UNIVERSITY OF COPENHAGEN

Technical indicators of financial performance in the dairy herd

Kristensen, Erling Lundager; Østergaard, Søren; Krogh, Mogens Agerbo; Enevoldsen, Carsten

Technical Indicators of Financial Performance in the Dairy Herd

E. Kristensen,[*1] **S. Østergaard,**[†] **M. A. Krogh,**[‡] **and C. Enevoldsen**[‡]
*StrateKo Aps, Gartnervaenget 2, DK-8680 Ry, Denmark
†Department of Animal Health, Welfare and Nutrition, Faculty of Agricultural Sciences, University of Aarhus, Research Centre Foulum,
PO Box 50, DK-8830 Tjele, Denmark
‡Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Grønnegaardsvej 2, DK-1870 Frederiksberg C,
Copenhagen, Denmark

ABSTRACT

Monte Carlo simulation was used to predict the long-term financial performance related to the technical performance of dairy herds. The indicators addressed were derived from data collected routinely in the herd. They indicated technical performance that can be affected by the farmer or the consultant, and they were derived from expected cause-effect relations between technical performance and financial performance at the herd level. The study included the indicators shape of lactation curve, reproduction efficiency, heifer management, variation between cows in lactation curve persistency, mortality in cows and calves, dynamics of body condition, and somatic cell counts. Each indicator was defined by 2 or 3 levels, and 2- and 3-factor interactions were included in the simulation experiment, which included 72 scenarios. Each scenario was replicated 200 times, and the resulting gross margin per cow was analyzed as the measure of financial performance. The potential effects of the selected indicators on the gross margin were estimated by means of an ANOVA. The final model allowed estimation of the financial value of specific changes within the key performance indicators. This study indicated that improving the shape of the herd-level lactation curve by 1 quartile was associated with an increase in gross margin of €227 per cow year. This represents 53% of the additional available gross margin associated with all the management changes included in the study. The improved herd-level lactation curve increased the gross margin 2.6 times more than improved reproduction efficiency, which again increased the gross margin 2.6 to 5.9 times more than improved management related to heifers, body condition score, mortality, and somatic cell counts. These results were implemented in a simple "metamodel" that used data extracted from ordinary management software to predict herd-specific financial performance related to major management changes. The metamodel was derived from systematic experiments with a complex simulation model that was used directly for advanced herd-specific decision support. We demonstrated the use of these key performance indicators to forecast the financial consequences of different "what-if" herd management options, with emphasis on herd health economics.

Key words: key performance indicator, benchmarking, financial performance, herd health economics

INTRODUCTION

The financial impact related to changing the levels of input factors (e.g., management changes or changes in housing) in the dairy herd must be documented (Ko-Net-Praksis, 2006). Usually, it is straightforward to estimate the direct costs associated with a change in one input factor, but for the following reasons, it is a very complex task for assessing the financial value of the technical effects of one or more changes in input factors that occur at the same time (Dijkhuizen et al., 1995; Tauer and Mishra, 2006). First, the dairy herd is a very complex system with numerous feedback mechanisms (Enevoldsen et al., 1995; Østergaard et al., 2000). Consequently, simple partial budgeting techniques are invalid in most situations (Dijkhuizen et al., 1995; Ferguson et al., 2000). Next, because of the long generation interval in cattle breeding, several years may pass before changes in individual animal performance (e.g., effects of the rearing period of dairy heifers) affect the financial performance of the herd as a whole (Mourits et al., 1997). During such a long time span, numerous other determinants of financial performance probably change as well (e.g., prices of inputs and outputs from the herd, and herd-level production constraints). Consequently, it is likely to be practically impossible to collect empirical data at the herd level from a sufficient number of herds and years that will allow a valid comparison of financial performance in herds with different levels of input factors of interest.

Still, extensive use of empirical evidence from studies on various relationships and combining the evidence into a modeling framework is a way of estimating the technical and economic consequences of health-control strategies in a dairy herd (Seegers et al., 2003).

Several decision support models for use at the herd level have been developed to solve the problems described above (Ferguson et al., 2000; Shalloo et al., 2004). It can be argued that simulation models lack the credibility of field experiments, but using a simulation model provides an opportunity to explore complex relationships between input factors that cannot be studied in any other way. By using a simulation model, it is possible to keep all input factors and herd-level constraints constant except for the input factor(s) of interest. The Monte Carlo-type models provide estimates of random variation associated with technical and financial output variables. Such estimates are essential for planning interventions (Shalloo et al., 2004). Nevertheless, it has been difficult to develop an analytical model that provides estimates that are perceived as trustworthy by farmers and consultants, including practicing veterinarians. One explanation may be related to the difficulty associated with providing relevant and valid herd-specific input parameters for the models when using a decision support model for herd-specific interventions (Østergaard et al., 2000).

Provided that the relations between measures of financial performance and some (key) indicators of technical performance are consistent and precise, such key performance indicators (**KPI**) may be used as indicators of financial performance (Kaplan and Norton, 1998). Nonetheless, Enevoldsen et al. (1996) found KPI to be correlated. Even if correlations among KPI are accounted for by means of suitable techniques, such as factor analysis or principal component analysis (Enevoldsen et al., 1996), they may not be independently related to financial performance, or the effect may be too small to be distinguished accurately from the very large variance in income caused by other factors (Dijkhuizen et al., 1984). This may lead to double counting of some financial effects of interventions (Østergaard et al., 2000). If the KPI are varied systematically in a simulation experiment (i.e., a sensitivity analysis), where it is possible to identify the existence of interactions (Shalloo et al., 2004) between KPI, then it would be reasonable to interpret the KPI as indicators of financial performance of the herd. The objective of this study was to define and rank technical KPI that were tightly related to long-term effects on the financial performance in dairy herds predicted by means of Monte Carlo simulation.

MATERIALS AND METHODS

Selection of KPI and Study Design

On the basis of experience with herd management and modeling (Enevoldsen et al., 1996; Østergaard et al., 2000, 2005) and theoretical considerations (Kleijnen and Sargent, 2000; Shalloo et al., 2004), 8 herd-level potential KPI were selected, which are described below. The following general criteria were used to select the potential KPI:

1. The level of the variable describing the potential KPI must be likely to be obtained from data collected within a typical herd management program.
2. The level of the variable must vary between herds.
3. The potential KPI must describe a technical variable that can be affected by the farmer or the advisor.
4. A cause-effect relation between the potential KPI and financial performance is plausible.

The potential KPI addressed were 1) shape of the lactation curves (**LC**), 2) reproduction efficiency (**RE**), 3) heifer management (**HM**), 4) variation between cows in lactation curve persistency (**LC-V**), 5) mortality in cows (**MCow**), 6) mortality in calves (**Mcalf**), 7) dynamics of body condition (BCS), and 8) SCC.

The potential KPI were defined in the context of the SimHerd model described by Sørensen et al. (1992) and implemented in the modeling framework presented by Østergaard et al. (2005), with some model modifications to address the current research questions. For each potential KPI, the 75th percentile and the 25th percentile were calculated as found in various Danish standard protocols. The term "high" ($_H$) was defined as applicable to "good farm management" and the term "low" ($_L$) to "pitiable farm management," and these corresponded with the 75th percentile and the 25th percentile. Because of the model design, the term "middle" ($_M$) was calculated as the average of $_H$ and $_L$ so that the numerical distance between $_L$ and $_M$ was equal to the numerical distance between $_M$ and $_H$. When the percentiles were not directly available, $_M$ was defined as the mean of the potential KPI and $_L$ and $_H$ were based on our expectations. The models ensured that the numerical distance between $_L$ and $_M$ was still equal to the distance between $_M$ and $_H$.

For ease of interpretation, $_H$, $_L$, and $_M$ were regarded as quartiles. The selected scenarios represented practically relevant levels of management and associated performance. In the context of the simulation model, gross margin was defined as sales income less variable costs (feed, AI, and other costs) for cows and heifers. "Other costs" included veterinary assistance, medicine, bed-

ding, and milk control. Labor and management costs were not included as variable expenses (Østergaard et al., 2005).

The selected levels of each potential KPI were varied systematically at levels $_H$, $_L$, and $_M$ for LC, RE, and HM and at levels $_H$ and $_M$ for the remaining potential KPI. The potential KPI were combined according to the following initial model specification:

Gross margin = LC$_{(H, M, L)}$ + RE$_{(H, M, L)}$ + HM$_{(H, M, L)}$

+ LC-V$_{(H, M)}$ + MCow$_{(H, M)}$ + Mcalf$_{(H, M)}$ + BCS$_{(H, M)}$

+ SCC$_{(H, M)}$ + all possible 2-factor interactions

+ the 3-factor interaction among LC, RE, and HM.

This initial model represented 72 scenarios of all possible scenarios. Each scenario was simulated 200 times (replicates) with the modified SimHerd model described below.

General Framework of the SimHerd Model

The applied simulation model (SimHerd) was a dynamic, mechanistic, and stochastic model predicting the production and states of a herd over time. Each cow and heifer was described by a state. The states were characterized by identification number, age, reproductive status, parity, DIM, genetic milk yield level, lactation curve parameters for the current lactation, BW, BCS, culling decision, health status on each simulated disease, milk withdrawal, and SCC. The prediction was made weekly for each animal in the herd. The state of the individual animal was updated, and the production and input consumption of the herd were calculated. Input and output of animals from the herd were also simulated. The drawing of random numbers by using relevant probability distributions triggered variation between animals and discrete events such as pregnancy and culling. The production and development within the herd were determined indirectly by simulation of the production and change in state of the individual animal.

Details About the Modified SimHerd Model and Parameterization

LC. The Wilmink function (Wilmink, 1987) was used to describe the fixed part of the lactation curves. The model of daily milk yield of a cow in SimHerd was modified for this study to represent empirically estimated lactation curve parameters more directly in the simulations. This implied that the feed intake was a consequence of energy needed to match the production

level. The new lactation curve model was based on the Wilmink function:

$$Y_{ijk} = W0_{ij} + W1_{ij} \times DIM_{ijk} + W2_j \times \exp(W3_j \times DIM_{ijk}),$$

where Y_{ijk} is the daily milk yield in kilograms of ECM, $W0_{ij}$ is the yield level (intercept), $W1_{ij}$ is the lactation curve slope after peak yield of cow$_i$ in lactation$_j$, DIM_{ijk} is the DIM of cow$_i$ in parity$_j$ at time step$_k$, $W2_j$ and $W3_j$ are parameters for the lactation curve shape until the peak in parity$_j$, and exp is the exponential function. $W0_{ij}$ and $W1_{ij}$ were drawn randomly at each calving from the 2-dimensional normal distribution:

$$N[(W0g\_3_i + W0_j - W0\_3), (SdW0g\_3^2 + SdW0e_j^2), W1_j, SdW1_j^2, \rho_j],$$

where ρ_j is the phenotypic correlation coefficient between yield level and lactation curve slope after peak yield. $W0g\_3_i$ is the permanent part of the yield level of the individual cow$_i$. $W0g\_3_i$ was drawn randomly for individual animal$_i$ at birth from the normal distribution $N(W0\_3, SdW0g\_3^2)$, where $W0\_3$ and $SdW0g\_3^2$ are the mean and variance for the yield level (intercept) in parity 3; $W0_j - W0\_3$ is the fixed effect of parity$_j$ on the yield level; $SdW0e_j^2$ is the environmental variance of the yield level in parity$_j$; and $W1_j$ and $SdW1_j^2$ are the mean and variance of the lactation curve slope after peak yield in parity$_j$.

The data used for parameterization of the model originated from 39 Danish dairy herds, which are described by Thomsen et al. (2007). To parameterize SdW0g_3 and SdW0e$_j$, a certain level of heritability and permanent environment was assumed. In Denmark, heritabilities of 0.43, 0.29, and 0.27 for kilograms of milk were found for parities 1, 2, and 3+, respectively, in Danish Holsteins (Danish Agricultural Advisory Service, 2005–2006). Those for protein and fat were slightly smaller. Jakobsen et al. (2002) reported similar estimates. These heritabilities originated from 305-d lactations. Jakobsen et al. (2002) showed that the heritability was lower in early lactation. For permanent environment, a heritability of 0.35 was used and the permanent environment accounted for 0.15, so that the repeatability accounted for approximately 0.50 of the total variance. The value of SdW0g_3 was fitted to 3.0 kg of ECM and subsequently fitted to the estimated variance components SdW0e$_j$, SdW1$_j$, and ρ_{W0W1j} (Table 1). From the variance components, the resulting repeatabilities of W0 were calculated (heritability and permanent environment) at 0.58, 0.33, and 0.25 for parities 1, 2, and 3+, respectively.

Table 1. Variance components $SdW0e_j$, $SdW1_j$, and ρ_{W0W1j} used to calculate repeatabilities (heritability and permanent environment) of W0 in the Wilmink function[1]

Variance component	$SdW0e_j = \sqrt{(SdW0g\_3^2 - SdW0_j^2)}$	$\rho_{W0W1j} = Cov_{W0W1j}/(SdW0_j \times SdW1_j)$	$SdW1_j = \sqrt{(SdW1_j^2)}$
Parity 1	2.57	−0.62	0.0147
Parity 2	4.31	−0.61	0.0216
Parity 3+	5.15	−0.65	0.0251

[1]The Wilmink function is $Y_{ijk} = W0g\_3_i + W0e_{ij} + W1_{ij} \times DIM_{ijk} + W2_j \times \exp(W3_j \times DIM_{ijk})$, where $W0g\_3_i$ is the permanent part of the yield level of the individual cow$_i$. $W0g\_3_i$ was drawn randomly for the individual animal at birth from the normal distribution $N(W0\_3, SdW0g\_3^2)$, where $W0\_3$ and $SdW0g\_3^2$ are the mean and variance for the intercept (W0).

The levels for LC were derived from Enevoldsen and Krogh (2006), where $_H$ and $_L$ were the averages of the upper and lower quartiles, respectively, from their study of herd-level lactation curves. The lactation curves for parity 3+ in the simulation experiment are illustrated in Figure 1.

RE. The levels of RE were derived from Freudendal and Strudsholm (2003). The end of the voluntary waiting period was set at 42 DIM. In this context, heat detection was defined as the probability of correctly identifying a cow in heat and that the farmer wanted this cow inseminated, and conception rate was the probability of conception following AI as measured by a pregnancy test at 42 d after AI. The levels were defined as products of heat detection and conception rate (Table 2).

HM. Intensified HM usually focused on feeding intensity, increased heat detection, and time to start of AI based on the size of the individual heifer. Heifers would calve at a younger age without a negative effect on milk
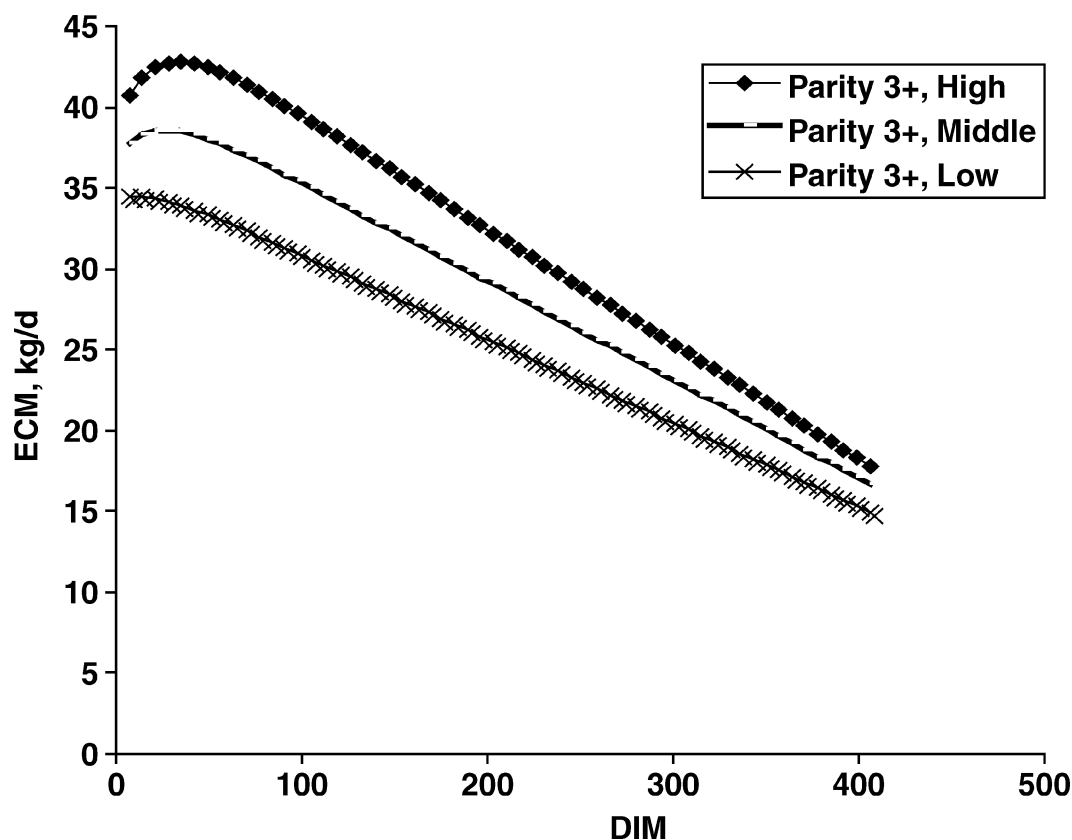

Figure 1. Visualization of the lactation curves at different levels of management for parity 3+ based on the Wilmink function applied in the simulation experiment. High is applicable to "good farm management"; low is applicable to "pitiable farm management"; and middle is the average of H and L.

Table 2. Values of variables that defined levels of reproduction efficiency in the simulation experiment

Reproduction efficiency level	Low	Middle	High
Heat detection rate	0.44	0.50	0.58
Conception rate at 42 d after AI	0.42	0.52	0.62

yield. The effect of calving age on milk yield among individuals was reduced because of the focus on time to start of AI based on the size of the individual heifer. The HM at level $_M$ was simulated by a heat-detection rate of 0.50, and at the resulting average calving age, a marginal effect was assumed of 1 extra day of calving age of 0.0066 kg of ECM/d in first lactation. At levels $_H$ and $_L$, heat-detection rates of 0.60 and 0.40, respectively, were assumed, and at the resulting average calving ages, marginal effects of 1 extra day of calving age of 0.0033 and 0.0099 kg of ECM/d, respectively, were assumed in first lactation (C. Enevoldsen, unpublished data; J. Ettema, Department of Animal Health, Welfare and Nutrition, Faculty of Agricultural Sciences, University of Aarhus, Research Center Foulum, Tjele, Denmark, 2006). In all scenarios, the marginal effect of an extra day of calving age was simulated to decline linearly to zero at a calving age of 30 mo.

LC-V. The between-cow variation within herds is highly variable between herds. One reason could be that social stress, meager housing design, or diseases such as lameness limit the feed intake of some cows. Consequently, LC-V could be a potential KPI. From an unpublished analysis (M. A. Krogh, unpublished data), the 10th percentile herd had a variance 50% the size of the variance in the 50th percentile herd, so the $SdW1_j$ was reduced accordingly for level $_H$:

Parity 1: $SdW1_j = \sqrt{(0.0147^2 \times 0.50)} = 0.0104$,

Parity 2: $SdW1_j = 0.0153$, and

Parity 3+: $SdW1_j = 0.0177$.

MCow. The estimates used were from an unpublished analysis by P. T. Thomsen (Department of Animal Health, Welfare and Nutrition, Faculty of Agricultural Sciences, University of Aarhus, Research Center Foulum, Tjele, Denmark, 2006) regarding mortality in Danish dairy cows for weekly estimates of incidence rate of cow death: $_L = 0.001233$, $_M = 0.000678$, and $_H = 0.00024$.

Mcalf. From Danish standard protocols, the probability of a calf surviving birth and the first 180 d postpartum for first parity was $_L = 0.77$, $_M = 0.84$, and $_H = 0.90$, and the probability of a calf surviving birth and the first 180 d postpartum for second parity and 3+ parity was $_L = 0.81$, $_M = 0.87$, and $_H = 0.93$.

BCS. The model for BW and BCS of a cow in SimHerd was modified for this study to empirically represent estimated parameters more directly in the simulations. This implied that the feed intake was a consequence of energy needed to match the production level. A Gompertz curve was used to describe the BW of the animal corrected to a BCS of 3.0 and excluded any weight of a fetus:

$$BW_{BCS=3.0} = MatureBW \times exp[-m \times exp(-n \times Age_{ik})],$$

where Age_{ik} is the age in days of animal$_i$ in the kth time step, and m and n are model parameters describing the shape of the curve. Based on the results from Nielsen et al. (2003), estimates were MatureBW = 680, m = 2.5483, and n = 0.00314.

Figure 2 shows the applied Gompertz curve describing the BCS-corrected BW. The BCS change of the cow was based on the model of Friggens et al. (2004). First, the cows were assumed to be driven to a certain BCS at the nadir after the first part of the lactation (phase 1). Second, the BCS would not change until pregnancy (phase 2). Finally, during pregnancy the cows were assumed to be driven to a certain BCS at calving (phase 3). Body condition was specified at the 2 different time points in the lactational cycle: at calving (a fixed BCS of 3.50) and at nadir, which was 70 d after calving. There was a relationship between BCS and fertility in the simulation model for the individual cow. If BCS dropped below 2.75, there was a reduced likelihood of the onset of the first ovulation. This relationship was based on the model described by Friggens and Chagunda (2005).

In the context of the simulation experiment, defined variables were high mobilization ($_H$), an expected mobilization ($_M$), and a low mobilization ($_L$). Input variables at the parity level are shown in Table 3.

SCC. The SCC input parameters were fitted from Østergaard et al. (2005) and represented estimates of SCC at the herd level, given different levels of management (Table 4). Because of the selection criteria, clinical mastitis was excluded. The reason for this decision was the large variation in practical management (i.e., medical treatment(s) or selective drying off of infected quarters) of clinical mastitis at the herd and cow levels (Vaarst et al., 2002), which makes comparison of the occurrence of clinical mastitis in different herds complicated. The impact of SCC on milk yield was mediated through LC. Consequently, SCC affected only gross margin through the milk price. The corrections for bulk tank SCC per milliliter on the milk price per kilogram of ECM (€0.291; Table 5) were +1.8, +0.9, −3.6, and −9.0% for bulk tank SCC of ≤200,000; 200,000 to
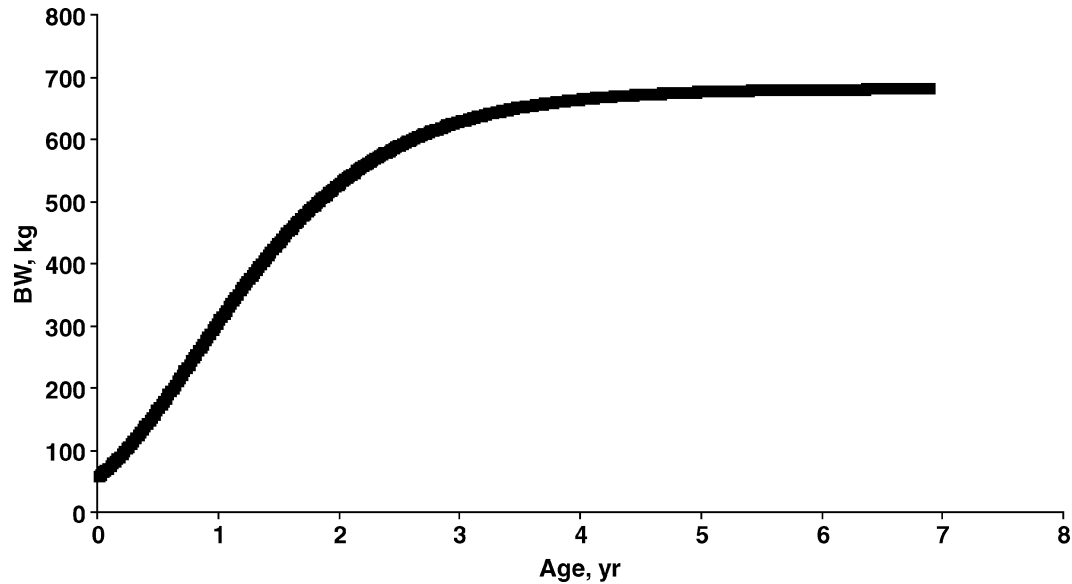
Figure 2. Visualization of the Gompertz curve describing the dynamics of BCS-corrected BW within the SimHerd model (see Table 3).

Table 3. Parity-specific values of BCS used to define management levels in the simulation experiment

BCS at parity level	Low	Middle	High
Parity 1, at calving	3.50	3.50	3.50
Parity 1, at nadir	3.25	3.00	2.75
Parity 2, at calving	3.50	3.50	3.50
Parity 2, at nadir	3.00	2.75	2.50
Parity 3+, at calving	3.50	3.50	3.50
Parity 3+, at nadir	3.00	2.75	2.50

Table 4. Parity-specific values of SCC used to define management levels in the simulation experiment

Level of SCC management	Low	Middle	High
Parity 1, × 1,000/mL	191	152	112
Parity 2+, × 1,000/mL	451	360	268

Table 5. Examples of prices or costs (€) applied in the simulation experiment

Variable	€
ECM per kg	0.291
AI per insemination	13.33
Heifer <1 yr per head (buy or sell)	453
Pregnant heifer per head (buy or sell)	1,067
Slaughter price per kg	0.893
Cost of dead cow, obliteration	49
Bull calf, price for sale at 14 d	173
Scandinavian Feeding Unit (SFU)	0.17

300,000; 400,000 to 500,000; and >500,000, respectively (Østergaard et al., 2005).

Simulation Experiment and Statistical Analysis

The simulation experiment was conducted as follows:

1. The choice of the potential KPI (e.g., lactation curves with different levels of peak yield and persistency) was assumed to mimic effects of a systematic change in input factors in the herd management programs.
2. Simulated management changes (expressed as changes in KPI levels) were translated into changes in gross margins through the SimHerd model.
3. The set of assumptions made SimHerd aim at keeping the herd size as close as possible to the maximum of 250 cows.
4. The simulation took place in a situation without any quota constraints; that is, the production was constrained by a maximum number of cows in the herd.
5. Cows and heifers were fed TMR with a standard feed price.
6. Prices were set (in €) according to typical prices in Denmark in 2006 (Table 5).
7. The output from the 10th simulation year was used for analysis, because initial exploration of the simulated data showed that in some scenarios, it took up to 9 yr to obtain steady state (i.e., the "burn-in"

Table 6. Key characteristics of the default herd (no changes in management) in the 10th simulation year after 200 independent replications

Selected variable	Mean	SD	Minimum	Maximum
Parity 1, n	93	6.6	73	109
Parity 2, n	59	6.1	42	79
Parity 3+, n	95	7.2	75	113
Heifers >1 yr, n	155	10.0	124	178
Heifers <1 yr, n	116	8.2	98	139
Cows per year,[1] n	248	0.2	247	248
Bulk tank SCC, × 1,000/mL	261	6	240	280
ECM per cow/yr, kg	9,735	77.7	9,530	9,972
Culling rate[2]	40	3.0	31	50
Calves born per cow/yr	1.08	0.03	0.96	1.15
AI per cow/yr	2.26	0.10	1.99	2.59
Age at first calving, mo	28.40	0.2	27.8	29.0
Total revenue, €	791,820	8,055	764,734	814,562
Total expenses, €	400,637	3,813	391,302	411,612
GM[3] per cow, €	1,578	25	1,508	1,642

[1]Cows per year = total number of cow days in a year/365.

[2]Calculated according to the Danish definition: (number of cows going into the herd + number of cows leaving the herd)/2/number of cows per year.

[3]GM = gross margin from the herd as a whole divided by the number of cows per year.

period reflecting the initial bias; Abate and Whitt, 1987; Chen and Kelton, 2003).

The simulation took place in a no-quota situation. The reason was that the quota system had become more liberal in Denmark and was expected to be lifted within the European Community in the near future.

Key characteristics of the default herd in the 10th simulation year after 200 independent replications are described in Table 6. This herd was defined by having all the KPI placed at level $_M$.

Statistical Analysis. The simulated results from the 10th simulation year were analyzed by means of an ANOVA using SAS PROC MIXED (Littell et al., 2006). The full model described above was reduced by backward elimination of the KPI and their interactions until the *P*-values of all factors were highly significant ($P < 0.0001$). The residuals from the resulting model were not normally distributed (Kolmogorov-Smirnov <0.01). A qualitative analysis was performed on each of the 72 scenarios, and 6 scenarios were identified in which the residuals were not normally distributed (Shapiro-Wilk <0.05). These 6 scenarios were excluded from the data set and were referred to a qualitative analysis of relevant herd-level variables (calving interval, cows per year, calf mortality, total milk yield, culling decision, heifers per year, age at first calving, total gross margin, gross margin per cow, total income, total expenses). The qualitative analysis revealed that the input variables for RE_L may have been too low compared with RE_H, making the data skewed within 4 scenarios of the initial model. Other systematic relationships between the "not acceptable" scenarios were not

identified, indicating that the stochastic elements of the SimHerd model probably were responsible for the remaining 2 scenarios. Subsequently, the 6 "not acceptable" scenarios were removed from the data set.

The potential effect of the selected variables was estimated on the gross margin with the ANOVA model by using Satterthwaite's approximation. The level of financial significance was set at €1.33/cow per year. The potential KPI and their interactions had to comply with both statistical and financial significance to be retained in the final model. Eight scenarios were removed because of financial nonsignificance.

Range of Effects. By using linear contrasts, the 2- and 3-factor interactions were dissolved to study the differences between KPI and their relation to the gross margin when changing the KPI levels. To compare the KPI, the differences from $_L$ to $_M$ and from $_M$ to $_H$ were used, making it possible to compare the 2-level KPI with the 3-level KPI. That is, the unit of KPI change was largely 1 quartile within the interquartile range. The design provided a direct link to data from benchmarking facilities in herd management programs, yet the study design did not allow us to draw conclusions on the contrasts between $_M$ and $_L$ for the 2-level KPI.

RESULTS

The Model

The initial ANOVA model was reduced (statistical significance: $P < 0.0001$) to the following final model (the metamodel):

Table 7. Estimates of statistically and financially significant effects on gross margin of a series of key indicators of technical performance (KPI)[1]

Main effect of KPI	Estimate, €	2-factor interaction	Estimate, €	3-factor interaction	Estimate, €
Intercept	1,578	$LC_H \times RE_L$	1.7	$LC_H \times RE_H \times HM_H$	−4.6
Shape of lactation curve (LC_H)	206.7	$LC_L \times RE_L$	7.4	$LC_H \times RE_H \times HM_L$	−3.5
Shape of lactation curve (LC_L)	−222.9	$LC_H \times HM_H$	−1.5	$LC_H \times RE_L \times HM_L$	−16.6
Reproduction efficiency (RE_H)	9.1	$LC_L \times HM_L$	3.7	$LC_L \times RE_H \times HM_H$	−3.9
Reproduction efficiency (RE_L)	−67.0	$RE_L \times HM_H$	25.4	$LC_L \times RE_H \times HM_L$	3.1
Heifer management (HM_H)	16.9	$RE_L \times HM_L$	−21.0	$LC_L \times RE_L \times HM_L$	21.0
Cow mortality ($MCow_H$)	19.8	$RE_H \times MCow_H$	3.5		
Calf mortality ($Mcalf_H$)	16.7	$RE_H \times Mcalf_H$	1.6		
SCC (SCC_H)	15.4	$RE_H \times BCS_H$	7.9		
		$RE_L \times BCS_H$	9.6		

[1]The intercept represents the gross margin in the default herd (see Table 6). $_H$ = good farm management; $_L$ = pitiable farm management. Mean SE, € = 28.6; assuming variance homogeneity, a confidence interval of the predictions can be estimated from $2 \times$ mean SE ~ €57.

$$\text{Gross margin} = LC_{(H, M, L)} + RE_{(H, M, L)} + HM_{(H, M, L)}$$
$$+ MCow_{(H, M)} + Mcalf_{(H, M)} + BCS_{(H, M)} + SCC_{(H, M)}$$
$$+ LC \times RE + LC \times HM + RE \times HM + LC \times RE \times HM$$
$$+ RE \times MCow + RE \times Mcalf + RE \times BCS.$$

KPI

Table 7 presents the KPI and interactions that complied with both statistical and financial significance (LC, RE, HM, MCow, Mcalf, BCS, SCC). The final model explained 96% of the variation in the simulated data. The within-scenario variation was negligible (0.5% of total variance; $P < 0.0001$) for practical purposes.

Example. An example use of KPI was derived from Table 7. A farmer owns a herd identical to the default herd (herd characteristics in Table 6) and asks about the expected financial performance if all the KPI change from $_M$ to $_H$. The answer is: We simply add or subtract the values (€) of the KPI and the interactions in question from each other: $206.7 + 9.1 + 16.9 + 19.8 + 16.7 + 15.4 − 1.5 + 3.5 + 1.6 + 7.9 − 4.6 \sim €291$. The 95% confidence interval for financial performance, given the specified changes in KPI, was €235 to €349 (based on the root mean standard error; Table 7).

In our setup, the default herd consisted of 248 cows, with a mean gross margin per cow per year equal to €1,578 (Table 6). Thus, the gross margin for the default herd equaled €391,344. The best-case scenario equaled an improvement of the gross margin by almost 20%. This value took into account important interactions among the KPI and prevented double counting because of the simulation design. Nonetheless, costs of increasing the quality of management, such as additional labor, management support, and quality of feed, necessary to obtain the changes were not included.

Interactions and Range of Effects

Tables 7 and 8 show that the relation between LC and gross margin was modified by RE and HM. The lowest difference between LC_H and LC_L was €332, which occurred when both RE and HM were $_L$. The remaining differences between LC_H (€206.7) and LC_L (€−222.9) were very similar (~€430). When RE and HM were $_L$, the difference was smaller because of the impact of the 3-factor interaction when all 3 KPI (LC, RE, and HM) were $_L$ (€21.0 in Table 7).

The relation between RE and gross margin was modified by LC and HM. The lowest difference between RE_H and RE_L was €52, which occurred when both LC and HM were $_L$. The remaining differences between RE_H and RE_L were €71 to €89. The contrasts revealed that the impact of RE on the gross margin was skewed, making losses associated with moving from $_M$ to $_L$ (€67) much larger than the gain associated with moving from $_M$ to $_H$ (€9), regardless of the levels of LC and HM. If RE was $_L$ and LC was kept constant, it was only possible to increase the gross margin slightly (€16 to €34) by improving HM.

Reproduction efficiency was involved in 2-factor interactions with MCow, Mcalf, and BCS. The smallest effect including MCow was €20, and for Mcalf it was €17. Both occurred when RE was different from $_H$. In contrast, the smallest effect of BCS (€15) was found at RE_M.

Somatic cell count was not significant in any interactions. The effect of SCC was €15 (the difference between SCC_H and SCC_M).

Table 8 ranks the KPI by the largest effects on gross margin, measured as quartiles within the interquartile ranges, and provides estimates of the relative financial performance of the KPI and the interactions.

Table 8. The numerical and relative importance of key indicators of technical performance (KPI) with respect to their long-term impact on the financial performance in dairy herds, measured as gross margin per cow

KPI	Interquartile range,[1] €	Percentage of total	Short description of the most important findings between levels of KPI
Shape of lactation curve (LC)	227	53	The range is independent of the levels of RE and HM.
Reproduction efficiency (RE)	89	21	Most affected by LC. At LC_L the effect decreases to €52.
Heifer management (HM)	34	8	The effect is €34 at levels LC_H and RE_L. Otherwise HM is €17.
BCS dynamics (BCS)	25	6	At BCS_M and RE_L, the gross margin is €10 if BCS_M improves to BCS_H.
Cow mortality (MCow)	23	5	The effect is €23 regardless of RE level.
Calf mortality (Mcalf)	18	4	The effect is €18 regardless of RE level.
SCC	15	3	SCC is not affected by interactions; thus, the effect is equal to the main effect.

[1]To compare the potential financial performance of the KPI, we used the differences from levels $_L$ to $_M$ and from $_H$ to $_M$, making it possible to compare the 2-level KPI with the 3-level KPI. These differences correspond largely to 1 quartile within the interquartile ranges. $_H$ = good farm management; $_L$ = pitiable farm management; $_M$ = middle (average of $_H$ and $_L$).

Examples. To illustrate the interpretation of the results, 2 small examples are presented and show a particularly interesting finding:

1. Figure 3 illustrates the 3-factor interaction between levels of RE, HM, and LC_H. The maximum value (€17) from moving HM 1 quartile was found at RE_L and LC_H when the movement was from $_L$ to $_M$. The maximum value (€89) of RE was found at HM_L and LC_H when the movement was from $_L$ to $_M$. Notice that at RE_M, there was no additional gross margin associated with moving HM from $_L$ to $_M$.

2. The differences in expected number of median days open between RE_L and RE_M and between RE_M and RE_H were calculated from the parameters in Table 1. The corresponding differences in the gross margin were divided by these numbers of days open. With this approach, it was possible to estimate the cost per day open within different levels of RE. At $MCow_H$, the average cost of moving RE_L to RE_M and RE_M to RE_H was €5 and €1 per day open, respectively. At HM_L, the average cost of moving RE_L to RE_M was €7.

The change in gross margin was calculated between scenarios that were only different regarding levels of MCow and Mcalf, respectively (data not shown). This provided the financial loss associated with 1 dead animal within different levels of LC. At $Mcalf_H$ and LC_H, the average cost of moving $MCow_M$ to $MCow_H$ was €1,013 for a dead cow. At LC_L, the cost was €863. With the same approach, the cost of a dead calf was estimated at €291 and €264, respectively, in herds with LC_H and LC_L.

Table 8 describes the most important findings based on quartiles within the interquartile ranges. All differences were significant ($P < 0.0001$) and were numerically larger than the level of financial significance (€1.33/cow per year).

DISCUSSION

Validation of the Metamodel with Respect to the Simulation Model (SimHerd)

The metamodel fit very well with the aggregated data from the simulation experiment conducted with Sim-Herd ($R^2 = 0.96$). The gross margin output from Sim-
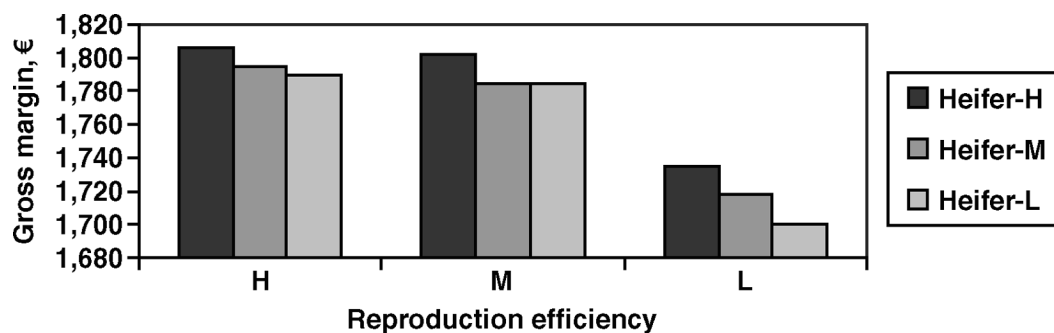
Figure 3. Visualization of the financial performance related to reproduction efficiency and heifer management, given the shape of lactation curves corresponding to a high management level. H is applicable to "good farm management"; L is applicable to "pitiable farm management"; and M is the average of H and L.

Herd, and consequently from the metamodel, responded to changes in KPI levels in the direction that agreed with prior qualitative knowledge about the simulated problem entity. In most cases, plausible explanations were provided for the rather complex interactions between KPI. These interactions provided more insight into the complex behavior of the herd as a system. The face validation of the pathways from assumed management adjustments, to KPI, to simulation input, to simulation output, and finally to metamodel output suggests that the metamodel provides a valid tool for herd advisors (Sørensen, 1990).

A word of caution when using the metamodel: the low reproduction efficiency within the 4 "not acceptable" scenarios made it impossible for SimHerd to maintain a steady number of cows in the simulated herd without frequent purchase of pregnant heifers. The same situation with unfortunate reproduction management could easily occur in real life, but SimHerd was probably too simple to simulate such extreme examples. SimHerd simply assumed that the farmer would wait until the herd size had dropped to a certain number of cows and then pregnant heifers were purchased, 1 heifer at a time, to ensure that the herd size did not drop further. Consequently, at present there are extreme scenarios that cannot be modeled in a satisfactory way with SimHerd. This may be due to the simple nature of the feedback mechanism for purchase in the SimHerd model. This is an important finding that has added further information to the validity of the SimHerd model.

General Discussion

Our study basically was a condensation of a series of herd simulations with the SimHerd model that provided a much more user-friendly, and nevertheless valid, tool for predicting the financial effect of the most relevant management adjustments in herd management. The chosen metamodel circumvented the problems related to obtaining the large number of input variables needed for complex simulation models for decision supports (Enevoldsen et al., 1995).

The financial performance associated with changes in herd management did not include labor and management costs or costs associated with needs for improved feed quality, which may be important costs in a real herd decision problem. In that case, these costs must be estimated and subtracted from the gross margin estimated with the metamodel. In the interpretation of the results, it should be mentioned that the difficulty or ease of achieving a certain management change is herd specific. For instance, it is likely that for some farmers, it is easier or less costly to obtain the gross

margin indicated by our study than for others. It would be easier to move most of the KPI from $_L$ to $_M$ than from $_M$ to $_H$.

In situations in which the milk quota is the major production constraint in the herd, rather than the number of cows (as we have assumed), the gross margin per kilogram of milk produced is a relevant measure of financial performance because of the extra costs of producing more or less than the milk quota. The general mechanism of a milk quota was that strategies that affected the milk yield were generally reduced; that is, the loss per dead animal dropped to about half compared with a no-quota situation (Sørensen and Enevoldsen, 1991). The reason for this is that by implementing preventive measures, a herd under an unadjustable quota can prepare for this situation (the dead animal) by increasing production; however, if the situation does not occur, the herd will need to be fit into the allowable production by reducing the cow numbers. The possibility of buying and selling quotas offers the farmer another option, which makes gross margin per kilogram of milk produced an incomplete financial measure. The European milk quota system is accelerating, and in Denmark it is now possible to buy and sell milk quotas 4 times a year. This provides the individual farmer with great flexibly to adjust to the quota situation. It would be very difficult, perhaps impossible, to implement this flexibility in the simulation. Because simulation under a quota restriction will not reflect reality and because of the long time intervals of some of the simulated management changes, we deliberately chose to simulate without adjusting for the financial effect of a milk quota.

The ranking of the KPI was based on the gross margin obtained after 10 yr of simulation, where the simulation experiment reached steady state. On the other hand, the financial value of a given management change obviously depends on the time span until full manifestation of the effects. That is, the gross margin obtained in all the simulation years ideally should be discounted and transformed into a net present value.

The planning horizon differs among farmers and within farmer, depending on the characteristics of the management change. Therefore, both short-term and long-term predictions will be relevant for the decision-making process, but the short-term behavior of the SimHerd model has not been studied in sufficient detail to allow this type of analysis. Consequently, the short-term consequences on gross margin until the time of steady state need to be explored further.

Implications from the Results

The results of this study are intended to support the prediction of the financial performance associated with

practically feasible changes in specified KPI. The constructed KPI levels cover the interquartile ranges of KPI obtained in Danish dairy herds reasonably well. Consequently, benchmarking facilities in efficient herd management software probably could produce the information needed to use the general results described in Table 8. The detailed descriptions of the modeling assumptions allow potential users to judge whether the metamodel is valid for contexts of interest to them.

The financial performance associated with RE is mediated through 2-factor interactions between RE and each of MCow, Mcalf, and BCS. This was an important finding, because this made RE at the herd level even more important than what was calculated if a simpler model was used (partial budget or similar).

The interaction between BCS and RE was explained by the effect of the period of negative energy balance on BCS postpartum and the likelihood of onset of estrus (Friggens and Chagunda, 2005), because a low BCS indicates postponed onset of estrous cycling. Body condition score thereby affects RE. If BCS drops below 2.75, then SimHerd links the negative energy balance with an increase in time to onset of estrous cycling of 1 wk. Increasing BCS from level $_M$ to $_H$ reduced the impact of negative energy balance postpartum on RE in the SimHerd model. Then again, in our scenarios only a few cows experienced a detrimental effect on RE. The interaction between BCS and RE was expected to be more pronounced if the BCS levels became lower than what we had simulated.

Variation between cows in lactation curve persistency was not significant in the metamodel. The changes related to LC-V may be too subtle to be identified by using only 1 quartile, or the modeling may have been too crude; that is, we assumed that the reduction in variance affected all cows.

It may seem rather surprising that SCC did not interact with culling or production level as would be expected in real life. On the other hand, the study design prevented us from drawing any conclusions regarding such interactions; that is, we included data that were possible to obtain from a normal herd health program, with focus on potential production improvements. They may be caused partly by (absence of) disease, yet the decreased milk production caused by SSC in real life may not be fully reflected when estimating the impact of SCC on financial performance. In other words, the model underestimated the financial impact caused by SCC.

The metamodel showed that more than 50% of the changes in additional gross margin could be obtained by means of improving the LC by 1 quartile. Next, RE represented 20%. The other KPI included represented approximately the same value (€15 to 20).

CONCLUSIONS

The result from a complex long-term simulation experiment was used to estimate the financial performance of specified key technical performance indicators, measured as gross margin per cow per year. The results from the simulation experiment were condensed into a metamodel to improve user-friendliness compared with the rather complex SimHerd model. The metamodel used data extracted from routinely collected management data to forecast the financial performance related to specified management changes in specific dairy herds characterized by very different sets of key technical performance indicators.

This study indicated that improving the shape of the herd-level lactation curve by 1 quartile was associated with a gross margin increase of up to €227/cow year in a no-quota situation. This was 2.6 times more than improved RE, which increased the gross margin 2.6 to 5.9 times more than improved HM, BCS, mortality, and SCC.

The results showed numerous significant interactions between the different combinations of technical performance indicators. This implies that financial performance related to certain management strategies will depend significantly on the management level in other areas of herd management. This is perhaps the most important finding of this study.

REFERENCES

Abate, J., and W. Whitt. 1987. Transient behaviour of regular Brownian motion. Adv. Appl. Probab. 19:560–631.

Chen, E., and W. Kelton. 2003. Determining simulation run length with the runs test. Simul. Model. Pract. Theory 11:237–250.

Danish Agricultural Advisory Service. 2005–2006. Årsstatitik Avl. Rapport no. 115. Dan. Agric. Advisory Serv., Landscentret, Aarhus, Denmark.

Dijkhuizen, A. A., R. B. M. Huirne, and A. W. Jalvingh. 1995. Economic analysis of animal diseases and their control. Prev. Vet. Med. 25:135–149.

Dijkhuizen, A. A., J. Sol, and J. Stelwagen. 1984. A three year herd health and management program on thirty Dutch dairy farms. III: Economic evaluation of fertility control. Vet. Q. 6:158–169.

Enevoldsen, C., J. Hindhede, and T. Kristensen. 1996. Dairy herd management types assessed from indicators of health, reproduction, replacement, and production. J. Dairy Sci. 79:1221–1236.

Enevoldsen, C., and M. A. Krogh. 2006. A random coefficient regression analysis of lactation curves for dairy herd management. Page 236 in Proc. Int. Soc. Vet. Epidemiol. Econ., Queensland, Australia. http://www.sciquest.org.nz/default.asp?pageid=69&pub=10&vol=11 Accessed Dec. 7, 2007.

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. 78:947–961.

Ferguson, J. D., D. K. Beede, R. D. Shaver, C. E. Polan, J. T. Huber, and P. T. Chandler. 2000. A method to analyze production responses in dairy herds. J. Dairy Sci. 83:1530–1542.

Freudendal, A. J., and F. Strudsholm. 2003. Hjælper aktivitetsmålerne ved brunstkontrol? KvægInfo no. 1254. Dan. Agric. Advisory Serv., Landscentret, Aarhus, Denmark.

Friggens, N. C., and M. G. G. Chagunda. 2005. Prediction of the reproductive status of cattle on the basis of milk progesterone measures: Model description. Theriogenology 64:155–190.

Friggens, N. C., K. L. Ingvartsen, and G. C. Emmans. 2004. Prediction of body lipid change in pregnancy and lactation. J. Dairy Sci. 87:988–1000.

Jakobsen, J. H., P. Madsen, and J. Pedersen. 2002. Multivariate covariance functions for test day production in Danish dairy breeds. Pages 95–102 in Proc. Interbull Mtg, Interlaken, Switzerland. http://agtr.ilri.cgiar.org/Library/docs/Interbull/bulletin29_files/docs/Jakobsen.pdf Accessed Dec. 6, 2007.

Kaplan, R. S., and D. P. Norton. 1998. The Balanced Scorecard: Translating Strategy into Action. Harvard Business School Press, Watertown, MA.

Kleijnen, J. P. C., and R. G. Sargent. 2000. A methodology for fitting and validating metamodels in simulation. Eur. J. Oper. Res. 120:14–29.

KoNet-Praksis. 2006. Epidemiology in practice. Proc. World Buiatrics Congr. XXIV, Nice, France. http://www.nice-acropolis.com/wbc2006/EN/ProgScient.php Accessed Dec. 6, 2007.

Littell, R., G. Milliken, W. Stroup, R. Wolfinger, and O. Schabenberger. 2006. SAS for Mixed Models. 2nd ed. SAS Inst. Inc., Cary, NC.

Mourits, M. C. M., A. A. Dijkhuizen, R. B. M. Huirne, and D. T. Galligan. 1997. Technical and economic models to support heifer management decisions: Basic concepts. J. Dairy Sci. 80:1406–1415.

Nielsen, H. M., N. C. Friggens, P. Løvendahl, J. Jensen, and K. L. Ingvartsen. 2003. Influence of breed, parity, and stage of lactation on lactational performance and relationship between body fatness and live weight. Livest. Prod. Sci. 79:119–133.

Østergaard, S., M. G. G. Chagunda, N. C. Friggens, T. W. Bennedsgaard, and I. C. Klaas. 2005. A stochastic model simulating pathogen-specific mastitis control in a dairy herd. J. Dairy Sci. 88:4243–4257.

Østergaard, S., J. T. Sørensen, and A. R. Kristensen. 2000. A stochastic model simulating the feeding-health-production complex in a dairy herd. J. Dairy Sci. 83:721–733.

Seegers, H., C. Fourichon, and F. Beaudeau. 2003. Production effects related to mastitis and mastitis economics in dairy cattle herds. Vet. Res. 34:475–491.

Shalloo, L., P. Dillon, M. Rath, and M. Wallace. 2004. Description and validation of the Moorepark Dairy System Model. J. Dairy Sci. 87:1945–1959.

Sørensen, J. T. 1990. Validation of livestock herd simulation models: A review. Livest. Prod. Sci. 26:79–90.

Sørensen, J. T., and C. Enevoldsen. 1991. Kvægsundhed, økonomi og mælkekvote. Dan. Vet. J. 74: 298–302. (In Danish)

Sørensen, J. T., E. S. Kristensen, and I. Thysen. 1992. A stochastic model simulating the dairy herd on a PC. Agric. Syst. 39:177–200.

Tauer, L. W., and A. K. Mishra. 2006. Dairy farm cost efficiency. J. Dairy Sci. 89:4937–4943.

Thomsen, P. T., S. Østergaard, J. T. Sørensen, and H. Houe. 2007. Loser cows in Danish dairy herds: Definition, prevalence and consequences. Prev. Vet. Med. 79:136–154.

Vaarst, M., B. Paarup-Laursen, H. House, C. Fossing, and H. J. Andersen. 2002. Farmers' choice of medical treatment of mastitis in Danish dairy herds based on qualitative research interviews. J. Dairy Sci. 85:992–1001.

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