Animal-wise prediction of milk, protein and fat yield using solutions from the Nordic test-day model

Pitkänen T.[1], Mäntysaari E.A.[1], Rinne M.[2], Lidauer M.[1], [1]MTT Agrifood Research Finland, Biotechnology and Food Research, [2]Animal Production Research, 31600 Jokioinen

Introduction

The purpose of this work was to make animal-wise predictions of milk, protein and fat yield for the Finnish Dairy cows using solutions from the Nordic test day yield evaluation (Lidauer et al, 2006). The prediction equations will be incorporated by the Finnish Agricultural Data Processing Centre (MLOY) into an interactive web-application where herd owners can obtain yield predictions for a cow or for a group of cows.

The Nordic test-day model describes all available test-day observations for milk, protein and fat yield from Denmark, Finland and Sweden from 1988 onwards. Evaluations are performed 4 times a year to provide estimates for all animal and environmental effects that affect on yield. Three breed-specific random regression test-day models (RRM) are used for evaluation of the main breeds. In this study the prediction equations have been built on the Red cattle model, which includes data from all Finnish cows in milk recording.

Predicting yields

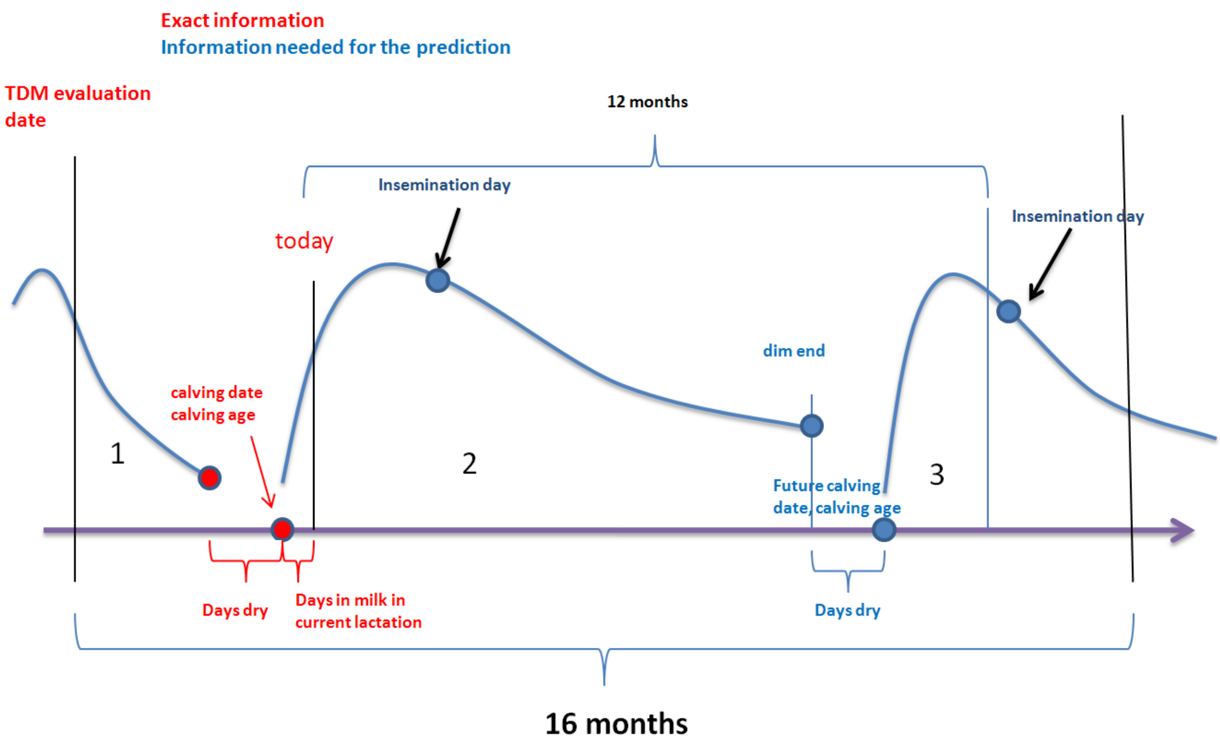

Figure 1. A sketch of three lactation curves within a 16 month interval with animal information needed to predict yields for 12 months within that interval. The information about cow's past events comes from the MLOY databases and future events are predicted. As the test-day model evaluations are preformed four times a year, prediction equations are made for 16 months to provide at least one year of prediction.

Prediction of a cow's future milk, protein and fat yield is based on the cow's genetic and non-genetic lactation curves, environmental-specific lactation curves, the cow's age, parity and breed, as well as on the prediction of the herd management effect on the production level. Finally, stage of lactation, calving date and insemination date are input parameters for the prediction (figure 1).

On each evaluation run, a large amount of solutions for effects in the test-day model are produced and only small proportion of those are needed for prediction of future yields. Therefore, test-day model solutions are first "filtered" where relevant solutions for the predictions are selected and processed so that the prediction is relatively straightforward to compute. For example, genetic and non-genetic lactation curves are needed only for cows alive but the RRM produces those to all animals in the pedigree. Genetic lactation curves are also needed for heifers that start producing milk within the prediction interval. In that case, heifers genetic lactation curves are their parent averages. The RRM will provide genetic and non-genetic lactation curves for heifers after they have made their first test-day observation.

Predicting herd management effects

The herd management effect needs to be predicted for future months. In the test-day model the herd management effect is modeled with fixed test-month, fixed herd year, and random herd*test-month effects. Observed herd management effects for each herd are obtained by adding up solutions for those effects (figure 2).

A random coefficient model is fitted separately for each herd and trait to model observed herd effects and it is used to predict the future management effects. The model used is based on the model presented by Koivula (2007). In the model a year is defined as production year which starts in June and therefore the first month is June and last month is May. The overall trend in herd management level is modeled by a regression line, which slope is allowed to be different for the last four production years. Monthly trends within a production year are considered to be random effects and are modeled by a linear trend and three sine functions.

The model equation for production year *y* and production month *m* is

$$Y_{ym} = \mu + \beta_{1b} * m_{88} + \beta_{1a} * m_{88} * I_{4year} + \left(\beta_2 + u_{1y}\right) * m + (\beta_3 + u_{2y}) * \sin\left(\frac{2}{12}\pi(m-6)\right)$$

$$+ (\beta_4 + u_{3y}) * \sin\left(\frac{2}{12}\pi(m-8)\right) + \left(\beta_5 + u_{y4}\right) * \sin\left(-\frac{2}{6}\pi(m+6)\right) + \epsilon_{ym}$$

where m_{88} is the months from January 1988, I_4 is the indicator indicating if ym is closer than 48 months from RRM evaluation date, β_{1b} is the for linear term for first years, β_{1a} is the linear term for the most recent four years, β_2 is the slope for within year linear term, β_3, β_4 and β_5 are the fixed coefficients for sine terms, u_{2y} , u_{3y} , u_{4y} and u_{5y} are random the coefficients for sine terms and ϵ_{ym} is error term. Random coefficients are assumed to follow multivariate normal distribution with mean 0 and unstructured covariance structure. The error term is assumed to be mutually independent of random coefficients and assumed to follow normal distribution with mean 0 and variance σ^2. The production year starts in June which is the first production month (m=1) and ends in may (m=12).

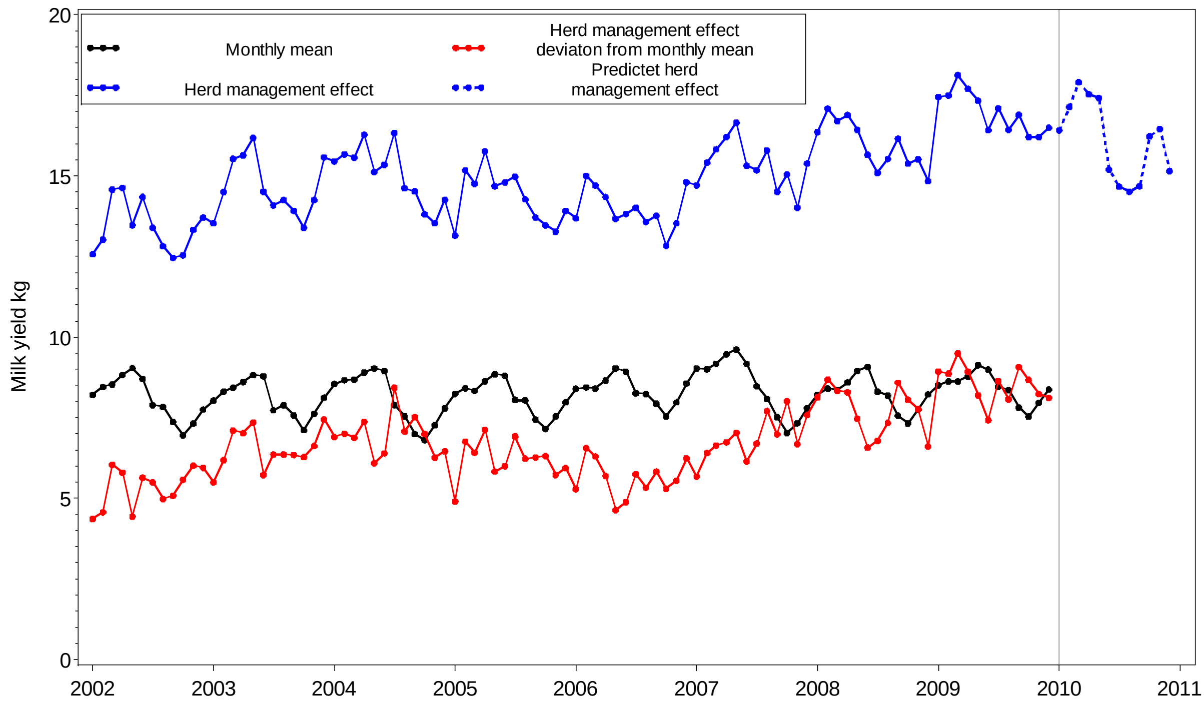

Figure 2. Prediction of herd management effect. The herd management effect consists of a monthly mean and a herd management effect deviation from the monthly mean. The future effects (dashed) are predicted by using linear mixed effects model.

Average animals

The RRM provides solutions for prediction only for animals in the milk recording system. In some situations predictions are needed for cows that are not in the evaluation or their identity is unknown. This can occur when herd owner plans to buy new animals and would like to have some kind of prediction for those. Therefore, we calculated mean animals for different birth year-groups that can be used for the prediction.

Accuracy of prediction

The cows' test-day observations in 2010 have been predicted by using solutions of effects from test-day model evaluation based on data until December 2009. The summary of test-day data in year 2010 is in the table 1. The stage of lactation and other cow-wise information needed for predictions were obtained from actual test-day data for year 2010. Hence, there was no need to predict calving and insemination dates.

The ability to predict test-day observations were measured by using correlation between observed and predicted test-day yields and the coefficient of determination R^2. The herd-level prediction ability was measured by comparing sum of test-day observations and sum of predictions within herd in different months. The correlation and R^2 is calculated monthly to see possible trends in accuracy of predictions.

Month	Number of animals	number of test-day observations	Number of herds	Animals per herd
January	180855	181447	7844	23.1
February	179596	180191	7730	23.2
March	180455	182396	7745	23.3
April	180368	181408	7689	23.5
May	181222	182620	7663	23.6
June	179316	180131	7579	23.7
July	179520	180659	7572	23.7
August	180867	182295	7586	23.8
September	179097	180266	7535	23.8
October	176595	178401	7492	23.6
November	175019	176136	7447	23.5
December	177090	178935	7421	23.9
TOTAL	**289484**	**2164885**	**7950**	**36.4**

Table 1. Number of different animals, test-day observations, herds having test day records in year 2010.

Results

For milk yield the correlation (figure 3) depends on parity and decreases from January to October. For mature cows the correlation is higher because of information from earlier lactations. In fat and protein yield the correlation followed similar pattern but were lower than for milk (table 2).

	Milk		**Protein**		**Fat**	
	Correlation	**R^2**	**Correlation**	**R^2**	**Correlation**	**R^2**
Animal level						
Lactation 1	**0.60 - 0.84**	**0.31 - 0.71**	**0.45 - 0.80**	**0.03 - 0.63**	**0.40 -0.64**	**0.06 - 0.40**
Lactation 2	**0.73 - 0.89**	**0.49 - 0.79**	**0.63 - 0.85**	**0.31 - 0.70**	**0.53 -0.74**	**0.20 - 0.55**
Later lactations	**0.74 – 0.89**	**0.50 -0.80**	**0.65 - 0.85**	**0.26 - 0.72**	**0.58 - 0.76**	**0.27 - 0.58**
Herd level						
Total yield across all animals within herds	**0.99**	**0.50-0.79**	**0.98-0.99**	**0.31- 0.72**	**0.97-0.99**	**0.26-0.56**

Table 2. Correlation and coefficient of determination ranges between observed and predicted test-day yields in 2010. Correlation and coefficient of determination (R^2) is calculated for each month in 2010 and range of values is reported. Total yield is sum of all test-day observations of animals within herd.

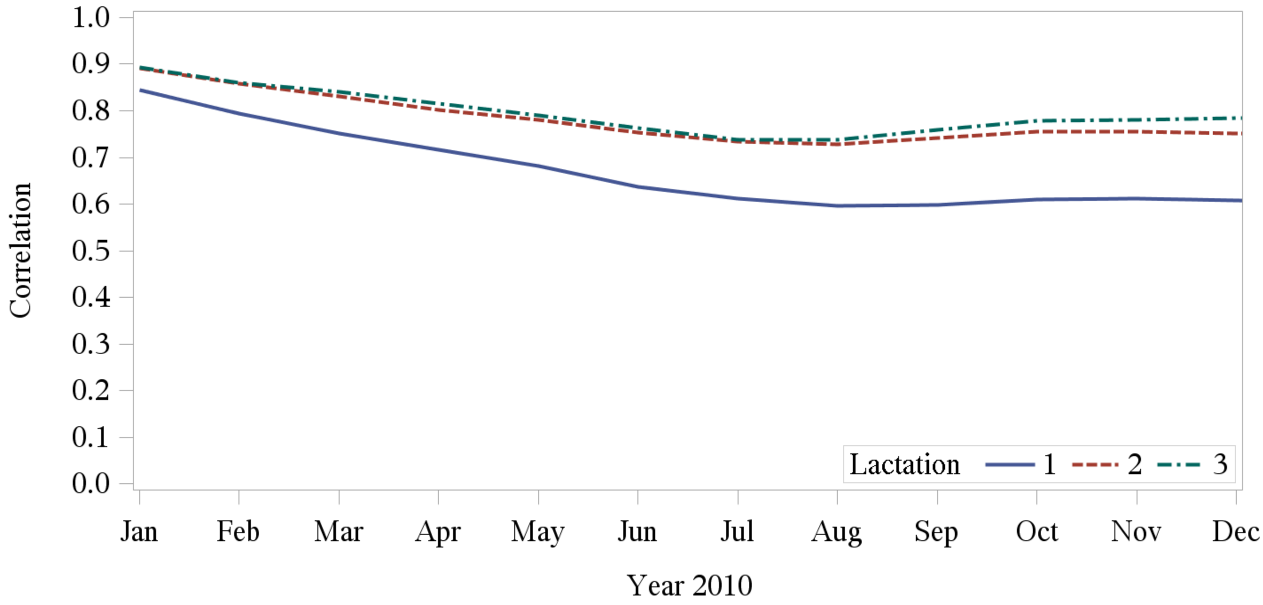

Figure 3. Correlation between test-day observations and predicted test-day milk yields in 2010. The correlation for mature cows is higher because of information from earlier lactations. Predictions are based on an evaluation with last test-day data in December 2009.

Conclusions

RRM solutions can be used to make short term prediction of cow and herd yields in near future with satisfactory accuracy. The accuracy decreases the further the predictions are made. The predictions can be used in web-application as a tool for the herd owners for management decisions.

Prediction of the future herd management effect was found to be difficult and main reason for the decrease of correlation and R^2. Also prediction of calving and insemination date is critical. Therefore, inclusion of occurred events and daily updates of the predictions is important.

References

Lidauer M., Pedersen J., Pösö J., Mäntysaari E. A., Strandén I., Madsen P., Nielsen U.S., Johansson K., Eriksson J.-Å. and Aamand G. P. 2006. Joint Nordic test day model: Evaluation model**.** Interbull Bull. 35:103–107

Koivula, M., Nousiainen, J. I., Nousainen, J, Mäntysaari, E. A. 2007. Use of Herd Solutions from a Random Regression Test-Day Model for Diagnostic Dairy Herd Management. Journal of Dairy Science Vol. 90, Issue 5: 2563-2568