10	1.13	0.16
3 Age at first calving	0.05	0.10	−0.03	0.03	−0.03	−0.19	−0.02	0.03	0.03	−0.08
4 Calving interval	0.05	0.01	0.97	−0.32	0.01	−0.07	−0.04	0.13	0.00	0.02
5 Calf survival	−0.01	−0.04	−0.01	−0.01	0.01	−0.09	0.02	−0.10	−0.03	0.16
6 Cow survival	−0.01	0.01	0.01	0.00	0.01	0.05	−0.01	−0.04	0.01	−0.04
7 Calving rate	−0.07	0.07	0.01	1.16	0.08	0.00	0.00	0.00	0.20	−0.18
8 Sales of live cattle	−0.03	−0.08	0.01	−0.02	0.04	−0.07	0.00	−0.35	0.01	0.08
9 Cow age	−0.03	−0.04	0.02	−0.02	0.03	0.09	0.00	−0.17	0.01	0.01
10 Breeding value	−0.05	0.01	0.06	−0.02	1.02	−0.07	0.02	0.05	0.14	0.01
11 First parity peak	0.04	0.04	−0.04	0.02	−0.04	−0.21	0.03	0.24	−0.07	0.23
12 Cow peak	0.90	−0.18	0.02	0.01	−0.09	0.46	0.00	−0.13	−0.60	0.68
13 First parity persistency	0.09	−0.11	0.03	−0.07	0.03	0.06	1.01	−0.09	−0.06	0.05
14 Cow persistency	0.11	0.25	−0.01	0.09	−0.02	−0.22	−0.17	−0.17	0.29	−0.97
15 CV of First parity peak	0.00	0.19	0.00	0.00	−0.01	0.06	−0.01	−0.08	0.02	−0.05
16 CV of Cow peak	0.02	0.48	−0.03	0.00	−0.06	0.02	0.01	−0.01	−0.02	0.10
17 Persistency IQ, first parity[2]	0.01	0.07	0.00	0.00	0.00	0.01	−0.01	−0.03	0.02	−0.07
18 Persistency IQ, cows[2]	0.03	0.23	−0.02	0.01	−0.04	−0.03	0.00	0.10	−0.01	0.01
19 Disease milk loss	−0.01	0.08	0.01	0.00	−0.02	0.16	0.01	0.15	−0.02	0.01
20 Mastitis level	−0.01	0.04	0.01	0.00	0.00	0.15	0.00	0.04	0.01	−0.06
21 Drug use	−0.09	−0.02	0.06	−0.04	0.04	0.53	0.01	−0.04	0.01	−0.13
22 Milk quota	−0.01	0.02	0.00	−0.01	0.00	−0.02	0.01	−0.03	−0.02	0.10
Squared multiple correlation	0.99	0.83	1.00	1.00	0.99	0.75	0.99	0.68	0.98	0.71

[1]Variable numbers and labels refer to Tables 1 through 4.

[2]IQ = Difference between 75th and 25th percentiles. Cows were in second and later parities.

Danish dairy herds that could be expected to be productive in the near future.

More precise information about cow size (height and weight) and information about body condition probably would be valuable but would be difficult to obtain at low cost on a larger scale.

Criteria for performing disease treatment and the frequency of chronic (long-lasting) health disorders probably also would be valuable. During the ongoing collaboration with the practitioners, options for collecting this type of information are currently being investigated.

The uniqueness of the indicator variables ($1 - h^2$) was strongly influenced by random or systematic error in measurement. The uniqueness of the milk quota constraint indicated that quota constraints were of little general importance or, probably more realistically, that the developed indicator variable was a poor estimator because of large random error. Further work is needed in this area of management assessment.

TABLE 8. Standardized factor scoring coefficients for second-order factors and squared multiple correlations of the first-order factors with each second-order factor.

First-order factor label	Second-order factor scoring coefficient				
	F1	F2	F3	F4	F5
1 Peak potential	0.00	0.00	1.02	−0.14	0.21
2 Variability of milk	0.00	1.00	−0.02	0.00	−0.19
3 Reproduction	0.57	0.01	0.20	−0.05	−0.54
4 Calving rate	0.31	0.00	0.00	0.01	0.27
5 Genetic potential	0.02	0.00	0.00	0.00	−0.01
6 Disease	−0.04	0.04	−0.14	1.03	−0.03
7 First parity persistency	−0.03	0.00	0.00	0.00	−0.03
8 Open, young herds	0.05	0.00	−0.03	0.01	0.18
9 Larger cows	0.06	0.00	−0.05	−0.02	0.31
10 Calf health	−0.24	0.00	0.01	−0.01	−0.27
Squared multiple correlation	0.80	1.00	0.98	1.00	0.58

TABLE 9. Final regression model of total annual milk production in 111 Danish dairy herds. Effects of five second-order factors and mean number of cows.

Variable[1]	df	Regression coefficient	SE	P
1 Replacement intensity	1	21,020	10,279	0.0435
2 Production variability	1	15,039	7224	0.0399
3 Potential peak production	1	−11,319	7881	0.1541
5 Larger cows in younger, open, and larger herds	1	2775	3961	0.4851
6 Cows in the herd, no.	1	6595	152	0.0001
1 × 3	1	10,794	4219	0.0120
3 × 5	1	−9724	3394	0.0051
1 × 6	1	−474	163	0.0044
2 × 6	1	−421	125	0.0010
3 × 6	1	639	130	0.0001
Intercept	1	8214	9201	0.3741

[1]Second-order factors are defined as standard deviation units. That is, one standard deviation unit increase in replacement intensity is associated with a 21,020-kg increase in herd milk production, which is conditional on the other effects in the model. Dependent mean = 420,752 kg of energy-corrected milk (ECM). Root mean SE = 27,545 kg of ECM (CV = 6.6%). Adjusted R^2 = 97.5%.

Calf and cow survival are other unique indicators that probably were exposed to much random error. Systematic error of measurement might be suspected in the measurement of drug use. In herds with poor health, disease recording might also be poor. This poor record keeping, in turn, affects the estimation of disease-related milk loss, which depends on the identification of healthy cattle. Systematic evaluations of the health states of individual animals performed by the veterinarians at regular herd visits would help to solve that problem.

Analytical Concept, Interpretation of Analytical Models, Model Specification, Validity, and Precision

This study was based on observed (nonexperimental) data. Such data do not meet the requirements for a valid hypothesis test of causal relationships (5). Therefore, the statistical analyses in this study should be regarded as purely descriptive. That is, the statistical analyses serve to identify structures in the data that are as simple as possible but still contain

TABLE 10. Predicted annual herd milk production (energy-corrected milk, ECM) for dairy herds given various combinations of values for second-order factors.[1]

Value of second-order factors, SD units				Predicted milk production, 1000 kg of ECM (total per cow)	
Peak production potential	Replacement intensity	Larger cows in younger, open, and larger herds	Production variability	40 cow-years	100 cow-years
All herds low values					
−1	−1	−1	−1	258 (6.5)	663 (6.6)
Effects					
Production variability					
−1	−1	−1	+1	255 (6.4)	614 (6.1)
Interactions					
−1	−1	−1	0	256 (6.4)	639 (6.4)
−1	−1	+1	0	275 (6.9)	657 (6.6)
−1	+1	−1	0	241 (6.0)	574 (5.7)
−1	+1	+1	0	260 (6.5)	593 (5.9)
+1	−1	−1	0	280 (7.0)	739 (7.4)
+1	−1	+1	0	270 (6.8)	728 (7.3)
+1	+1	−1	0	302 (7.6)	711 (7.1)
+1	+1	+1	0	292 (7.3)	701 (7.0)

[1]Regression coefficients from Tables 7 to 9 were applied for prediction of herd milk production at either 40 or 100 cow-years in a herd conditional on various combinations of second-order factor values given in standard deviation units. The main effect of a two standard deviation-unit increase in production variability at 100 cows is 614,000 − 663,000 = −49,000 kg of ECM, for instance.

most of the original information. Such results of the statistical analyses allow statements such as this one: the highest milk production in larger herds was obtained in herds with a high potential peak production and low replacement intensity.

The study design, therefore, does not allow direct causal interpretation of a regression coefficient with a statement such as this: the effect of a high peak production potential was X kilograms per cow. Such an interpretation assumes randomization of the study units. Instead, causal inference must be based on a biological and technical interpretation of the data structures revealed by the analysis. Predictions from the regression model, similarly, serve to provide a reference value for a given herd. This reference value is then based on all the information from all of the herds in the data. Phrased alternatively, if all of the herds, contrary to fact, were exposed to exactly the same conditions as the actual herd, then the herd result would be X kilograms of milk.

Factor analysis was the major analytical tool in this study. A major criticism of factor analysis is that it contains several subjective elements, such as choice of prior communalities, number of factors, and, in particular, naming of factors. Traditional generalized linear models at first appear to be more objective, primarily because numerous statistical tests are available for hypothesis tests. However, in studies of complex systems such as dairy herds, development of traditional linear models is also a highly subjective process (for instance, specification of dependent versus independent variables, number of interactions to examine, nonlinearities, and model selection strategies). Basically, factor analysis can be regarded as an analytical linear model that is much less restrictive than a traditional linear model with one or a few dependent variables. In addition, factor analysis is one solution to the often severe multicollinearity problems in this type of data. The maximum likelihood method is regarded as scale invariant (10).

Factor analysis has been widely applied in many scientific disciplines during most of the 20th century. Applications in animal and veterinary science, however, are few, but some examples confirm the usefulness of the analytical approach (6, 13, 14, 15).

There was some indication of validity problems related to the factor models. In a few instances, final communality was less than the prior communality (e.g., genetic potential), and final communality should, in general, be higher than prior estimates. Some Heywood cases occurred. The standard deviation of residuals was much less than $1/\sqrt{n-1}$. These problems could indicate overfactoring, but results from alternative estimation options showed comparable factor structures. Factor analysts disagree about the importance of overfactoring (11). There were no convergence problems, which could be regarded as a goodness of fit criterion (10).

The simple criteria for structure were not met completely, and estimates of multiple correlation of regression factor scores were below the critical limit in some instances. Consequently, factor interpretation and factor score estimation was imprecise in some instances (first-order factors 6, 8, and 10 and second-order factor 5). The relationships that determined these factors, consequently, need further study. However, the simple structure criteria apparently are difficult to satisfy in real data (10).

The validity of the final regression model was fully acceptable. The interactions with herd size provided important pieces of information. Compared with other studies of field data, the coefficient of determination was high, but that high value was clearly caused by the inclusion of herd size. However, the decrease in coefficient of variation caused by the inclusion of the second-order factors and the interactions showed that the final model had provided a marked improvement in predictive ability compared with that of a simple model such as mean milk production per cow. This analysis also showed that the latter model was invalid.

Biological and Technical Interpretation of Factor Patterns

The very strong correlations between peak production and lack of persistency probably could be explained by the generally applied feeding strategies in Danish dairies. Usually a short period with flat rate feeding is associated with a higher energy concentration of the ration in early lactation, thus producing a strong correlation between peak production and lack of persistency.

The indicators for variability of milk production were strongly interrelated. This study did not indicate whether these correlations reflected deliberate management routines or managerial flaws. Some herds apply summer grazing when a low concentrate allocation is intended and winter feeding with fodder beets, concentrates, and only small amounts of silage for ad libitum consumption. Such a combination of rations may be economically attractive but will produce a large variation over the year.

The study indicated that the lack of persistency of first parity heifers constituted a separate entity, which could have been due to management of replacement and springing heifers. More detailed recordings are needed to elucidate such conditions. Type, size, age, and body condition measurements around first calving probably could facilitate the interpretation of this and the BW relationships to the factor space.

Genetic potential was not directly related to milk production factors, which might seem surprising, but, because breeding schemes consistently aim at eliminating environmental components, a high degree of uniqueness might be expected. The increase in breeding value with younger age at first calving should be expected because a young age at first calving should reflect the general increase in breeding values of sires.

Cow age, reproduction, mortality, and sales indicators were logically related, but the analysis confirmed that the interrelationships among these variables were complex. Isolated interpretation of the individual indicators was, therefore, hardly possible. The replacement intensity factor appeared to be a valid indicator that efficiently combined the available information while it remained interpretable.

The health indicators were logically related; the lack of relationships with calf survival was unexpected. An explanation could be that most disorders in relation to dystocia were included in the treatment records. A positive association with the peak production factors was also expected. An explanation for the lack of an association could be that those herds with higher peak milk production might also manage emerging disease symptoms more efficiently.

Relations Among Second-Order Factors and Herd Milk Production

Variability of milk production (second-order factor 2) was associated with low herd milk production in larger herds, but no such relationship was revealed in smaller herds. The explanation of this interaction could be that feeding is more variable in smaller herds or that management flaws in larger herds have a larger impact on production because problems may be more difficult to detect and solve.

The potential for high peak production was virtually always associated with the highest milk production per herd and per cow (adjusted to the same herd size), as expected.

The smallest herds with the highest herd milk production were characterized with a high peak milk production combined with a high replacement intensity, which was expected because more cows in the herd were in early lactation when replacement intensity is high.

In smaller herds with low peak milk production, low replacement intensity apparently was a characteristic of high herd production. Because low peak milk production was so strongly related to persistency, this interaction is explainable.

In larger herds, high replacement intensity was a consistent characteristic of lower herd production. A possible explanation could be that high replacement intensity was a necessity in those herds because of a management strategy or management flaws that caused low milk production. When the rate of involuntary culling was high, retention of cows with lower production would possibly be necessary.

Because of the relatively small effects of the factor of larger and younger animals and the relatively high imprecision of the solution for this factor, the association between this factor and milk production is uninterpreted.

The interactions among herd size, potential peak production, and replacement present several interesting patterns or types of herds worth further detailed investigation in the herds. Thorough studies of representatives of these types might reveal weak and strong points of management. The identification of these types of management is the major contribution of this study. The derived coefficients for scoring factors also allow assessment of the type of a new herd or a future status of one of the herds from this study. Such assessments are useful to monitor production.

CONCLUSIONS

Twenty-two correlated indicators of herd health, reproduction, replacement, and milk production were reduced to five new uncorrelated variables through a second-order factor analysis. These five new variables were valid indicators of replacement intensity, variability of milk production, potential peak milk production, disease, and a complex pattern related to herd size and age, cow size, and live cattle sales. Potential milk production, peak replacement intensity, and variability of milk production were strong predictors of herd milk production. Interactions with herd size were important. The derived factor scoring coefficients allowed assessment of the management type of a given herd.

ACKNOWLEDGMENTS

The collaboration of practicing veterinarians B. Elgaard, J. Helbo, K. H. Nielsen, and J. S. Philipsen and the constructive suggestions of the reviewers were greatly appreciated.

REFERENCES

1 Blom, J. Y. 1993. Diseases and feeding in Danish dairy herds. Acta Vet. Scand. (Suppl.)89:17.

2 Enevoldsen, C., and Y. T. Gröhn. 1994. Disease effects on milk

yield in dairy cows: development and reliability assessment of a simplified within herd and within test day approach for estimation. Page 418 *in* Proc. 7th Int. Symp. Vet. Epidemiol. Econ. G. J. Rowlands, M. N. Kyule, and B. D. Perry, ed. ILRAD, Nairobi, Kenya. (ISSN: 0256-5161.)

3 Enevoldsen, C., J. T. Sørensen, I. Thysen, C. Guard, and Y. T. Gröhn. 1995. A diagnostic and prognostic tool for epidemiologic and economic analyses of dairy herd health management. J. Dairy Sci. 78:947.

4 Francis, J. D. 1993. Multidimensional measurement and classification. Lecture notes, Dep. Rural Sociol., Cornell Univ., Ithaca, NY.

5 Greenland, S. 1990. Randomization, statistics, and causal inference. Epidemiology 1:421.

6 Korver, S., G. Averdunk, and B. B. Andersen. 1987. Performance testing of AI bulls for efficiency and beef production in dairy and dual-purpose breeds. Page 142 *in* Proc. Eur. Assoc. Anim. Prod. Sem. Study Commission Cattle Prod. Anim. Genet., Wageningen, The Netherlands.

7 Kristensen, T., M. Vaarst, and E. Steen Kristensen. 1994. Analysis of the production in Danish organic and conventional dairy herds. Page 7 *in* Proc. 10th Int. Fed. Organic Agric. Movement (IFOAM) Conf., Christchurch, New Zealand.

8 Markusfeld, O., and E. Ezra. 1993. Body measurements, metritis, and postpartum performance of first lactation cows. J. Dairy Sci. 76:3771.

9 Morant, S. V. 1989. The statistical analysis of a field study of the reproductive performance of dairy herds. Page 13 *in* Proc. Soc. Vet. Epidemiol. Prev. Med. G. J. Rowlands, ed. Univ. Exeter, Exeter, UK.

10 Morrison, D. F. 1976. Multivariate Statistical Methods. 2nd ed. McGraw-Hill Book Co., New York, NY.

11 Rummel, R. J. 1970. Applied Factor Analysis. Northwestern Univ. Press, Evanston, IL.

12 SAS® User's Guide: Statistics, Version 6.03. Edition. 1988. SAS Inst., Inc., Cary, NC.

13 Schwabe, C. W., H. P. Riemann, and C. E. Franti. 1977. Epidemiology in Veterinary Practice. Lea & Febiger, Philadelphia, PA.

14 Sieber, M., A. E. Freeman, and P. N. Hinz. 1987. Factor analysis for evaluating relationships between first lactation type scores and production data of Holstein dairy cows. J. Dairy Sci. 70:1018.

15 Sieber, M., A. E. Freeman, and P. N. Hinz. 1988. Comparison between factor analysis from a phenotypic and genetic correlation matrix using linear type traits of Holstein dairy cows. J. Dairy Sci. 71:477.

16 Thysen, I. 1988. Application of event time analysis to replacement, health, and reproduction data in dairy cattle research. Prev. Med. 5:239.