Proc. Estonian Acad. Sci. Biol., 1993, 42, 4, 241—248
https://doi.org/10.3176/biol.1993.4.01

PREDICTION OF ADDITIVE GENETIC VALUE OF DAIRY CATTLE

1. PEDIGREE INDEX BASED ON ANCESTORS' RECORDS

Rein TEINBERG

Eesti Teaduste Akadeemia Eksperimentaalbioloogia Instituut (Institute of Experimental Biology, Estonian Academy of Sciences), Instituudi tee 11, EE-3051 Harku, Harjumaa, Eesti (Estonia)

Received May 18, 1993; accepted May 18, 1993

Abstract. The theoretical background of calculating a pedigree index for predicting the breeding value of young dairy bulls and cows is described. The principles and formulas for calculating weight factors for the pedigree index are presented. The results of simulation studies on computer are given. The maximum theoretical accuracy of prediction when only the pedigree information (breeding values of the dam, maternal granddam, sire and maternal grandsire) was used was 0.62. The addition of cow's own records (5 lactations) increased the accuracy to 0.78. The data on maternal grandsire and granddam result in an insignificant increase in the accuracy of prediction if we have the data on the dam and the sire's progeny testing.

The pedigree index recommended for practical use with Estonian cattle breeds was

$$PI = 0.45(BV_S - \bar{P}_S) + 0.15(BV_D - \bar{P}_D) + 0.22(BV_{mGS} - \bar{P}_{mGS}),$$

where BV_S, BV_D and BV_{mGS} are the breeding values of sire, dam and maternal grandsire, and $\bar{P}_S$, $\bar{P}_D$ and $\bar{P}_{mGS}$ are the population means in the sire's, dam's and maternal grandsire's generations, respectively.

Key words: dairy cattle, breeding value estimation, pedigree index.

INTRODUCTION

The most crucial factor in genetic improvement of dairy cattle is the accurate prediction of the additive genetic value of cows, bulls and young cattle. For selection the breeder must correctly rank his animals on the basis of their breeding value evaluation.

As the selection characters of cattle are influenced by genes at many loci, each with a relatively small effect, the phenotypic expression of such traits is usually measured quantitatively. Production traits in cattle are quantitative in terms of measurement and in the number of genes involved. The sum of all genes' effects is considered the additive genetic value or breeding value of the animal.

A quantitative trait can be expressed in phenotypic parameters as follows (Van Vleck et al., 1987; Henderson, 1984):

$$P_{ij} = \mu + G_i + PE_i + TE_{ij},$$

where P_{ij} is the jth record of the ith animal;

μ is the average value of the genes that all animals in the population have in common which determines the constant level of performance;

G_i is the sum of genetic values;

PE_i is the sum of effects of environmental factors, which permanently influence the performance of animal i; and

TE_{ij} is the sum of random (temporary) environmental effects which affect only the jth record of animal i.

The expression characterizes the basic problem of genetic selection: only the phenotype, P_i, can be observed, but animals that have the best genetic values, G_i, must be selected. The environmental factors, E_i, however, prevent direct measurement of G_i. The use of records of the animal and its relatives to predict breeding value will be discussed here.

Breeding value estimation of a selection trait involves a proper weighing of phenotypic records of the animal itself and its relatives. Records of animals more closely related to the animal being evaluated have more weight than records of distant relatives, depending on common genes and gene effects (Van Vleck, 1983; Aitchison, 1989).

The calculation of appropriate weights for the animal's own records and for its relatives' records is called the selection index procedure. The index, denoted as $\widehat{A}_i$ (an estimate of A_i, the true additive genetic value of an animal i), is (Becker, 1985; Pirchner, 1983; Mäki-Tanila and Juga, 1988)

$$\widehat{A}_i = b_1X_1 + b_2X_2 + \ldots + b_NX_N,$$

where b_1 is the weight (standardized linear regression coefficient) for X_1, the record on relative 1;

b_2 is the weight for the record X_2 of relative 2; and b_N is the weight for X_N, the record on the Nth relative.

Selection index characterizes most accurately the true additive genetic value of an animal. On its basis animals can be ranked as accurately as possible according to their breeding values (additive genetic values). Of course, the true A_i for a single animal, as all statistical parameters, is never known precisely. However, statistical and quantitative genetic theory allows the maximizing $r_{A,\widehat{A}}$ — the correlation between true A_i and $\widehat{A}_i$, or the prediction of A_i (Hillers et al., 1982).

Chronologically, the first source of information for predicting the breeding value of an animal is its pedigree—ancestors' records. In dairy cattle the phenotype of the main production character — milk yield — can be recorded only in females. Therefore, estimating bulls' breeding value is possible only on the basis of their ancestors (in the case of young bulls) or the average phenotype of progeny. In cows their own phenotype (milk recording data) can be added (Vinson and White, 1982).

In this part of the paper only the use of ancestors' records in predicting the breeding value of cows and bulls will be discussed.

MATERIAL AND METHODS

It was decided to change the old system of estimation of the breeding value in Estonia as of January 1993 and to use two types of breeding value indexes: the pedigree index, *PI*, which gives a preliminary estimate of the breeding value of young cows and bulls, and the selection index, *SI*, which includes weights for three main selection traits — milk yield, protein yield and fat yield. The experimental procedures included: 1) simulation studies on computer of various pedigree indexes and their testing on cattle records, and 2) simulation studies of different models of selection indexes for several characters, based on the cow's own records only. In the next part of the paper the results of these studies as well as the results of testing the models using field data from the data bank of cows' records will be presented.

Simulation studies were performed on a personal computer using special programmes. For the testing of models data on the production control of the Estonian Black and White (EBW) and Estonian Red (ER) cattle breeds were used (about 220,000 cows).

The mathematical methods of quantitative genetics were used for predicting the breeding value of animals on the basis of ancestors' records. In dairy cattle, as a rule, three (sometimes four) main sources of pedigree information are used (Van Vleck et al., 1987; Funk and Hansen, 1988):

1) sire's breeding value (progeny test results, half sisters of animal being evaluated);

2) maternal grandsire's breeding value (also on the basis of progeny test, half sisters of the dam of animal being evaluated);

3) dam's breeding value, estimated by her own lactation records (one or several lactations);

4) maternal granddam's breeding value; however, if the first three sources of information are available, maternal (or paternal) granddam's data will add practically nothing to the exactness of the breeding value estimate. We did not use this source of information.

The method of path coefficients was used in estimating the value of the various criteria contributing to the breeding value evaluation by pedigree information of an animal. This method assumes that an animal gets half of genes from each parent, so the path coefficient between the genotype (additive genetic value) of parent and progeny will be 0.5. Correlation (path coefficient) between the genotype and phenotype of an animal has been denoted as h, which can be determined as the square root of heritability h^2 (Lush, 1965; Dempfle, 1982; Van Vleck, 1983). To obtain the correlation between the animal's breeding value (additive genetic value A_i) and the phenotype of its relative P_i, it is necessary to multiply all the intervening paths. For example, for the dam it will be $0.5h$, for granddam or half sister $0.25h$, and for a daughter (or son) $0.5h$. The regression of A_i on the phenotype of a relative (e.g. dam) is obtained by multiplying the correlation coefficient by the ratio of the standard deviation of A_i and the phenotypic standard deviation of the dam, $\sigma_{A_i}/\sigma_{P_D}$. This ratio is equal to h and consequently the regression of A_i on P_D is $0.5h^2$.

When the animal has many sibs or progeny, its breeding value can be estimated with more accuracy. The correlation between A_i and the phenotypic mean of n half-sisters, also daughters, is calculated by multiplying the correlation between A_i and the phenotype of one member of a group by $n/\{1+(n-1)r_gh^2\}$, where r_g is the coefficient of relationship between family members and r_gh^2 is the phenotypic correlation in a randomized environment. The regression coefficient is accordingly

$$nr_gh^2/\{1+(n-1)r_gh^2\}.$$

In dairy cattle more accurate information about the breeding value of the sire than that of the dam is usually available, because bulls are progeny tested on a large number of daughters while the cow has only her own production records. Therefore, the question is how much weight to give to the various criteria (ancestors' records) in the *PI*. This question is rather complicated and involves calculating the partial regression of the additive genetic value, A_i, on each criterion used (dam's records, maternal grandsire's and sire's records). The standard partial regression coefficients k_i (or weight factors for multiplying the ancestors' records) between the ancestor's breeding value and the individual's phenotype are calculated on the basis of heritability, h^2, of the trait (Danell, 1981; Dempfle, 1982).

In the case of several records (for example several lactations) for the animal, h^2_m was calculated as follows:

$$h^2_m = \frac{mh^2}{1+(m-1)R},$$

where R is repeatability, m is the number of records, and h^2 is heritability for one record per animal.

Table 1

Weight factors, k_i, for the records of ancestors (expressed as phenotypic deviations from population mean) and the multiple correlation (accuracy), R_{AI}, between the breeding value of the individual, A_i, and calculated index *PI*

Source of information	Weight factors, k_i	Accuracy, R_{AI}
Individual	$k_x = h^2_x$	$\sqrt{h^2_x}$
Sire, dam	$k_1 = 0.5h^2_1$ $k_2 = 0.5h^2_2$	$\sqrt{0.5(k_1+k_2)}$
Dam, maternal granddam	$k_2 = \frac{h^2_2(4-h^2_6)}{2(4-h^2_2 \cdot h^2_6)}$ $k_6 = \frac{2h^2_6(1-h^2_2)}{2(4-h^2_2 \cdot h^2_6)}$	$\sqrt{0.5k_2+0.25k_6}$
Dam, maternal grandsire and granddam	$k_2 = \frac{h^2_2(4-h^2_5-h^2_6)}{2\{4-h^2_2(h^2_5+h^2_6)\}}$ $k_5 = \frac{2h^2_5(1-h^2_2)}{2\{4-h^2_2(h^2_5+h^2_6)\}}$ $k_6 = \frac{2h^2_6(1-h^2_2)}{2\{4-h^2_2(h^2_5+h^2_6)\}}$	$\sqrt{0.5k_2+0.25(k_5+k_6)}$
Four grandparents	$k_3 = 0.25h^2_3$; $k_4 = 0.25h^2_4$; $k_5 = 0.25h^2_5$; $k_6 = 0.25h^2_6$	$\sqrt{0.25(k_3+k_4+k_5+k_6)}$
Individual, sire, dam	$k_x = \frac{h^2_x(4-h^2_1-h^2_2)}{4-h^2_x(h^2_1+h^2_2)}$ $k_1 = \frac{2h^2_1(1-h^2_x)}{4-h^2_x(h^2_1+h^2_2)}$ $k_2 = \frac{2h^2_2(1-h^2_x)}{4-h^2_x(h^2_1+h^2_2)}$	$\sqrt{k_x+0.5(k_1+k_2)}$

Symbols:	h^2_i	k_i
Sire (S)	h^2_1	k_1
Dam (D)	h^2_2	k_2
Paternal grandsire (pGS)	h^2_3	k_3
Paternal granddam (pGD)	h^2_4	k_4
Maternal grandsire (mGS)	h^2_5	k_5
Maternal granddam (mGD)	h^2_6	k_6
Individual (X)	h^2_x	k_7

If the sire had progeny his h^2_n was calculated as follows:

$$h^2_n = \frac{0.25nh^2}{1+(n-1)0.25h^2}.$$

In models varying h^2 estimates as well as repeatabilities and lactation and progeny numbers were used.

We used the coefficients k_i calculated according to the recommendations of Eriksson and Danell (1984), Van Vleck et al. (1987), and Johansson and Rendel (1968). The calculation formulas for k_i and the multiple correlations, R_{AI}, between the true breeding value, A_i, and calculated PI are given in Table 1.

RESULTS AND DISCUSSION

Using the formulas given above we calculated the theoretical accuracy, R_{AI}, of estimating the breeding value on the basis of the pedigree index (records of ancestors). The results of some variants most appropriate for dairy cattle are presented in Table 2.

Table 2 shows that the maximum accuracy of predicting the breeding value of a cow (bull) on the basis of only pedigree information is only $R_{AI}=0.62$. So, the breeding value of a young cow can be estimated with the accuracy only slightly over 0.5.

The use of embryo transfer and deep freezing of embryos (cows will be produced from embryos after their paternal half sisters have completed their first lactation) enables to increase the accuracy, but not more than by 10 to 20 per cent (Van Vleck et al., 1987). Besides that, the generation interval will be prolonged and the selection effect reduced. The data on the breeding value of maternal granddam and grandsire give only an insignificant increase in the accuracy (R_{AI} increases only by 0.02) when sire is progeny tested. By increasing the number of lactations of dam and maternal granddam to 5 (instead of 3) and progeny number to 200 (instead of 100), we again achieve only a small increase — from 0.59 to 0.62 — in the accuracy of prediction.

Table 2

Accuracy, R_{AI}, of breeding value estimates using ancestors' phenotypic records ($h^2=0.25$ and repeatability $R=0.40$)

Sources of information					Accuracy, R_{AI}
D	mGD	S	mGS	X	
No. of lactations, m		No. of progeny, n		No. of lactations, m	
3	0	100	0	0	0.57
3	3	100	100	0	0.59
5	5	200	200	0	0.62
1	0	100	0	3	0.72
3	3	100	100	3	0.75
5	5	200	200	5	0.78

When cow's own records are considered together with ancestors' data the accuracy of predicting its breeding value increases maximally to 0.78 (Table 2). An increase in the number of lactations from 3 to 5 increases the accuracy of predicting the breeding value only by 0.03. This is of little use as the generation interval also increases by 2 years. Maternal granddam's and maternal grandsire's data add but little to the predicting accuracy: it increases from 0.72 to 0.75.

The calculations gave the result that in the case of $h^2=0.3$ and only four grandparents' breeding values being used, we can predict the breeding value of an animal (young cow or bull) with the accuracy of 0.27 only. An increase in h^2 to 0.6 gives the accuracy of 0.39.

It can be concluded from our study that in predicting the breeding value of an animal only on the basis of its grandparents' milk, milk protein or milk fat production ($h^2=0.3$) the accuracy of the estimate will be rather low ($R_{AI}<0.3$). Inclusion of the parents' breeding value figures raises the accuracy to $R_{AI}\approx 0.6$, which can be considered the maximum accuracy for milk yield on the basis of pedigree information only. If the cow's own milk production data are available, the accuracy will increase to $R_{AI}\approx 0.8$, which is practically the maximum for a dairy cow.

Proceeding from theoretical calculations a pedigree index for predicting young bulls' and heifers' breeding value (additive genetic value) was proposed for Estonian cattle breeds. The breeding value estimates of sire, S, dam, D, and maternal grandsire, mGS, were included in the index. This index can be used for predicting the breeding value with reference to milk yield as well as milk protein and fat yield (or their sum). The heritability in first lactation (h^2_1) was established at $h^2=0.3$.

Index weights, k_i, were calculated as follows:

1. For the data on sire:

$$k_1=0.5\cdot h^2_1,$$

where

$$h^2_1=\frac{n\cdot h^2\cdot 0.25}{1+(n-1)0.25h^2},$$

where n is the number of daughters tested and $h^2=0.3$.

Example: when $n=100$ and $h^2=0.3$,

$$h^2_1=\frac{100\cdot 0.3\cdot 0.25}{1+(100-1)0.25\cdot 0.3}=0.890,$$

$$k_1=0.5\cdot 0.890=0.445.$$

2. For the data on dam:

$$k_2=0.5\cdot h^2_2,$$

where $h^2_2=0.3$ (even when dam has data on several lactations).

Example: $k_2=0.5\cdot 0.3=0.150$.

3. For the data on maternal grandsire:

$$k_5=0.25\cdot h^2_5,$$

where

$$h^2_5=\frac{n\cdot h^2\cdot 0.25}{1+(n-1)0.25h^2}.$$

Example: when $n=100$ and $h^2=0.3$, $h^2_5=0.890$, and

$$k_5=0.25\cdot 0.890=0.223.$$

4. The pedigree index is

$$PI = k_1(BV_S - \bar{P}_S) + k_2(BV_D - \bar{P}_D) + k_5(BV_{mGS} - \bar{P}_{mGS}),$$

where $k_1 \ldots k_5$ are weight factors (regression coefficients) for ancestors' records;

BV_i are the breeding values of ancestors;

$\bar{P}_i$ are the population (contemporaries') means by generations.

In our example the pedigree index is

$$PI = 0.445(BV_S - \bar{P}_S) + 0.150(BV_D - \bar{P}_D) + 0.223(BV_{mGS} - \bar{P}_{mGS}).$$

5. An example of calculating the pedigree index for a young bull.

Conditions: breeding values: $BV_S = 500$ kg milk protein + fat; $BV_D = 600$ kg, $BV_{mGS} = 450$ kg; population means: $\bar{P}_S = 400$ kg, $\bar{P}_D = 400$ kg, $\bar{P}_{mGS} = 400$ kg.

The pedigree index for the bull is:

$$PI = 0.445(500 - 400) + 0.150(600 - 400) + 0.223(450 - 400) = 85.65 \text{ kg},$$

which predicts the difference in the production of fat+protein per lactation between the contemporaries and a young bull's daughters.

6. The accuracy of prediction:

$$R_{AI} = \sqrt{0.5(k_1 + k_2) + 0.25k_5},$$

in our example

$$R_{AI} = \sqrt{0.5(0.445 + 0.150) + 0.25 \cdot 0.223} = \sqrt{0.353} = 0.594.$$

It means that the accuracy of the predicted breeding value is about 0.6.

The pedigree index procedure using the index presented above was introduced for predicting the breeding value of young *AI* bulls and heifers from the beginning of 1993.

REFERENCES

Aitchison, T. E. 1989. The USDA Animal Model Genetic Evaluations. Cooperative Extension Service Bull. No. 3091, Iowa State Univ.

Becker, W. A. 1985. Manual of Quantitative Genetics. 4th ed., Washington State Univ., Pullman.

Danell, B. 1981. Evaluation of sires on first-lactation yield of Swedish dairy cattle. Report 51, Swedish Univ. Agric. Sci., Uppsala.

Dempfle, L. 1982. Problems in estimation of breeding values. — In: 2nd World Congr. Genetics Applied to Livestock Prod., Madrid, 5, 104—117.

Eriksson, J.-A. and Danell, B. 1984. Sire evaluation for milk yield corrected for non-random matings within herds by considering the maternal grandsire of the daughter. — Acta Agric. Scand., 34, 84—96.

Funk, D. C. and Hansen, L. B. 1988. Predictability of pedigree indexes for bulls from two, three, and four generations. — J. Dairy Sci., 71, 3148—3153.

Henderson, C. R. 1984. Applications of Linear Models in Animal Breeding. Univ. of Guelph, Guelph, Ont.

Hillers, J. K., Bernard, J. W. and Williams, G. F. 1982. Predicting Genetic Merit of Young Dairy Bulls. Washington State Univ. Press, Extension Service Bull. No. 778, 102—104.

Johansson, I. and Rendel, J. 1968. Genetics and Animal Breeding. Oliver and Boyd, Edinburgh—London.

Lush, J. L. 1965. Animal Breeding Plans. Iowa State Univ. Press, Ames.

Mäki-Tanila, A. and Juga, J. 1988. Use of the individual animal model in the genetic evaluation of breeding animals. — Ann. Agric. Fenniae, 27, 231—246.

Pirchner, F. 1983. Population Genetics in Animal Breeding. 2nd ed., Plenum Press, New York—London.
Van Vleck, L. D. 1983. Notes on the Theory and Application of Selection Principles for the Genetic Improvement of Animals. Dept. of Animal Sci., Cornell Univ. Press, Ithaca.
Van Vleck, L. D., Pollak, E. J. and Oltenacu, E. A. B. 1987. Genetics for the Animal Sciences. Freeman and Co., New York.
Vinson, W. E. and White, J. M. 1982. Prediction of modified contemporary comparisons from pedigree information for artificial insemination and natural service bulls. — J. Dairy Sci., 65, 1578—1582.

PIIMAVEISTE ADITIIVSE GENEETILISE VÄÄRTUSE PROGNOOSIMINE

1. Eellaste andmetel põhinev põlvnemisindeks

Rein TEINBERG

On esitatud noorpullide ja -lehmade aretusväärtuse hindamiseks kasutatava põlvnemisindeksi arvutamise teoreetilised põhimõtted. On toodud valemid indeksi osaregressioonikordajate arvutamiseks ja arvutil saadud indeksivariandid. Aretusväärtuse prognoosi täpsus (korrelatsioon tegeliku aretusväärtusega) oli põlvnemisandmete põhjal (ema, emaisa, isa ja isaisa aretusväärtuse andmed) maksimaalselt 0,62. Kui indeksisse võeti ka lehma enda toodanguandmed (5 laktatsioonil), tõusis hindamistäpsus 0,78-ni. Kui noorlooma kohta on olemas ema jõudlusandmed ja isa järglaste (noorlooma poolõdede) hinnang, lisavad vanavanemate andmed hindamistäpsusele vähe. Jõudlusindeksit on soovitatud praktikas rakendada Eesti veisetõugude puhul.

ПРОГНОЗИРОВАНИЕ АДДИТИВНОЙ ГЕНЕТИЧЕСКОЙ ЦЕННОСТИ У МОЛОЧНОГО СКОТА

1. Индекс происхождения, полученный на основе данных о продуктивности предков

Рейн ТЕЙНБЕРГ

Рассматриваются теоретические принципы прогнозирования племенной ценности молодых быков и нетелей на основании индекса происхождения. Приведены формулы для вычисления весовых коэффициентов индекса и результаты имитации разных индексов на компютере. Максимальная точность прогнозирования племенной ценности с использованием только данных предков (мать, отец матери, отец и отец отца) составляла 0,62. Включение в этот индекс собственной продуктивности коровы (5 лактаций) увеличивает точность определения до 0,78. Если имеются данные по матери и проверки по потомству отца, то данные прародителей только незначительно повышают точность прогноза. Индекс рекомендуется для практического использования у эстонских пород скота.