PEDIGREE INDEXING OF YOUNG HOLSTEIN BULLS BASED ON ANIMAL MODEL EVALUATIONS

W. C. ZHANG, L. R. SCHAEFFER[2], J.C.M. DEKKERS[2] and H. DEBOER

Semex Canada, 130 Stone Road W. Guelph, ON, Canada
[2]Department of Animal & Poultry Science University of Guelph Guelph, ON, Canada

SUMMARY

Four methods of pedigree indexing young bulls for artificial insemination were compared relative to subsequent progeny tests on the same bulls. Because many young bulls entered in Canadian AI units are progeny of parents from the United States, formulas for converting US genetic evaluations into Canadian equivalents were utilized. Indexing Method 1 averaged the (converted) genetic evaluations of the sire and dam. Method 2 was based on derived coefficients from the regression of progeny genetic evaluations on sire and dam genetic evaluations. Method 3 was the same as Method 2 except Canadian dam genetic evaluations were adjusted to eliminate the influence of their son(s) on their genetic evaluations due to an automatic correlation between sons and dams arising from animal model evaluation procedures. Method 4 was the same as method 1 except the adjusted dam index was used for parental averages. Pedigree indexes from Method 1 overpredicted young bull merit on average, while the other three methods gave means similar to progeny test evaluations. Method 2 gave more weight to Canadian dam genetic evaluations, because of the auto-correlation with sons' evaluations, than on US dam genetic evaluations, which did not include their Canadian sons' information. Method 3 gave the best predictions, but rankings were similar to those of Method 1. Correlations among the pedigree indexes and eventual proofs for traits in the Lifetime Profit Index ranged from .6 to .8, and were similar for all four methods. Of the top 40% of young bulls on the basis of pedigree indexes, 86% (93 out of 108) were also in the top 10% on actual proofs. Not one bull with a PI in the bottom 30% had an actual proof that reached the top 10%, except 2 bulls when PI were computed by method 4.

INTRODUCTION

In dairy cattle breeding, potentially superior young bulls need to be identified early so that enough doses of semen can be collected and stored prior to the demand for that bull once his progeny test becomes available. Pedigree indexes (PI) have been utilized for this purpose for many years. One of the weaknesses of PI has been the accuracy of estimated transmitting abilities (ETA) for bull-dams. Although the implementation of the animal model greatly increased the accuracy of cow ETA, there is still a tendency of overpredicting ETA for bull-dams (Ferris et al. 1991, Graham et al. 1991, Mao et al. 1991, Uimari et al. 1992,).

PI could be produced simultaneously by the animal model evaluation itself by including future young bulls in the relationship matrix, even though they do not yet have any daughters in the data. This PI is equivalent to the average ETA of the sire and dam. Another method would be to regress son's ETA on sire and dam ETA, with allowances for Canadian or US sources of information (Schaeffer 1981, Jansen et al. 1986). However, in animal model evaluation procedures, the ETA of the dam includes a major influence from her progeny, particularly if that progeny is an AI sire with many progeny. Thus, regressing sons' ETA on dams' ETA tends to bias the regression coefficient upwards on dams' ETA values. Ideally, one would need to regress ETA of the son on sire and dam ETA obtained from before the son received progeny test proof. However, there are some practical difficulties if the information were not from the same evaluation run. Otherwise, the influence of the son's ETA on the dam's ETA needs to be removed prior to use in regression procedures. For an analysis of Canadian bulls, only bulls with Canadian dams affected because Canadian data were until recently not included in USA dam ETA.

The objectives of this study were to derive a method for removing the influence of the son's ETA on the dam's ETA and to compute PI on young bulls by different methods and compare them on the basis of accuracy and bias when predicting eventual progeny test ETA values. Another objective was to examine the probability of success (PS) in using PI for ranking young unproven bulls, PS is defined as the number of bulls that ends up in top 10% eventual proof LPI over number selected on the PI for LPI at different percentiles.

MATERIAL AND METHODS

From 1984 to 1993 a total of 2942 Holstein bulls were entered into Canadian AI centres for progeny testing, of which 1235 have now received official progeny test evaluations (Table 1). The traits considered in this study were those included in the Lifetime Profit Index (LPI), which has the formula

$$LPI = 6 [3 \text{ Fat} + 8 \text{ Protein}] + 4 [3 \text{ FC} + 4 \text{ MS} + 2 \text{ FL} + \text{CAP}]$$

where Fat and Protein are the yield ETA values in BCA units, FC, MS, FL, and CAP are ETA of conformation traits for final class, mammary system, feet and legs, and capacity, respectively. ETA information from the spring 1993 was obtained from Agriculture Canada for fat and protein yields, from Holstein Canada for conformation traits, and from the American Holstein Association for production and type PTA of USA sires and dams of Canadian AI bulls. American final score, udder composite, feet and legs composite, and body depth were chosen to correspond to the Canadian traits of final class, mammary system, feet and legs, and capacity, respectively. US dams did not have ETA values for feet and legs composite or body depth, and therefore ETA values of the maternal grandsires (MGS) were used. Pedigree indexes were derived using 4 methods.

<u>Method 1</u> requires averaging the ETA values of the sire and dam of each bull. To apply this method, ETA values of USA sires and dams had to be converted to Canadian equivalents. The 1993 official conversion formulas in Canada for the traits in the LPI were used (Robinson 1993, Lohuis and Burnside 1993).

<u>Method 2</u> was a regression procedure which followed the model

$$ETA(son)_{ijk} = CS_i + CD_j + b_{1i}ETA(sire) + b_{2j}ETA(dam) + e_{ijk}$$

where CS_i is the country effect of the sire (Canadian or US origin of ETA), CD_j is the country effect of the dam, b_{1i} and b_{2j} are regressions on the sire and dam ETA with different regressions depending on the country of proof. Formulas were also derived using sire and MGS ETA for all bulls, so that bulls with missing dam ETA could be indexed.

<u>Method 3</u> was the same as Method 2 except that the ETA values of Canadian dams were adjusted to remove the autocorrelation effect of their son's ETA on their own ETA. The adjustment procedure is based on re-creating the right hand side (RHS) element of the mixed model equation for the dam, using her ETA and those of her sons, and dividing that amount by the diagonal of the mixed model equations for that dam when ignoring the fact that she had any sons at all. The approximate RHS element is given by

$$RHS(dam) = [NL + k(2 + .5(NS + ND))]ETA(dam) - \Sigma k\, ETA(sons)$$

where NL is the number of lactations on the dam, NS is the number of sons of the dam, ND is the number of daughters, and k is the ratio of residual to additive genetic variances used in the animal model. To obtain a new dam ETA without the influence of her sons, divide RHS(dam) by

$$[NL + k(2 + .5(ND))]$$

which is the appropriate diagonal had the dam not had any sons. The new dam ETA values of CAN dams were used in the regression approach of Method 2. There were both CAN and USA sires mated to CAN dams, and their contributions to PI were estimated by method 2.

<u>Method 4</u> was a modification of method 1, in which for the CAN dam group, the adjusted dam ETA from method 3 was used to obtain the parental averages. This method was only applied to group with CAN dams. Otherwise, method 1 was used for other groups.

Bulls with CAN dams were used to compare all four methods of generating PI. To measure the goodness of fit between the four methods and actual proofs, the following formula was used:

$$X = \Sigma (PI - proof)^2 / \Sigma\, proof^2$$

The smaller the value of X (fitness), the better would be the method.

In order to examine the success of using the different PI methods, success has to be defined.

Using the PI for each trait, a PI for LPI can be calculated. Of the 1078 bulls (with all PI and proof information available) there were 108 bulls classified as successful bulls (in top 10% on proof LPI).

Table 1. Bulls entered in Canadian AI units in current analysis

Year of entry	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	Total
no. entered	138	140	210	299	302	334	341	375	395	408	2942

RESULTS

The estimated regression coefficients for CAN bull dam from Methods 2 and 3 are shown in Table 2. The coefficients for Canadian dam ETA in Method 3 were significantly lower for all traits, and correspondingly, the R^2 values for the regressions. The change in regressions on Canadian dams caused slight increases in coefficients for Canadian sires and also changes in coefficients on US information, but without a particular trend upwards or downwards. The relative values of dam to sire regressions for protein yield, for example, were 1.26 and .67 for Canadian and USA parents respectively by Method 2, and were correspondingly .67 and .66 for Method 3. Thus, the adjustment of Canadian dam ETA for the autocorrelation with their son's ETA seems to be effective in terms of restoring the relative weighting of sire and dam information.

Figure 1 shows mean PI by all four methods for 389 bulls and their new proof in 1994. These bulls were not included in the regression analysis because they had not yet received a progeny test in Spring, 1993. Method 1, the average of parent ETA, continuously over-predicted the actual LPI while Methods 2 and 3 followed actual LPI better. Method 4 also tended to overpredict the actual LPI, but to a lesser extent than method 1. The correlations among PI and proof LPI were .59 for first three methods, and .56 for method 4.

Results from the goodness of fit test on bulls from CAN dams (Fig. 2) showed that there were no substantial differences among the first three methods, but method 4 showed larger disagreement between PI and actual proof LPIs. The trend was that the biases were higher in earlier years than that in late years (Fig. 2).

Probability of success (PS) was estimated for the 5th and every 10th percentile. Results (Fig. 3) showed that all four methods had a similar prediction of PS. Although in the first 10th percentile of PI for LPI, method 2 was slightly better (PS was .44 by method 2, .42 by method 3 and method 1). But for the top 20th percentile, method 1 seemed better (PS was .33 for method 1, .32 for method 2, and .30 for method 3 and method 4). On average, the top 40th percentile of PI for LPI included 86% of the successful bulls (93/108). For the first three methods, the bottom 30th percentile did not include a single successful bull, except 2 bulls when PI were computed by method 4. Method 4 was not as good as the other methods in predicting the PS.

Figure 1. Mean pedigree index LPI and proof

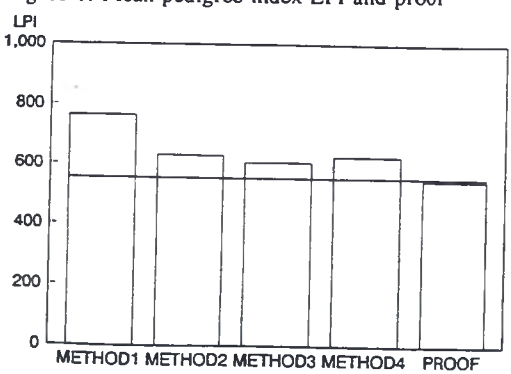


Figure 2. Fitness test between PI and Proof

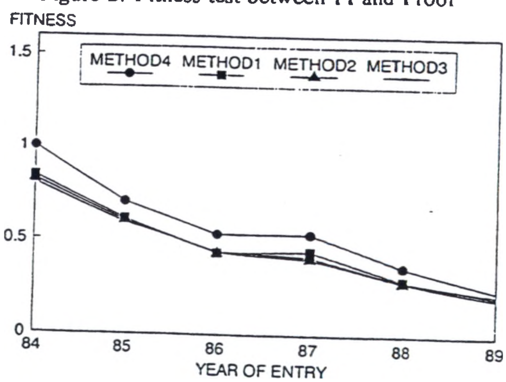
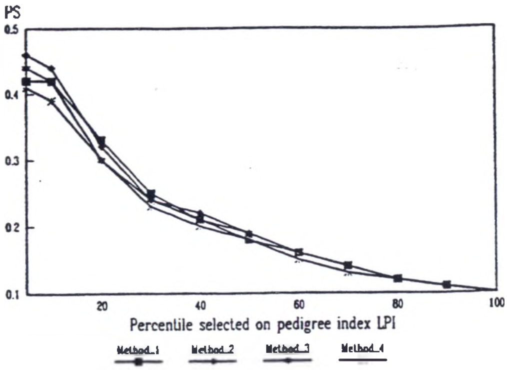


Table 2. A comparison of DAM regression factors before and after adjustment for Canadian bull dams.

Trait	Method	a (intercept)	b (slope)	R^2
Fat	2	-1.10176	.53830	.35
	3	- .45660	**.34318**	.28
Protein	2	-.69367	.53740	.41
	3	-.22325	**.30215**	.34
Final Class(FC)	2	-.06640	.59130	.57
	3	.22485	**.50332**	.52
Mammary System(MS)	2	.01036	.63076	.52
	3	.27829	**.53246**	.45

Figure 3. Probability of success (PS, bulls whose proof LPI reaches top 10% over the number selected on PI LPI).



DISCUSSION

Method 1 overpredicted son's ETA in USA and Finland as well as in Canada (Uimari et al. 1992, Mao et. al. 1991. Ferris et al. 1991. Graham et. al. 1991). The lack of accuracy of bull dam ETA, primarily due to preferential treatment, is the major problem in identifying superior cows. More extensive usage of MOET maybe one of the measures to increase the accuracy of cow EBV by progeny testing cows through a lot of daughters. The autocorrelation between son and dam ETA values becomes a problem for regression only after the sons have received their proofs. Using the procedure in this study does provide a reasonable dam ETA for regression analysis. Directly applying the adjusted ETA in a parental average was less favourable compared to other methods. Generally speaking, in terms of ranking bulls based on PI, although Method 1 had inflated PI (Fig. 1) it ranked bulls reasonably well. If only ranking within sampled bulls is the main concern, method 1 provides a easy approach. However, method 3 could provide a good prediction in both the mean ETA and ranking of young bulls. Although it is not easy to apply at farm level and for AI personnel who need a simple method for day to day decision making.

The probability of success estimated on real data in the current study were higher than Lohuis et al. (1992) reported based on theoretical studies.

REFERENCE

FERRIS, T.A. and G.R. WIGGANS. (1991) J. of Dairy Sci. 74 (Suppl.):265 (Abstr.).
GRAHAM, N.J., C. SMITH and J. GIBSON. (1991) Canadian J. of Animal Aci. 71:21.
JANSEN, GG, L.R. SCHAEFFER, E.B. BURNSIDE, H.M. WILSON (1986). Can. J. Anim. Sci. 66:889.
LOHUIS, M.M., J.C.M. DEKKERS, and C. SMITH. (1992) J.Dairy Sci. 75:1660.
LOHUIS, M.M. and E.B. BURNSIDE. (1993) USA-Canada conversion (type traits), CGIL.
MAO, I.L., M.C. DONG and C.E. MEADOWS. (1991) J. of Dairy Sci. 74:2747.
ROBINSON. A. (1993a) USA-Canada conversion (production traits). Agriculture Canada.
SCHAEFFER, L.R. (1981) Can. J. Anim. Sci. 61:261.
UIMARI, P. and E.A. MANTYSAARI. (1992) J. Dairy Sci. 75 (Suppl.):246 (Abstr.).