Predicted breeding values for litter size with an animal model used in the Danish pig breeding program

Daniel Sorensen

STATENS HUSDYRBRUGSFORSØG

Foulum, Postboks 39, 8830 Tjele. Telf.: 86 65 25 00. Fax: 86 65 24 97

Statens Husdyrbrugsforsøg, oprettet 1883, er en institution under Landbrugsministeriet.

Statens Husdyrbrugsforsøg har til formål at gennemføre forskning og forsøg og opbygge viden af betydning for erhvervsmæssig husdyrbrug i Danmark og bidrage til en hurtig og sikker formidling af resultater til brugerne.

Der skal i forsknings- og forsøgsarbejdet lægges vægt på ressourceudnyttelse, miljø og dyrevelfærd samt husdyrprodukternes kvalitet og konkurrenceevne.

Abonnement på Statens Husdyrbrugsforsøgs Beretninger og Meddelelser kan tegnes ved direkte henvendelse til Statens Husdyrbrugsforsøg på ovenstående adresse.

Der er følgende afdelinger:

Dyrefysiologi og Biokemi	Forsøg med Fjerkræ og Kaniner
Forsøg med Kvæg og Får	Forsøg med Pelsdyr
Forsøg med Svin og Heste	Centrallaboratorium
Administration	Landbrugsdrift

NATIONAL INSTITUTE OF ANIMAL SCIENCE

Foulum, P.O. Box 39. DK-8830 Tjele
Tel: +45 86 65 25 00. Fax: 86 65 24 97

The National Institute of Animal Science was founded in 1883 and is a governmental research institute under the Ministry of Agriculture.

The aim of the institute is to carry out research and accumulate knowledge of importance to Danish animal husbandry and contribute to an efficient implementation of the results to the producers.

The research work puts emphasis on utilization of resources, environment and animal welfare and on the quality and competitiveness of the agricultural products.

For subscription to reports and other publications please apply direct to the above adress.

The institute consists of the below departments:

Animal Physiology and Biochemistry	Research in Poultry and Rabbits
Research in Cattle and Sheep	Research in Fur Animals
Research in Pigs and Horses	Central Laboratory
Administration	Farm Management & Services

Daniel Sorensen

Predicted breeding values for litter size with an animal model used in the Danish pig breeding program

Med dansk sammendrag

FORORD

Et væsentligt element i det danske svineavlssystem er en vel organiseret databank, der blev udviklet som et samarbejdsprojekt mellem Statens Husdyrbrugsforsøg og Danske Slagterier. Databanken muliggør anvendelse af moderne metoder til beregning af genetiske parametre samt gennemførelse af forskningsprojekter baseret på dataanalyse.

Nærværende beretning dokumenterer den model og den metode, der anvendes til beregning af avlsværdier for antal fødte grise pr kuld, sådan som disse er iværksat i det danske svineavlsprogram. Thorkild Vestergaard havde i en årrakke ihærdigt arbejdet med at nyttigøre forskningsresultater i avlsprogrammet. Hans entusiasme i de diskussioner, han deltog i tilknyttet avlsprogrammets udvikling, spillede en væsentlig rolle i nærværende arbejde.

Manuskriptet er renskrevet og forberedt til trykning af Aase Sørensen.

Statens Husdyrbrugsforsøg, Foulum
Juli 1991.

LIST OF CONTENTS

 ammendrag

I denne beretning dokumenteres den model, som anvendes i det danske svineavlsprogram til beregning af avlsværdier for antal fødte grise per kuld. Modellens anvendelse illustreres med et lille eksempel. Modellen betegnes som en enkelt-dyr-model med gentagne målinger (repeatability animal model), og parametrene estimeres ved hjælp af miksede modelmetoder.

I den oprindelige enkelt-dyr-model antages det, at avlsværdierne stammer fra en population, hvis gennemsnit er nul. Indenfor Landrace og Yorkshire er mange dyr blevet importeret fra lande, hvis gennemsnit for kuldstørrelse afviger indbyrdes, og dette medfører, at den ovennævnte antagelse ikke opfyldes. For at råde bod på dette inkluderes en gruppeeffekt i modellen. Den anvendte model er følgende:

$$Y_{ijklmno} = H_i + S_j + K_{kl} + \sum_m x_m G_{km} + a_{ijklmn} + P_{ijklmn} + e_{ijklmno}$$

hvor $Y_{ijklmno}$ repræsenterer antal fødte grise fra kuld o og so n. H_i, S_j og K_{kl} er systematiske effekter, der repræsenterer henholdsvis besætning x år x befrugtningsmetode (KS eller naturlig befrugtning), årstid og endelig race x kuldnummer. $\sum_m x_m G_{km}$ er gruppeeffektens bidrag til antal fødte grise indenfor race k ($\sum_m x_m = 1$). a_{ijklmn}, P_{ijklmn} og e_{ijklmn} er tilfældige effekter, der repræsenterer henholdsvis avlsværdier, permanente miljøeffekter og endelig resteffekten fra kuld o. Arveligheden og gentagelseskoefficienten antages at være henholdsvis 0,10 og 0,15. Inversen af slægtsskabskorrelationsmatricen er beregnet efter en algoritme afledt af Henderson (1976).

I beretningen redegøres for nogle matematiske og EDB-mæssige overvejelser tilknyttet opbygningen og løsningen af den miksede models ligningssystem.

En række antagelser, som ligger til grund for estimation af avlsværdier, er diskuteret. En antagelse, som muligvis ikke er opfyldt, er, at antal fødte grise kun påvirkes af gener, som har en additiv virkning. Dette medfører, at gruppeeffekter kombineres på en additiv måde. Det er hensigten, at heteroseeffekter vil blive indarbejdet i proceduren i fremtiden.

Summary

A model for selection for litter size used in the Danish pig breeding progra is described, and an example of its use is given. The model is based on repeatability reduced animal model with groups and the parameters are estimate using mixed model methodology.

The usual assumption of the animal model is that breeding values are sample from a population with zero mean. Within the Landrace and Yorkshire breeds, considerable amount of importation has taken place over the years and it is clea that there are rather large differences in litter size among the countries o origin. The purpose of introducing groups in the model is to take account of thi fact. The model which is used is:

$$Y_{ijklmno} = H_i + S_j + K_{kl} + \sum_m x_m G_{km} + a_{ijklmn} + p_{ijklmn} + e_{ijklmno}$$

where, $Y_{ijklmno}$ is the oth record (number of born piglets) of sow n, H_i, S_j, K_k are fixed effects of herd x year x type of fertilisation, season and breed : parity number, respectively, $\sum_m x_m G_{km}$ is the fixed group contribution nested within breed k ($\sum_m x_m = 1$) and a_{ijklmn}, p_{ijklmn}, and e_{ijklmn} are breeding values, permanent environmental effects, and residuals, respectively (random effects). The heritability and repeatability used as priors are 0.10 and 0.15, respectively. The inverse of the additive genetic relationship matrix is computed using an algorithm from Henderson (1976).

Some mathematical aspects associated with the development of the model are discussed, as well as several computing details involved with the building and solution of the mixed model equations.

A number of assumptions of the model are described. One of the most contentious assumptions is that the trait is affected by genes that act additively, within and between loci. Group effects combine therefore in an additive manner. This assumption is hardly tenable in a trait like litter size, and it is hoped that heterosis effect will be incorporated in the model in the future.

1. Introduction

It is well established that increased litter size improves the economic efficiency of pig production. This would be further enhanced if production traits such as leanness were to be approaching an optimum for a given slaughter weight, so that further genetic improvement as carcass fat declines becomes difficult to achieve (Hill and Webb, 1982).

Reasons why litter size has not been included in the breeding goal of the Danish pig breeding program have been, firstly, that experimental evidence showing that it can be enhanced by selection has been lacking. Secondly, it has been argued that genetic improvement of litter size using traditional selection techniques would lead to very slow progress, essentially, due to the low heritability of the trait and its sex limited expression (Smith, 1964).

Recently however, some encouraging results have become available. Le Roy et al. (1987) presented the results of a hyperprolific sow experiment carried out under farm conditions, where the estimates of realised heritability for total number of born piglets and total number of liveborn piglets were 0.14 ± 0.05 and 0.10 ± 0.05, respectively. Also, Avalos and Smith (1987) have shown that use of a family index incorporating several sources of information can lead to expected annual responses of up to 0.51 pigs per litter.

With these results as a background, in June 1988 the National Institute of Animal Science made available to the Danish pig breeding program, predicted breeding values for total number of born piglets of breeding animals. The trait was used partly because overall reproductive efficiency can be enhanced most effectively by increasing litter size (Smith et al, 1983) and partly, because this is the most reliable data available on reproduction traits in the Danish databank for pig production data.

The predicted breeding values were derived using mixed model techniques, and the model used is known as the repeatability animal model with groups. This evaluation procedure takes into account the heritability and repeatability of the trait, and makes use of all the records in different parities of the animal itself, and of all the records from the animal's relatives present in the data set. There can be more than one thousand records available from an animal's relatives, and all the information is combined in an optimum manner in order to evaluate its breeding value. Both males and females receive a predicted breeding value. These predicted breeding values are simultaneously corrected for such

effects as herd-year, season, parity number, breed and type of conception
(artificial insemination versus natural mating). Account is also taken of the fact
that some animals are imported from countries whose means differ. For example, the
population of Finnish Landrace sows has a mean litter size that differs from th
population of Danish Landrace sows. Rather than assuming that all breeding value
are sampled from one population, the group effect in the model reflects that th
sampling process involves different populations, and even mixture of populations,
whose means for the trait in question differ.

The first years are considered an exploratory period during which alternativ
models will be fitted to the data and the sources of variation contributing t
number of born piglets will be investigated. During this initial stage, the inde
will not be part of the breeding goal.

The purpose of this publication is to describe and document the methods
followed in the development of these predicted breeding values.

2. The individual animal model with groups

The predicted breeding values for total number of born piglets are obtained
using mixed model techniques on a repeatability reduced animal model with groups.
The computer program was written in PL/I and it can accommodate an arbitrary
number of fixed and random effects.

In the individual animal model excluding groups, the phenotypic value is
written in terms of fixed effects and the contribution from the animal's additive
genetic and non-genetic values. For example, assuming an additive genetic model,
for a given animal, j, say, that makes a record (y_{ij}) in herd (H) i, we write:

$$y_{ij} = H_i + a_{ij} + e_{ij}$$

where a_{ij} is the additive genetic value and e_{ij} is the residual non-genetic
effect. The usual assumption in this model with respect to the breeding values is
that they have null means.

In order to take into account the fact that animals often originate from
populations whose means differ, a group effect is included in the model. This idea

was introduced in the context of a sire model by Thompson (1979) and extended to an individual animal model by Robinson (1986), Westell and Van Vleck (1987) and Quaas (1988). The concept is simple and intuitively appealing and it arises from the fact that an animal receives half of its genes from one parent and the other half from the other. If animals 1 and 2, which are sampled from populations with means g_1 and g_2, respectively, are parents of animal 3, then under the group model, the expected breeding values of 1, 2, and 3 are g_1, g_2, and $1/2$ $(g_1 + g_2)$, respectively.

A matrix representation of the individual animal model with groups is:

$$y = Xb + Gg + Za + e \tag{1}$$

where y is the vector of observations, b is the vector of fixed effects excluding groups, g is the vector of group effects, and a and e are random vectors. The matrices X, G and Z are known design matrices. The matrix G has numbers of columns equal to the number of groups, and each row sums to one. The elements of a row of G are the proportion with which a group effect contributes to the given individual. They can also be interpreted as the proportion of genes originating from the different groups.

The first and second moments of (1) are:

$$E \begin{bmatrix} y \\ a \\ e \end{bmatrix} = \begin{bmatrix} Xb + Gg \\ 0 \\ 0 \end{bmatrix}$$

and

$$Var \begin{bmatrix} y \\ a \\ e \end{bmatrix} = \begin{bmatrix} ZAZ'\sigma^2_a + R & ZA\sigma^2_a & R \\ (ZA)'\sigma^2_a & A\sigma^2_a & 0 \\ R' & 0 & R \end{bmatrix} \tag{2}$$

In (2), A is the numerator relationship matrix and σ^2_a is the additive genetic variance in the populations from which base animals are sampled. R is usually assumed to be equal to $I \sigma^2_e$, where I is the identity matrix and σ^2_e is variance of the residual effect. Thus, the model assumes that groups have common variance. The important concept in the group model is that breeding values are given by:

$$a* = Qg + a \tag{3}$$

The assumption in (3) is that $E(a*) = \mathbf{Qg}$,

where $\mathbf{ZQ} = \mathbf{G}$.

Like any other fixed effect, groups may not be estimable. While predicted genetic values are always estimable, expressing them in the form of (3) may render them non-estimable. Differences among predicted breeding values expressed as in (3) can be estimable if group effects are not nested within other fixed effects. If group effects are nested within breeds, for example, then estimable functions of differences among breeding values, expressed as in (3), can be made within breeds only.

The total number of equations that have to be solved to obtain the predicted genetic values is equal to the number of fixed effects plus the total number of animals in the system. With large data sets, this can be a very large number and computing costs can be substantial.

3. The reduced animal model with groups

An alternative formulation is to make use of Mendelian theory, and to reparameterise the model, so that the total number of breeding values is reduced from the total number of animals, to the number of animals that have offspring. This is known as the reduced animal model, originally proposed by Quaas and Pollak (1980).

Let a_s (a_d) be the additive genetic effect of the sire (dam). Then, ignoring groups for the moment, the additive genetic effect of an individual, a_o, for the case of both parents known, one parent known, the sire say, and none of the parents known is written in (4), (5) and (6):

$$a_o = 1/2 \ a_s + 1/2 \ a_d + m_2 \tag{4}$$
$$a_o = 1/2 \ a_s + m_1 \tag{5}$$
$$a_o = m_o \tag{6}$$

In (4), m_2 is the Mendelian sampling deviation and its variance is equal to:

$$Var(m_2) = 1/2 \ \sigma^2_a \ (1 - 1/2(F_s + F_d)) \tag{7}$$

(Foulley and Chevalet (1981)) where σ^2_a is the additive genetic variance of the population from which the breeding values were sampled and F_s (F_d) is the inbreeding coefficient of the sire (dam), such that $Var(a_s) = \sigma^2_a(1 + F_s)$. The variance of m_1, $Var(m_1)$ can be shown to equal:

$$Var(m_1) = 3/4 \ \sigma^2_a(1 - 1/3 \ F_s) \tag{8}$$

where F_s is the inbreeding coefficient of the only known parent, the sire.

The variance of m_o, $Var(m_o)$ is equal to

$$Var(m_o) = \sigma^2_a \tag{9}$$

The Mendelian sampling deviation, m_2 in (1) can also be viewed as the sum of the deviations of each parental gamete, with respective contributions of $1/4 \ \sigma^2_a(1 - F_s)$ and $1/4 \ \sigma^2_a \ (1 - F_d)$ to $Var \ (m_2)$.

In the reduced animal model, animals that have no progeny are called non-parent records, and are written in terms of equations (4), (5), or (6), depending on the number of known parents, and this approach leads to having to solve directly, only the breeding values of animals that have progeny, which are called parent records. This is particularly beneficial in a species like pigs, where most animals contribute as non-parents and therefore savings in computing costs can be substantial.

The introduction of groups in the reduced animal model leads to changes in (5) and (6). With one parent known, (5) becomes:

$$a^*_o = 1/2 \ a^*_s + 1/2 \ g_d + m_1 \tag{10}$$

With both parents unidentified, (6) becomes:

$$a^*_o = 1/2 \ g_s + 1/2 \ g_d + m_o \tag{11}$$

If both parents are identified, the effect of groups is included in the parental contribution, and thus (4) still holds, provided that the breeding values are interpreted as in (3). Absence of pedigree information on sire or dam must be supplemented with information on the group which the missing parent originates from.

8

The repeatability reduced animal model is written in matrix notation as follows:

$$
\begin{bmatrix} y_f \\ y_o \end{bmatrix} = Xb + \begin{bmatrix} G_f \\ G_o \end{bmatrix} g + \begin{bmatrix} Z_f \\ T \end{bmatrix} a_f + Mm + Wp + e \tag{12}
$$

where,

y_f is the vector of parental records;

y_o is the vector of non-parent records;

b is the vector of all fixed effects (for parental and nonparental records);

a_f is the vector of random parental breeding values;

g is the vector of fixed group effects;

m is the vector of random Mendelian sampling deviations;

p is the vector of random permanent environmental effects;

e is the vector of random pure environmental effects.

X, G_f, G_o, Z_f, M and W are known design matrices and T is a matrix whose number of rows is equal to the number of rows in y_o, and number of columns equal to the number of elements in a_f, with row elements equal to 1/2 or 0, relating non-parental to parental breeding values. The vector of non-parental breeding values (a^*_o) is from (3) and (12):

$$
a^*_o = T_o\, a^*_f + (Q_o - T_o\, Q_f)\, g + m \tag{13}
$$

In (13), we notice that $a^*_f = Q_f\, g + a_f$, and that $Q_o - T_oQ_f = 0$ if the non-parent has both parents idenfified. The matrices Q_f and Q_o result from partitioning Q in (3) such that parent records precede non-parents. Q_o is of order (number of non-parental breeding values x number of groups), Q_f has order (number of parental breeding values x number of groups) and T_o is a subset of T which associates non-parental with parental breeding values and has order (number of non-parental breeding values x number of parental breeding values).
Thus:

$$
\begin{bmatrix} a^*_f \\ a^*_o \end{bmatrix} = \begin{bmatrix} Q_f \\ Q_o \end{bmatrix} g + \begin{bmatrix} a_f \\ a_o \end{bmatrix}
$$

The expected value of $(y_f\ y_o)$ in (12) is assumed to be equal to

$\mathbf{Xb} + \begin{bmatrix} \mathbf{G_f} \\ \mathbf{G_o} \end{bmatrix} \mathbf{g}$, and the second moments of the model are:

$$\text{Var}(\mathbf{a_f}) = \mathbf{A_f} \, \sigma^2_a \tag{14}$$

$$\text{Var}(\mathbf{m}) = \mathbf{D} \, \sigma^2_a \tag{15}$$

$$\text{Var}(\mathbf{p}) = \mathbf{I} \, \sigma^2_p \tag{16}$$

$$\text{Var}(\mathbf{e}) = \mathbf{I} \, \sigma^2_e \tag{17}$$

where,

$\mathbf{A_f}$ is the additive genetic relationship matrix among parental breeding values;

$\mathbf{D}$ is a diagonal matrix whose elements are given by (7), (8) or (9) with σ^2_a deleted;

$\mathbf{I}$ is the identity matrix. σ^2_a is the additive genetic variance of the population from which the base animals were conceptually sampled. σ^2_p and σ^2_e are the variance due to permanent environmental effects and due to pure environmental effects, respectively. All covariances between $\mathbf{a_f}$, $\mathbf{m}$, $\mathbf{p}$ and $\mathbf{e}$ are zero. The heritability (h^2) and repeatability (r) are defined as:

$$h^2 = \sigma^2_a / (\sigma^2_a + \sigma^2_p + \sigma^2_e) \tag{18}$$

$$r = (\sigma^2_a + \sigma^2_p) / (\sigma^2_a + \sigma^2_p + \sigma^2_e) \tag{19}$$

Letting $\mathbf{Z}' = (\mathbf{Z_f} \ \mathbf{T})'$ and $\mathbf{G}' = (\mathbf{G_f} \ \mathbf{G_o})'$, then the mixed model equations corresponding to the model are:

$$\begin{bmatrix} \mathbf{X'X} & \mathbf{X'G} & \mathbf{X'Z} & \mathbf{X'M} & \mathbf{X'W} \\ \mathbf{G'X} & \mathbf{G'G} & \mathbf{G'Z} & \mathbf{G'M} & \mathbf{G'W} \\ \mathbf{Z'X} & \mathbf{Z'G} & \mathbf{Z'Z}+\mathbf{A_f}^{-1}k_1 & \mathbf{Z'M} & \mathbf{Z'W} \\ \mathbf{M'X} & \mathbf{M'G} & \mathbf{M'Z} & \mathbf{M'M}+\mathbf{D}^{-1}k_2 & \mathbf{M'W} \\ \mathbf{W'X} & \mathbf{W'G} & \mathbf{W'Z} & \mathbf{W'M} & \mathbf{W'W}+\mathbf{I}k_3 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{g}} \\ \hat{\mathbf{a}}_f \\ \hat{\mathbf{m}} \\ \hat{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'y} \\ \mathbf{G'y} \\ \mathbf{Z'y} \\ \mathbf{M'y} \\ \mathbf{W'y} \end{bmatrix} \tag{20}$$

where, $\mathbf{A}^{-1}_f$ is the inverse of the additive genetic relationship matrix among parental breeding values,

$$k_1 = k_2 = \sigma^2 e / \sigma^2 a = (1-r)/h^2 \tag{21}$$

$$k_3 = \sigma^2 e / \sigma^2 p = (1-r)/(r-h^2) \tag{22}$$

Under the conditions of the model described above, the system in (20) provide best linear unbiased estimators (BLUE) of estimable functions of fixed effects an best linear unbiased predictors (BLUP) of random effects (Henderson, 1973).

In (20) the matrix G is dense and therefore all the blocks containing G will b costly to build. The resulting dense coefficient matrix is also computationall demanding to solve iteratively. Secondly, the predicted breeding values as define in the group model are $a^*_f = Q_f g_f + a_f$, and these are not directly obtained fro (20). Quaas and Pollak (1981) have suggested a transformation that leads to a computationally less demanding system of equations.

Define a matrix T^*, such that:

$$T^* = \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & Q_f & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix}$$

Then, inserting $T^{*-1}T^*$ between the coefficient matrix and the solution vector in (20), and premultiplying both sides by $(T^{*-1})'$, one obtains:

$$\begin{bmatrix} X'X & X'L & X'Z & X'M & X'W \\ L'X & L'L+Q'_f A^{-1}_f Q_f k_1 & L'Z-Q'_f A^{-1}_f k_1 & L'M & L'W \\ Z'X & Z'L-A^{-1}_f Q_f k_1 & Z'Z+A^{-1}_f k_1 & Z'M & Z'W \\ M'X & M'L & M'Z & M'M+D^{-1}k_2 & M'W \\ W'X & W'L & W'Z & W'M & W'W+Ik_3 \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{g} \\ \hat{a}^*_f \\ \hat{m} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} X'y \\ L'y \\ Z'y \\ M'y \\ W'y \end{bmatrix} \quad (23)$$

where $L = G - ZQ_f = \begin{bmatrix} G_f \\ G_o \end{bmatrix} - \begin{bmatrix} G_f \\ TQ_f \end{bmatrix} = \begin{bmatrix} 0 \\ G_o - TQ_f \end{bmatrix}$

The matrix L is equal to 0 when non-parent records have both parents identified and thus all blocks in (23) containing L become 0.

From (13), solution to non-parent breeding values are obtained from (24):

$$\hat{a}^*_o = T\hat{a}^*_f + (G_o - TQ_f) \hat{g} + \hat{m} \quad (24)$$

A remarkable property of the system in (23) is that Q_f does not have to be built up explicitly. Quaas (1988) shows how the structure of Q_f and A^{-1} can be

exploited so that the blocks containing Q_f can be constructed as records are read in, following a given set of rules.

The predicted value of individuals' future records is given by $\hat{a}* + \hat{p}$, where $\hat{a}*' = (\hat{a}*_f \ \hat{a}*_o)'$.

Inspection of the coefficient matrix of (23) shows that $M'M + D^{-1}k_2$ and $W'W + Ik_3$ are diagonal matrices. Therefore, to reduce the size of the coefficient matrix, the equations belonging to $\hat{p}$ and $\hat{m}$ can be absorbed. This absorption process leads to the system of equations (25):

$$
\begin{bmatrix}
X'PFX & X'PFL & X'PFZ \\
L'PFX & L'PFL + Q'_f A^{-1}_f Q_f k_1 & L'PFZ - Q'_f A^{-1}_f k_1 \\
Z'PFX & Z'PFL - A^{-1}_f Q_f k_1 & Z'PFZ + A^{-1}_f k_1
\end{bmatrix}
\begin{bmatrix}
\hat{b} \\
\hat{g} \\
\hat{a}*_f
\end{bmatrix}
=
\begin{bmatrix}
X'PFy \\
L'PFy \\
Z'PFy
\end{bmatrix}
\tag{25}
$$

where,

$$P = I - W(W'W + Ik_3)^{-1} W'$$
$$F = I - M(M'PM + Ik_2)^{-1} M'P$$

Solutions to non-parental breeding values and to permanent environmental effects are obtained using back-solving techniques. To solve for $\hat{m}$ and then for $\hat{a}*_o$ using (24), we proceed as follows. Absorbing the equation for $\hat{p}$ in (23) and solving for $\hat{m}$ yields:

$$\hat{m} = (M'PM + D^{-1}k_2)^{-1} \ (M'Py - M'PX\hat{b} - M'PL\hat{g} - M'PZ\hat{a}*_f) \tag{26}$$

It can be shown by expansion of (26) that the ith element of $\hat{m}$ is given by:

$$
\hat{m}_i = h_i \ c_i \ (\sum_{j=1}^{n_i} Y_{ij} - \sum_{j=1}^{n_i} x_{ij} \ \hat{b} - n_i (v_s 1/2 \ \hat{g}_m + v_d 1/2 \ \hat{g}_n) -
$$
$$
- n_i ((1-v_s) \ 1/2 \ \hat{a}*_s + (1-v_d) \ 1/2 \ \hat{a}*_d)) \tag{27}
$$

where h_i = 1 if individual i has one or more records,

= 0 if it has no record;

$c_i = k_3/(n_i k_3 + d_i k_2 \ (n_i + k_3))$

In (27),

$k_2 = (1 - r)/h^2$;

$k_3 = (1 - r)/(r - h^2)$;

n_i = number of records of individual i;

d_i = inverse of the ith element of D, where the elements of D are given in (7), (8), and (9) with σ^2_a deleted;

Y_{ij} = the jth record of individual i;

x_{ij} = row of matrix X corresponding to jth record of individual i;

$\hat{b}$ = solution vector of fixed effects in (25);

$v_s(v_d)$ = 1 if sire (dam) of i not known or

= 0 if sire (dam) of i is known;

$\hat{g}_m(\hat{g}_n)$ = group effect of parent m (n);

$\hat{a}^*_s(\hat{a}^*_d)$ = predicted breeding value of sire (dam) of i.

If in (23) $\hat{p}$ is not absorbed, then the solution for $\hat{m}_i$ is as in (27), but with $c_i = (n_i + d_i k_2)^{-1}$ and with $-n_i \hat{p}_i$ included in the term in brackets.

With $\hat{m}$ available, solution to non-parental breeding values are readily obtained from (24). For example, for the ith non-parent with parental predicted breeding values $\hat{a}^*_s$ and $\hat{a}^*_d$,

$$\hat{a}^*_i = (1-v_s)\ 1/2\ \hat{a}^*_s\ + (1-v_d)\ 1/2\ \hat{a}^*_d + v_s\ 1/2\ \hat{g}_n + v_d\ 1/2\ \hat{g}_m + \hat{m}_i \qquad (28)$$

When the non-parent does not have a record, the estimate of its Mendelian sampling deviation is zero and its predicted breeding value is obtained from (28) with $\hat{m}_i$ deleted.

The solution for the permanent environmental effects is obtained directly from (23). The last equation in (23) is:

$$\hat{p} = (W'W + Ik_3)^{-1}\ (W'y - W'X\hat{b} - W'L\hat{g} - W'Z\hat{a}^*_f - W'M\hat{m}) \qquad (29)$$

By expansion of (29), it can be shown that the ith element of $\hat{p}$ is:

$$\hat{p}_i = w_i\ h_i(\sum_{j=1}^{n_i} Y_{ij} - \sum_{j=1}^{n_i} x_{ij}\ \hat{b} - n_i\ \hat{a}^*_i) \qquad (30)$$

where $w_i = 1/(n_i + k_3)$ and all other terms are identified in connection with (27).

If i is a non-parent, $\hat{a}*_i$ is obtained from (28).

If an individual does not have a record, the estimate of its predicted permanent environmental effect is zero.

From a computational point of view, it should be stressed that the system of equations (25) can be built directly, without having to actually carry out the absorption of m. This is achieved by reparamenterising model (12) as follows:

$$\begin{bmatrix} y_f \\ y_o \end{bmatrix} = \begin{bmatrix} X_f \\ X_o \end{bmatrix} b + \begin{bmatrix} G_f \\ G_o \end{bmatrix} g + \begin{bmatrix} Z_f \\ T \end{bmatrix} a_f + \begin{bmatrix} W_f \\ W_o \end{bmatrix} p + \begin{bmatrix} e_f \\ m + e_o \end{bmatrix} \tag{31}$$

$$\text{where Var} \begin{bmatrix} e_f \\ m + e_o \end{bmatrix} = \begin{bmatrix} I_f & 0 \\ 0 & D\sigma^2 a/\sigma^2 e + I_o \end{bmatrix} \sigma^2 e \tag{32}$$

Writing the mixed model equations for (31) and (32), applying the transformation suggested by Quaas and Pollak (1981) and absorbing $\hat{p}$ leads to (25).

4. The model for prediction of breeding values for number of born piglets.

The model that had been in operation during 1988/89 is:

$$Y_{ijlno} = H_i + S_j + K_l + a_{ijln} + p_{ijln} + e_{ijlno} \tag{33}$$

where,

Y_{ijlno} : the oth record (number of born piglets) of sow n from breed x parity l, farrowing in season j, belonging to herd x year x type of fertilisaton i;

H_i: fixed effect of herd x year x type of fertilisation i (type of fertilisation: artificial insemination or natural mating);

S_j: fixed effect of season j (4 seasons per year);

K_l: fixed effect of breed x parity number l;

a_{ijln}: random effect of breeding value of animal n;

p_{ijln}: random effect of permanent environmental effect of animal n;

e_{ijlno}: random residual effect of record o.

Higher interaction terms that are likely to be biologically appropriate, like HS_{ij}, could not be fitted because the number of observations in a large proportion of this herd x year x type of fertilisation x season interaction becomes too small. Similar considerations precluded inclusion of breed in this interaction.

During 1990 this model has been extended to include groups. This implies that (33) is changed to (34):

$$Y_{ijklmno} = H_i + S_j + K_{kl} + \sum_m x_m G_{km} + a_{ijklmn} + p_{ijklmn} + e_{ijklmno} \qquad (34)$$

where the parameters are defined in connection with (33). The parameter K_{kl} has explicity two subscripts where k represents breed and 1 represents parity number, and G_{km}, the mth group effect is nested within breed k. The known constant x_m, is the contribution of the mth group, such that $\sum x_m = 1$.

The total genetic value, is given by:

$$a^*_{ijklmn} = \sum_m x_m G_{km} + a_{ijklmn} \qquad (35)$$

For non-parent records, a_{ijkln} in (34) is substituted by $(a_s + a_d)/2 + m_{ijklmn}$, where a_s (a_d) is the breeding value of the sire (dam) of n, and the term in m is the Mendelian sampling effect. Collecting the first three terms in (34) in a vector b, the fourth in g and so on, (34) is written in matrix notation as shown in (12) with first moment equal to $\mathbf{Xb} + \mathbf{Gg}$ and second moments given by (14), (15), (16), and (17). The heritability and repeatability of total number of liveborn piglets are for the time being assumed to be 0.10 and 0.15, respectively. It is assumed also that these figures hold with all possible combination of fixed effects in the model.

The inverse additive genetic relationship matrix among all animals in the system is computed following an algorithm by Henderson (1976). The program can easily accommodate Quaas´ (1976) algorithm to allow for inbreeding, but at higher computing cost.

5. Assumptions of the model

Two sets of assumptions will be discussed. One set is rather general and is associated with the mixed model equations (20) and (25). The other set is

particular to the model described in section 4. These two sets will be discussed in turn.

The starting point in (1) is that the vector of observations y is a function of a set of variables that act additively. The variables in the vectors b and g are fixed effects, while those in the vectors a and e are random effects. That is, a and e are each considered as one sample from a population of vectors a and e which has been drawn into the sample space associated with the data vector y. The random vectors a and e determine that the data vector y is also regarded as a random variable sampled from a conceptual population. In repeated sampling, the vectors a and e have mean zero and the vector y has a mean of $Xb + Gg$ and a variance given by the first row and column of (2).

The sampling of a and e is assumed to be carried out at random - selection as it operates in a breeding program is not allowed for.

The vector g has parameters representing group effects. Knowledge of an animal's ancestors defines the relative contribution of groups to its performance. Pedigrees must be complete and the parents of those animals that are not identified must have an identified group of origin. These unidentified parents are assumed to be average representatives (i.e. unselected) from their groups. Groups combine additively on their effect on performance.

Model (12) does not introduce new concepts. Here the stochastic variables are a_f, m, p, and e. The derivation of (20) from (12) assumes that a_f, m, p, and e have null mean and that their covariance matrix is known at least to proportionality. In our context, this means that the heritability and repeatability are known without error.

In going from (12) to (25), no assumption is made about the form of the distribution of the stochastic elements in the model. It is assumed though that the conceptual populations they are sampled from have null means. This is acceptable for m, p, and e. The vector m represents deviations due to Mendelian sampling. If there is no selection acting at the gametic stage, before or after fertilisation, then these deviations from the parental average should add up to zero.

The assumption that p and e are vectors sampled from populations with zer
expectation is not problematic. Permanent environmental effects and residua
effects are not likely to affect the records in any particular direction, o
average, especially when these effects are unspecified.

Other set of assumptions implied in (12), (14), (15), (16), and (17) are that
variances of random effects are constant across all levels of fixed effects that a
strictly additive genetic model holds and that the correlation between records o
the same animal is a constant (the repeatability) irrespective of whether th
records are adjacent or further apart.

Under these assumptions, given that the model is correct, solution of th
system (25) yields BLUE (best linear unbiased estimators) of estimable functions
of fixed effects and BLUP (best linear unbiased predictors) of random effects. If
the variances of the random effects, or their ratios, are not known and one
substitutes estimates of them, then the resulting solution to the random effects
are not BLUP, but are still unbiased (Kackar and Harville, 1981).

The assumption that the vector of genetic values, a_f is a random sample from a
population with a certain mean is not tenable when selection is known to have
operated on a trait whose heritability is larger than zero. One wishes to know
though, if there exists a set of conditions that if satisfied, lead to the result
that the system (25) leads to predictors of genetic merit with good properties,
even though the data have been generated by selection. This is a difficult problem
which has not been totally solved yet. Henderson (1975) invoking early results by
Pearson (1903) have shown that these set of conditions exist and that these are:

a) the vectors of random variables in the model follow a multivariate normal
   distribution;

b) the variances of the random variables, or their ratios, are known:

c) selection does not operate across levels of fixed effects;

d) selection does not operate on traits correlated with the data vector not
   specified in the model.

The genetic model implied by (a) is one in which the metric trait is determined

by a very large number of additive (nonepistatic) loci. An important assumption of this model is that, with directional selection, gene frequency changes are infinitesimally small and changes in the additive genetic variance are due to joint disequilibrium (Bulmer, 1971) and in small populations, due to inbreeding. With finite number of loci, particularly with genes of large effects, the assumption of normality does not hold and with data spanning several generations, gene frequency changes may be substantial. Under these conditions, the predictors obtained from (25) are likely to be biased.

The condition specified in (b) requires the correct specification of the covariance matrix of the random effects in the model. In the context of an animal model, this means that the relationship matrix among all the individuals in the data is complete and that the heritability in the population from which the base animals were conceptually sampled is known. If the heritability is not known, simulation results indicate that use of a REML (restricted maximum likelihood) estimator under an animal model in the solution of (25) leads to predictors of random effects with no detectable bias (Sorensen and Kennedy, 1986).

Condition (c) implies that the data had been adequately corrected for fixed effects prior to selection or that selection operated within fixed effects. The introduction of groups in the model guarantees in fact that this condition is violated. In attempting to solve one problem - that is, acknowledging that animals originate from different populations - one introduces a new one: when breeding values are the selection criterion and these include a group effect, selection operates across fixed effects.

Condition (d) requires that observed selection differentials were not the result of selection of a trait not specified in the model, and correlated with the vector **y**. Selection for daily gain within litters prior to sending pigs to a test station is an example of a violation of (d).

Gianola et al. (1988) has recently questioned the validity of some of the assumptions of the Pearson model in the context of genetic selection. Using a Bayesian approach, they arrive at a set of conditions which are not all in agreement with Henderson's. Some of these points were further expanded by Im et al. (1989).

The second set of assumptions relates more specifically to the model proposed for genetic evaluation of number of liveborn piglets in (34).

The most important of these assumptions are:

a. the trait (number of liveborn piglets) is not influenced by the sire the sow is mated to.

It is fairly well established that litter size at birth is influenced by the litter's genotype and as such, the assumption of no paternal influence does not strictly hold. However, published analysis indicate that the proportion of variation in litter size due to sire of litter is very small (Christensen, 1978; Van der Steen and Kock, 1987).

b. an additive genetic model holds. This implies that the trait does not exhibit inbreeding depression or heterosis.

There is ample data showing that litter size traits exhibit heterosis and inbreeding depression (Hill and Webb, 1982). These results are not compatible with a simple additive genetic model, but rather dominance and/or various forms of epistasis must be invoked. A unified mixed model approach which takes account of the effects of inbreeding and dominance on the mean and variance is not yet developed, although work in this area has started (Mäki-Tanila and Kennedy, 1986; Smith and Mäki-Tanila, 1990).

c. the genetic correlation of the trait in different parities is 1.

Vangen (1986) shows results that indicate that this hypothesis may not hold, and that the genetic correlations decrease as the distance between parities increases. On the other hand, in a recent review, Haley et al. (1988) conclude that genetic correlations between adjacent parities are high, although less than one, but the estimates that they quote imply that the genetic correlation between parities 1 and 4 is considerably less than one. They argue that the available estimates may be biased by selection. Until a multiple trait restricted maximum likelihood estimator of variance and covariance components, with an animal model, becomes available, the estimates must be interpreted with reservation. Even then, one could argue that estimates are biased by exclusion of non-additive gene action from the model. In view of the considerable computational difficulties associated with a multitrait evaluation system, one can take refuge in some of the conclusions of Haley et al. (1988) and justify the univariate approach as a temporary operational compromise.

d. the heritability and repeatability are known and equal to 10% and 15%, respectively, and these values hold for all breed x parity combinations and for all groups.

Associating these sets of assumptions with what is known of the biology of the trait and with what goes on in a typical breeding program may help towards understanding the discrepancy often observed between observed and expected selection responses!

6. Possible developments of the model

Progress in the areas of biology, methods of estimation of genetic parameters and in computer science makes it possible to develop a more refined model for genetic evaluation of number of liveborn piglets. Recently, a REML (restricted maximum likelihood) algorithm using animal models has become available (Meyer, 1989). This allows to test whether (a) and (d) hold. REML under an animal model provides estimates of genetic parameters with well defined statistical properties, especially when used in selected data given that the correct model can be specified. Statistical analysis could confirm whether genetic parameters are constant across breeds, and yield estimates that could be used as priors in the solution of the mixed model equations. Use of REML estimates obtained from the databank and used as variance ratios, would be a significant improvement over the present approach, where the heritability and repeatability used as priors are obtained from the literature.

The performance of crosses among animals originating from different groups (countries) should be analysed for presence of heterosis. Should this be confirmed, the present model can be extended - at least from an operational point of view - to allow for non-additivity. This can be accomodated including the expected proportion of heterosis in the model as a covariate.

7. An example

This example will illustrate the computation of breeding values and how some of the expressions developed in section 3 are used.

Consider the records of 14 animals shown in Table 1. There are two fixed effects, A (with 2 levels) and B (with 3 levels), and animals 9, 10, and 13 have repeated records.

Animals 1, 2, 3, 4, 11, and 14 have both parents missing; 7, 8, and 10 have one parent missing. Even though these parents are unknown, one must assume that the originate from a given group. The groups of origin of the missing parents ar shown in table 2. In this table, the sign " - " means that the parent is no known.

Individuals 1, 2, 3, 4, 5, 6, and 11 are parents; the remaining individuals d not have offspring and are therefore nonparents.

Assume the model for parent records is:

$$Y_{ijklm} = A_i + B_j + \sum_k x_k G_k + a_{kl} + p_{kl} + e_{ijklm} \qquad (35)$$

and for non-parent records, a_{ijkl} is replaced by $1/2\ a^s + 1/2\ a^d + m_{ijkl}$, where, A and B are fixed effects, G are fixed group effects, a represents genetic values (random) and e is the residual peculiar to record m. For non-parents, their genetic value is written in terms of their parental genetic values (a^s and a^d) and the Mendelian term, m.

By collecting fixed effects A and B in a vector b, group effects in a vector g, the genetic values of parent records in a vector a_f, Mendelian terms in m and permanent environmental effects in p, and ordering the data such that parent records precede non-parents, one can express (35) in the matrix formulation (12). Matrices T, T_o, G_f, and G_o are:

$$T=\begin{bmatrix} 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} ; \quad T_o=\begin{bmatrix} 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$G'_f = \begin{bmatrix} 0 & 0 & 1 & 0 & 3/4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1/4 & 0 & 0 \end{bmatrix}$$

$$G'_o = \begin{bmatrix} 1 & 3/8 & 1/2 & 1/2 & 1/2 & 1/2 & 1/2 & 1/2 & 5/8 & 5/8 & 1/2 \\ 0 & 5/8 & 1/2 & 1/2 & 1/2 & 1/2 & 1/2 & 1/2 & 3/8 & 3/8 & 1/2 \end{bmatrix}$$

The second row of T corresponds to animal 8. The $1/2$ in column 5 indicates that individual 5 is 8's parent. The second column of G'_o gives the group contribution to 8. Individual 8's unknown parent originates from group 2 – this contributes with $1/2\ G_2$. The known parent, 5 is the offspring of 1 and 2. Individual 1 contributes $1/4$ to G_1, while 2 contributes with $1/8$ to G_1 and with $1/8$ to G_2. This results in individual 8's group composition of $3/8\ G_1$ and $5/8\ G_2$. The total genetic value of 8 is thus:

$$a*_8 = 3/8\ G_1 + 5/8\ G_2 + 1/2\ a_5 + m_8.$$

The variance of $a*_8$ is $1/4\ \sigma^2 a + 3/4\ \sigma^2 a = \sigma^2 a$. The elements of the diagonal matrix D are: $\text{diag } D = (3/4\ \ 3/4\ \ 1/2\ \ 3/4\ \ 1/2\ \ 1/2\ \ 1)$, corresponding to the non-parental individuals.

The inverse of the additive genetic relationship matrix among the 6 parent-records is:

$$A^{-1} = \begin{bmatrix} 3/2 & 1/2 & 0 & 0 & -1 & 0 & 0 \\ 1/2 & 3/2 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 3/2 & 1/2 & 0 & -1 & 0 \\ 0 & 0 & 1/2 & 3/2 & 0 & -1 & 0 \\ -1 & -1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Assuming that the heritability and repeatability are $h^2 = 0.10$ and $r = 0.15$, then $k_1 = k_2 = 8.5$ and $k_3 = 17$.

The coefficient matrix of (25) is:

$$
\begin{array}{cccccccccccccc}
5.55 & 0 & 2.45 & 2.45 & 0.65 & 0.85 & 0.43 & 0.43 & 0 & 1.93 & 0.98 & 0.94 & 0 & 0 \\
 & 5.60 & 2.33 & 1.48 & 1.79 & 0 & 1.16 & 0 & 0 & 0.74 & 0.94 & 1.16 & 1.16 & 0.4? \\
 & & 5.27 & -0.38 & -0.12 & 0.43 & 0.37 & 0.43 & 0 & 0.70 & 0.33 & 1.31 & 0.79 & 0.4? \\
 & & & 4.42 & -0.12 & 0.43 & 0.79 & 0 & 0 & 1.64 & 0.33 & 0.37 & 0.37 & 0 \\
 & & & & 2.68 & 0 & 0.43 & 0 & 0 & 0.33 & 1.27 & 0.43 & 0 & 0 \\
 & & & & & 21.68 & 4.46 & -8.29 & -4.25 & -8.50 & 0 & 0 & 0 & -4.2? \\
 & & & & & & 13.54 & 0 & -4.25 & 0.37 & -8.50 & 0.21 & 0 & -4.2? \\
 & & \text{SYMMETRIC} & & & & & 12.96 & 4.25 & 0 & 0 & -8.50 & 0 & 0 \\
 & & & & & & & & 12.75 & 0 & 0 & -8.50 & 0 & 0 \\
 & & & & & & & & & 14.55 & 4.74 & 0 & -8.50 & 0 \\
 & & & & & & & & & & 14.18 & 0 & -8.50 & 0 \\
 & & & & & & & & & & & 18.53 & 0.37 & 0 \\
 & & & & & & & & & & & & 17.58 & 0.2 \\
 & & & & & & & & & & & & & 8.71
\end{array}
$$

The right hand side of (25) is:

(48.92 45.56 35.87 35.77 22.85 8.08 14.84 3.83 0 22.30 15.44 18.14 8.89 2.98)

The solution vector is:

$$
\begin{bmatrix}
\hat{A}_1 \\
\hat{A}_2 \\
\hat{B}_1 \\
\hat{B}_2 \\
\hat{B}_3 \\
\hat{G}_1 \\
\hat{G}_2 \\
\hat{a}^*_1 \\
\hat{a}^*_2 \\
\hat{a}^*_3 \\
\hat{a}^*_4 \\
\hat{a}^*_5 \\
\hat{a}^*_6 \\
\hat{a}^*_{11}
\end{bmatrix}
=
\begin{bmatrix}
4.89 \\
4.38 \\
1.89 \\
3.24 \\
4.14 \\
1.64 \\
0.25 \\
1.71 \\
0.98 \\
1.50 \\
0.09 \\
1.37 \\
0.79 \\
0.94
\end{bmatrix}
$$

The solution to the Mendelian sampling terms is obtained from (27). For animal 7, $\hat{C}_7 = 0.0769$, $n_7 = 1$ and the term in brackets is: $9 - 4.89 - 1.89 - 1((1/2)$ $1.64) - 1 ((1/2) 1.71) = 0.55$. Then, $\hat{m}_7 = (0.0769)(0.55) = 0.04$.

The remaining Mendelian sampling terms are:

$\hat{m}_8 = 0.13$ $\hat{m}_9 = 0.03$ $\hat{m}_{10} = 0.02$

$\hat{m}_{12} = -0.01$ $\hat{m}_{13} = -0.00$ $\hat{m}_{14} = 0.09$

The solution to the total genetic values are obtained from (28). For animal 7, noting that $v_s = 0$ and $v_d = 1$, we have:

$\hat{a}*_7 = (1/2)\ 1.71 + (1/2)\ 1.64 + 0.04 = 1.72.$

The remaining solution to the total genetic values are:

$\hat{a}*_8 = 0.94$ $\hat{a}*_9 = 0.83$ $\hat{a}*_{10} = 0.90$

$\hat{a}*_{12} = 0.86$ $\hat{a}*_{13} = 1.08$ $\hat{a}*_{14} = 1.04$

The solution to the total genetic values can also be obtained from (13):

$$
\begin{bmatrix} \hat{a}*_7 \\ \hat{a}*_8 \\ \hat{a}*_9 \\ \hat{a}*_{10} \\ \hat{a}*_{12} \\ \hat{a}*_{13} \\ \hat{a}*_{14} \end{bmatrix} =
\begin{bmatrix}
1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1/2 & 0 & 0 \\
0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\
0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} 1.71 \\ 0.98 \\ 1.50 \\ 0.09 \\ 1.37 \\ 0.79 \\ 0.94 \end{bmatrix} +
\begin{bmatrix}
1/2 & 0 \\
0 & 1/2 \\
0 & 0 \\
0 & 1/2 \\
0 & 0 \\
0 & 0 \\
1/2 & 1/2
\end{bmatrix}
\begin{bmatrix} 1.64 \\ 0.25 \end{bmatrix} +
\begin{bmatrix} 0.04 \\ 0.13 \\ 0.03 \\ 0.02 \\ -0.01 \\ -0.00 \\ 0.09 \end{bmatrix} =
\begin{bmatrix} 1.72 \\ 0.94 \\ 0.83 \\ 0.90 \\ 0.86 \\ 1.08 \\ 1.04 \end{bmatrix}
$$

Finally, the vector $\hat{p}$ is obtained from (29) or (30). For animal 7, $w_7 = 0.055$, the term in brackets in (30) is $(9 - 4.89 - 1.89 - 1\ (1.72)) = 0.5$ and therefore $\hat{p}_7 = w_7\ (0.5) = 0.027$. The remaining 9 permanent environmental effect solutions are:

$\hat{p}_3 = -0.09$ $\hat{p}_4 = -0.09$ $\hat{p}_5 = -0.01$

$\hat{p}_8 = 0.08$ $\hat{p}_9 = 0.03$ $\hat{p}_{10} = 0.02$

$\hat{p}_{12} = -0.01$ $\hat{p}_{13} = -0.00$ $\hat{p}_{14} = 0.05$

TABLE 1.

Example. Records of 14 animals.

ID	SIRE	DAM	A	B	Measurement
1	–	–	–	–	–
2	–	–	–	–	–
3	–	–	1	2	8
4	–	–	2	3	7
5	1	2	1	1	8
6	3	4	–	–	–
7	1	–	1	1	9
8	–	5	2	3	11
9	3	4	1	1	8
9	3	4	1	2	9
9	3	4	1	3	10
10	3	–	2	1	7
10	3	–	2	2	9
11	–	–	–	–	–
12	6	11	2	1	7
13	6	5	2	1	7
13	6	5	2	2	9
14	–	–	1	2	10

TABLE 2. Group of origin of the unknown parents

ID	Paternal group	Maternal group
1	1	1
2	1	2
3	1	1
4	2	2
7	–	1
8	2	–
10	–	2
11	1	2
14	2	1

References

Avalos, E; Smith, C. (1987). Genetic improvement of litter size in pigs. Animal Production 44, 153-164.

Bulmer, M.G. (1971). The effect of selection on genetic variability. Amer. Nat. 105, 201-211.

Christensen, A. (1978). Statistiske undersøgelser vedrørende kuldstørrelse hos Dansk Landrace. Matematisk Institut, KVL.

Foulley, J.L.; Chevalet, C. (1981). Methode de prise en compte de la consanguinité dans un modèle simple de simulation de per- formance. Ann. Génét. Sél. Anim. 13, 189-196.

Gianola, D., Im, S. and R.L. Fernando (1988). Prediction of breeding value under Henderson´s selection model: a revisitation. J. Dairy Sci. 71, 2790-2798.

Haley, C.S.; Avalos, E.; Smith, C. (1988). Selection for litter size in the pig. Anim. Breed. Abstr. 56, 317-332.

Henderson, C. (1973). Sire evaluation and genetic trends. Proc. Anim. Breed. Genet. Symp. Honor of Dr. J.L. Lush, Am. Soc. Anim. Sci., Am. Dairy Sci. Assoc., Champaign, Il.

Henderson, C.R. (1975). Best linear unbiased estimation and prediction under a selection model. Biometrics 31, 423-449.

Henderson, C.R. (1976). A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32, 69-83.

Hill, W.G.; Webb, A.J. (1982). Genetics of reproduction in the pig. In: Control of Pig Reproduction. Ed. D.J.A. Cole and G.R. Foxcroft. London, U.K. Butterworths.

Im, S., R.L. Fernando and D. Gianola (1989). Likelihood inferences in animal breeding under selection: a missing-data theory view point. Genet. Sel. Evol. 21, 399-414.

Kackar, R.N. and D.A. Harville (1981). Unbiasedness of two-stage estimation an prediction procedures for mixed linear models. Commun. Statist. Theor. Math A10, 1249-1268.

Le Roy, P.; Legault, C.; Gruand, J.; Ollivier, L. (1987). Realised heritability for litter size in selection of "hyperprolific" sows. Génét. Sél. Évol. 19 351-363.

Meyer, K. (1989). Restricted maximum likelihood to estimate variance component for animal models with several random effects using a derivative-fre algorithm. Genet. Sel. Evol. 21, 317-340.

Mäki-Tanila, A.; Kennedy, B.W. (1986). Mixed model methodology under geneti models with small number of additive and non-additive loci. Proc. 3rd World Congr. Genet. Appl. Livest. Prod. XII: 443-448.

Pearson, K. (1903). On the influence of natural selection on the variability an correlation of organs. Phil. Trans. Roy. Soc. Lond. A200, 1-66.

Quaas, R.L. (1976). Computing the diagonal elements and inverse of a larg numerator relationship matrix. Biometrics 32, 949-953.

Quaas, R.L. (1988). Additive genetic model with groups and relationships. J. Dair Sci. 71, 1338-1345.

Quaas, R.L.; Pollak, E.J. (1980). Mixed model methodology for farm and ranch bee cattle testing programs. J. Anim. Sci. 51, 1277-1287.

Quaas, R.L. and E.J. Pollak (1981). Modified equations for sire models wit groups. J. Dairy Sci. 64, 1868-1872.

Robinson, G.K. (1986). Group effects and computing strategies for models fo estimating breeding values. J. Dairy Sci. 69, 3106-3111.

Smith, C. (1964). The use of specialised sire and dam lines in selection for mea production. Animal Production 6, 337-344.

Smith, C.; Dickerson, G.E.; Tess, M.W.; Bennett, G.L. (1983). Expected relative responses to selection for alternative measures of life cycle economic efficiency of pork production. J. Anim. Sci. 56, 1306-1314.

Smith, S.P. and A. Mäki-Tanila (1990). Genotypic covariance matrices and their inverses for models allowing dominance and inbreeding. Génét. Sél. Évol. 22, 65-91.

Sorensen, D.A. and B.W. Kennedy (1986). Analysis of selection experiments using mixed model methodology. J. Anim. Sci. 63, 245-258.

Thompson, R. (1979) Sire Evaluation. Biometrics 35, 339-353.

van der Steen, H.A.M.; Kock, M.G.M. (1987). Sire of litter and sire of dam effects on categorical and normally distributed litter size traits in pigs. 38th Annual Meeting of the European Association for Animal Production. Lisbon, Portugal.

Vangen, O. (1986). Genetic control of reproduction in pigs from parturition to puberty. Proc. 3rd World Congr. Genet. Appl. Livest. Prod. VII. 288-292.

Westell, R.A., Quaas, R.L. and L.D. Van Vleck (1988). Genetic groups in an animal model. J. Dairy Sci. 71, 1310-1318.