Comparison of Persistency Definitions in Random Regression Test Day Models

G. J. Kistemaker
Canadian Dairy Network
150 Research Lane, Suite 207
Guelph, Ontario, Canada, N1G 4T2

Introduction

The Canadian Test Day Model currently calculates persistency as the breeding value for milk yield on day 280 minus the breeding value for milk yield on day 60, which is then expressed as the yield on day 280 as a percentage of the yield on day 60. This measure of persistency reflects the producer's idea of persistency and is well accepted within Canada.

Some concerns have been expressed internationally as to the validity of this measure when used in a test day model based on Legendre Polynomials. Canada is implementing a test day model based on Legendre Polynomials for the Holstein breed in May 2003. Therefore, this research was initiated to see if other measures of persistency that are currently in use in other countries are more suitable for use in Canada when using a test day model based on Legendre Polynomials.

Materials and Methods

Canadian Holstein test day records and pedigree information were extracted from the database at the beginning of October 2002. These files were used to estimate breeding values using the Canadian Test Day Model that had been modified to use 4th order Legendre Polynomials rather than the Wilmink curve. Resulting solutions for milk production (five parameters for each of three lactations) from all 2,930,494 animals were used to calculate breeding values for 305-day lactation yields and different measures of persistency.

Measures of Persistency

Several different measures of persistency are currently in use worldwide. In Canada the measure of persistency is essentially the breeding value for milk yield on day 280 minus the breeding value for milk yield on day 60. Test day models using Legendre polynomials allow for more variability than models using the Wilmink curve therefore this formula was changed to use intervals rather than single daily yields when applied to a test day model using Legendre Polynomials. The formula used to calculate persistency based on the Canadian method (P_{CAN}) can be written as:

$$P_{CAN} = \frac{1}{51}\sum_{i=255}^{305} EBV_i - \frac{1}{21}\sum_{i=50}^{70} EBV_i$$

where

EBV_i is the estimated breeding value for milk yield on the i^{th} DIM.

In Finland, persistency is calculated for Milk, Fat and Protein using a formula (P_{FIN}) that can be written as follows: (Pösö, 2003)

$$P_{FIN} = \frac{1}{200}\sum_{i=101}^{300} EBV_i - EBV_{100}$$

The formula used in the Netherlands to calculate persistency (P_{NLD}) can be written as follows: (De Roos et al, 2001)

$$P_{NLD} = \frac{1}{245}\sum_{i=61}^{305} EBV_i - EBV_{60}$$

The measure of persistency used in the Netherlands is based on combined grams of fat and protein. Canada will continue to use a measure of persistency based on milk yield and therefore all measures of persistency were applied to milk yield in this research.

Results

In order to determine if these three measures of persistency resulted in similar rankings of animals the correlations between the measures of persistency were calculated. Correlations between persistency measures were high (Table 1) but were considerably less than unity indicating that the different measures resulted in re-ranking of animals. Correlations between P_{CAN} and the other two measures were similar and consistent across lactations. Correlations between P_{FIN} and P_{NLD} were relatively low, especially in the first lactation. This may be due to when peak production occurs since in first lactation, peak production (in this dataset) occurs on average at 68 DIM. Cows with a later peak will tend to have a lower EBV_{60} and higher EBVs after day 60 and as a result have a higher P_{NLD}, at the same time EBV_{100} is increased considerably due to the late peak and as a result P_{FIN} tends to be lower. In second and third lactations peak production occurs on average at 50 and 53 DIM respectively, and this would tend to cause fewer differences between the two methods.

Table 1. Correlation between the three different measures of persistency in the first three lactations.

	Lactation 1	Lactation 2	Lactation 3
P_{CAN}-P_{FIN}	0.960	0.963	0.952
P_{CAN}-P_{NLD}	0.958	0.962	0.965
P_{FIN}-P_{NLD}	0.894	0.936	0.931

An important characteristic of the persistency measure used is the correlation with 305-day lactation EBVs. All three measures of persistency used in this research had low correlations with 305-day lactation milk yield EBVs (Table 2). In all three lactations the correlation between P_{CAN} and the lactation EBVs was closest to zero and ranged between 0.06 and 0.08. P_{NLD} had a correlation with lactation EBVs that ranged from 0.12 to 0.19, in contrast, P_{FIN} had a negative correlation with lactation EBVs that ranged from –0.12 to -0.09.

Table 2. Correlation between the three measures of persistency and 305-day lactation milk yield EBVs in the first three lactations.

	P_{CAN}	P_{FIN}	P_{NLD}
Lactation 1	0.060	-0.100	0.186
Lactation 2	0.061	-0.116	0.122
Lactation 3	0.082	-0.094	0.127

Correlations between the same measure of persistency in different lactations are shown in Table 3. This table shows that correlations between second and third lactation persistency proofs were the highest and correlations between persistency proofs in first and third lactation were lowest. The correlations between persistency proofs in different lactations when using P_{CAN} or P_{FIN} were very similar but correlations for P_{NLD} were somewhat lower. This indicates that P_{NLD} is less consistent across lactations than the other two measures. This may be related to when peak production occurs and shows that P_{NLD} is less desirable for calculating persistency of milk production in Canada.

Table 3. Correlations between persistency proofs in different lactations for the three measures of persistency.

Lactation	P_{CAN}	P_{FIN}	P_{NLD}
1 and 2	0.693	0.684	0.604
2 and 3	0.860	0.865	0.829
1 and 3	0.542	0.530	0.453

It is important to look at how important the 'cutoff' points in the formula are when calculating persistency proofs. In order to look at this the formulae used to calculate P_{CAN} and P_{NLD} were changed so that they compared persistency relative to DIM 70 and DIM 80 rather than the 60 DIM, which both formulae currently use. In P_{CAN} the second part of the formula was changed so that this part contained the 21 days centered at 70 (or 80)

DIM. In P_{NLD} the first part of the formula was changed so that this interval started at 71 (or 81) DIM and was divided by 235 (or 225) rather than 245, the second part was changed to use EBV_{70} (or EBV_{80}) instead of EBV_{60}. Correlations between the different versions of the same persistency formula are shown in Table 4. This table shows that P_{NLD} is more sensitive to minor changes in the formula than P_{CAN}.

Table 4. Correlations between persistency proofs when using different DIM in the formulas used to calculate persistency.

1st Lactation	P_{CAN}	P_{NLD}
60 DIM - 70 DIM	.995	.989
70 DIM - 80 DIM	.997	.992
60 DIM - 80 DIM	.984	.962

Discussion and Conclusions

From the three measures of persistency studied in this research it is clear that P_{NLD} has several undesirable properties, at least when applied to milk yield in the Canadian Holstein population. It has the highest correlations with 305-day Lactation EBVs (although they are still low), has the lowest correlations across lactations and is more sensitive to minor changes in its definition. The poor performance of P_{NLD} in this research seems to be due to the location of peak production in this data set, which in first lactation is after 60-DIM. Another reason could be that in this research the formula is applied to milk EBVs rather than fat plus protein EBVs as is done in the Netherlands.

The performance of P_{FIN} and P_{CAN} was very similar with slightly better properties for P_{CAN} in most cases. P_{CAN} has been used in Canada for more than four years and the other persistency measures that were investigated did not show any benefits over the current procedure. Therefore, there is no reason for Canada to change it persistency definition to any of the other measures investigated.

References

Pösö, J. 2003. Personal communication.

De Roos, A.P.W., Harbers, A.G.F. & de Jong, G. 2001. Random regression test-day model in The Netherlands. *Interbull Bulletin 27*, 2001.