Progeny group size for evaluating natural service bulls using AI reference sires

J. L. FOULLEY and Françoise CLERGET-DARPOUX*

Station de Génétique quantitative et appliquée
Centre national de Recherches zootechniques, I.N.R.A.,
78350 Jouy-en-Josas
*I.N.S.E.R.M., U155 Chateau de Longchamp
75016 Paris

Summary

Use of on-farm beef performance recording for evaluating NS sires might be a means to improve the efficiency of beef sire selection. However, the validity of ranking sires from a large number of herds remains restricted as long as systematic connections between herds have not been realized.

This study determines the sizes of the progeny groups to be tested in a system where these ties result from the wide diffusion of AI sires.

Progeny group sizes giving a specified accuracy where calculated and optimized taking account of the number of NS[1] sires used, assuming the use of a single AI[1] reference sire. Thus, for a given cœfficient of determination of 0.4 calculated on the basis of a heritability $h^2 = 0.2$, the following progeny numbers n (progeny per NS sire) and m (progeny of AI reference sire) have to be recorded in the herd : $(n = 25; m = 25)$, $(n = 23; m = 19)$ and $(n = 21; m = 12)$ in order to evaluate one, two or three sires respectively.

The effects of various factors such as number of reference sires, common environmental effects and mode of sampling sires are also discussed.

Introduction

For about the last twenty years, rational selection of beef cattle in France has mainly been intended for AI bulls. This trend has led in particular to the progressive setting up of integrated selection schemes of males used for either terminal crossing or producing breeding females. In the latter situation, a rather large fraction in the population, variable according to breeds from 85 p. 100 for *Charo-*

---

[1] N.S., A.I. : natural service and artificial insemination respectively.

lais to 30-50 p. 100 for *Limousin* and *Blonde d'Aquitaine* breeds is still bred by natural service. The value of this population both for selection within itself and for the diffusion of selected genes after progeny testing on breeding qualities remains limited, however, as long as the breeding value of NS sires has not been calculated objectively.

To that end, the data gathered from field recording of growth to weaning may be used with the aim of evaluating sires on their progeny. However, the immediate use of these data will suffer from a lack of exchanges of bulls between herds. In fact, because of the very low number of service bulls used in a herd, the mean genetic level of the sires may vary considerably fom one herd to another. An evaluation of sires based on the classical contemporary comparison does not appear in these conditions to be relevant since a sire compared to sires of contemporaries of a high genetic level will be disadvantaged relative to another compared to sires of a lower level. SPIKE and FREEMAN (1977) have emphasized, in relation to this, the value of correcting for herd effects including environmental effects only (and not the genetic effects) especially if bulls have been evaluated on a small number of progeny.

For that purpose, different types of design of connections between herds can be imagined. A large scale diffusion of AI sires appears to be an attractive and realistic solution. This system is already (or will be) applied in the main countries possessing specialized beef herds. With respect to the progeny numbers to be recorded, some standards have already been supplied such as those of the Beef Improvement Federation in the USA (ANONYMOUS, 1972). However in our opinion, the justification given seems to be incomplete from a theoretical point of view; therefore the aim of this study is to discuss this problem with a more rational approach.

Mathematical model

A. — *Generalities*

This study is based on the assumption that sires are evaluated by BLUP procedures of HENDERSON (1973). BLUP, the predictor with minimum mean square error of prediction in the class of unbiased linear predictors maximizes the probability of a correct ranking on true genetic values (s_i), when individuals are ranked on predicted $(\hat{s}_i)$ values (HENDERSON, 1973).

Therefore, we shall discuss the numbers of progeny to be recorded by referring to the minimum $E(\hat{s}_i - s_i)^2$ criterion or equivalently to the maximum coefficient of determination (CD).

The model used is the following :

$$y_{ijkl} = \mu + s_i + h_{jk} + e_{ijkl} \qquad (1)$$

where:

μ is the population mean,

s_i the effect of the ith sire $(i = 1, 2, ..., p)$,

h_{jk} the combined effect of the jth herd and the kth year $(j = 1, 2, ..., q;$ $k = 1, 2, ..., t)$

e_{ijkl} the residual effect for the lth progeny $(l = 1, 2, ..., n_{ijk})$ of sire i in herd-year jk.

In addition, we assume that :

— the s_i are independent with expectation zero and variance σ^2_s,
— the h_{jkl} are fixed effects,
— the e_{ijkl} have a homogeneous variance σ^2_e and are at first, independent.

In these conditions, it has been shown that:
i) BLUP solutions are obtained by solution of

$$A\ \hat{s} = B \qquad (2)$$

where

— A is a symmetric (p, p) matrix relative to the p sire effects which is a function of the progeny numbers and of the ratio $\lambda = \dfrac{\sigma^2_e}{\sigma^2_s}$.

$$\left. \begin{aligned} A_{ii} &= \sum_{jk} n_{ijk}\left(1 - \frac{n_{ijk}}{n_{0jk}}\right) + \lambda \\ A_{ii}{}' &= -\sum_{jk} \frac{n_{ijk}\ n_{i'jk}}{n_{0jk}} \text{ for } i \neq i' \end{aligned} \right\} \qquad (3)$$

with $n_{0jk} = \sum_i n_{ijk}$

* $\hat{s}$ is the vector of BLUP solutions for the p sires to be evaluated,
* B the right-hand side vector of the system defined by :

$$B_i = \sum_{jkl} y_{ijkl} - \sum_{jk} \frac{n_{ijk}}{n_{0jk}} \left(\sum_{il} y_{ijkl} \right)$$

ii) The variance-covariance matrix of prediction errors is given by:

$$E[{}^t(\hat{s} - s)(\hat{s} - s)] = A^{-1}\ \sigma^2_e \qquad (4)$$

Hence, in particular the expression of the coefficient of determination for sire i is

$$CD_i = 1 - \lambda\ w_i \qquad (5)$$

w_i being the ith diagonal term of A^{-1}.

B. — Formulation of the simple case with only one reference sire

It is useful to begin by discussing the simple situation where connections are made by one only AI sire.

In addition, the form of the matrix A is simplified under the following hypotheses:

— NS sires have progeny in only one herd and the different sires of a given herd have the same number of progeny;

— the AI reference sire has progeny each year in all herds and the proportion of calves derived from $AI(\theta_j)$ is constant from year to year in a given herd j.

Under these conditions, the use of (3) enables the matrix A to be written in partitioned form, with the NS sires arranged by herd ($p'j$ bulls in herd j) in one part, and the reference sire (index r) in the other part.

$$
A = \begin{bmatrix}
A_1 & & & & & R_1 \\
 & \cdot & & & O & \vdots \\
 & & \cdot & & & \vdots \\
 & & & A_j & & R_j \\
O & & & & \cdot & \vdots \\
 & & & & A_q & R_q \\
\hline
R'_1 & \cdots & R'_j & \cdots & R'_q & u_r
\end{bmatrix}
\tag{6}
$$

where A_j, R_j, are square blocks and vectors of size p'_j respectively such as:

$$
A_j = \begin{bmatrix}
u_j & & \\
 & \cdot & u'_j \\
 & \cdot & \\
u'_j & & u_j
\end{bmatrix} \; ; \qquad
R_j = \begin{bmatrix}
u''_j \\
\vdots \\
u''_j
\end{bmatrix}
$$

with (see 3)

$$
u_j = \sum_k n_{jk}\left(1 - \frac{n_{jk}}{n_{0jk}}\right) + \lambda = n_{j0}\left(1 - \frac{1 - \theta_j}{p'_j}\right) + \lambda
$$

$$
u'_j = -\sum_k \frac{n^2_{jk}}{n_{0jk}} = \frac{1 - \theta_j}{p'_j}\, n_{j0}
$$

$$
u''_j = -\sum_k \frac{n_{jk}\, m_{jk}}{n_{0jk}} = -\,\theta_j n_{j0}
$$

$$
u_r = \sum_{jk} m_{jk}\left(1 - \frac{m_{jk}}{n_{0jk}}\right) + \lambda = \sum_j p'_j\,\theta_j n_{j0} + \lambda
$$

n_{jk}, m_{jk} indicate the numbers of progeny for each NS sire and for the reference sire in the herd-year jk respectively and
n_{0jk} represents the progeny number in the herd-year jk and n_{j0} the total progeny number in herd j.

Results

A. — *Calculation of the accuracy*

An algebraic expression of the inverse of matrix A defined in 6 can be obtained by inverting A by blocks. Details are given in appendix 1. The part of matrix A^{-1} relative to p'_j NS sires of herd j is

$$
W_j = \frac{1}{n_{j0} + \lambda}\, I p'_j + \left[\frac{\tau_j}{n_{0j} + \lambda}\cdot\frac{1 - \theta_j}{p'_j\theta_j} + \frac{\tau^2_j}{\left(\sum_j p'_j\tau_j + 1\right)}\right] J p'_j
\tag{7}
$$

where $I p'_j$ the unit matrix of size p'_j.

$J p'_j$ the square matrix of size p'_j with all elements one

τ_j defined as $\tau_j = \dfrac{\theta_j n_{j_0}}{\theta_j n_{j_0} + \lambda}$

In most practical situations, the last term of (7) can be neglected. Hence, we obtained the following expression of the term w_j (variance of prediction errors for a NS sire of herd j)

$$w_j = \frac{1}{n_{j0} + \lambda}\left(1 + \frac{n_{j0}}{(1 - \theta_j)}\frac{p'_j}{\theta_j n_{j0} + \lambda}\right) \tag{8}$$

The coefficient of determination can then be written, in simplified notations as follows

$$CD = \frac{n}{n + \lambda}\left(1 - \frac{\lambda n}{nm + \lambda N}\right) \tag{9}$$

where :

— n, m are the numbers of progeny per herd of a NS sire and of the reference sire respectively,

— N the total number of recorded progeny in a herd

$$N = m + p'n \tag{10}$$

— $\lambda = \dfrac{4}{h^2} - 1$; h^2 being the heritability of the trait considered.

It is to be noticed that assuming the existence of a sufficient number of sires in the evaluation program (at least 85) the accuracy depends only on the characteristics of the herd using the sire to be evaluated.

B. — Numerical application

If (λ, p', n and m) are known, it is possible using formula (8) to calculate the accuracy of the predictor s_i. In fact, a more interesting goal is to try to optimize the progeny numbers (n, m) according to the level of accuracy desired. For a fixed size N of herd and number p' of NS sires, there is an optimum distribution of the numbers (n, m) maximizing the CD. This is an optimization with one only variable since n, m are related by (10). A graph is given in figure 1 showing for $h^2 = 0.2$ the curves giving maximum accuracy according to N and p'. The graph also shows the iso-n and iso-m curves. Using this graph, it is possible to determine for a given CD the number N of calves to be recorded and the optimum distribution (n, m) of the latter knowing the number of NS bulls used. Thus, for instance, in herds having two NS sires, it is necessary to test a total of 100 calves to get a CD of 0.5, these 100 calves being distributed into $2n = 68$ and $m = 32$.

In the practical situation of French beef breeds subjected to performance recording, growth and conformation traits at weaning exhibit a heritability of about 0.2 (MOLINUEVO and VISSAC, 1972) and the number of NS bulls used per herd varies between 1 and 3. In these conditions, table 1 gives the optimum numbers of progeny of the reference sire (m) and per NS sire (n) for various levels of accuracy (CD = 0.30; 0.35 and 0.40). Thus, on the basis of a CD of 0.40 which may be considered as a reasonable threshold of accuracy for evaluating this type of sire, it is necessary to test the following numbers of progeny (n, m) per herd : (25, 25); (23, 19) and (21, 12) where herds use one, two or three NS bulls respectively.

TABLE I

Progeny numbers to be recorded per herd for given accuracy when evaluating NS sires by using an AI reference sire

	Number of NS bulls to be evaluated (p')		
	1	2	3
CD = 0.30 [1] : — Total progeny number (N)	35	40	45
— Distribution of progeny $(n,\ m)$ [2] optimum.	17; 18	15; 10	13; 6
suboptimum [3]	24; 11	19; 2	14; 3
CD = 0.35 [1] : — Total progeny number (N).	40	50	60
— Distribution of progeny (n, m) [2] optimum	20; 20	18; 14	17; 9
suboptimum [3]	13; 27 27; 13	12; 26 23; 4	12; 24 19; 3
CD = 0.40 [1] : — Total progeny number (N)	50	65	75
— Distribution of progeny (n, m) [2] optimum	25; 25	23; 19	21; 12
suboptimum [3]	16; 34 34; 16	15; 35 29; 7	14; 33 24; 3
with c^2 effects ($c^2 = 0.05$) [4].	45; 45	40; 30	30; 15

[1] CD : coefficient of determination calculated for a cœfficient of heritability $h^2 = 0.20$.
[2] n : progeny number to be recorded for NS bull in a herd.
    m : progeny number of the AI reference sire to be recorded per herd.
[3]   : progeny numbers providing an accuracy of 0.9 CD.
[4]   : standards for a model with c^2 effects and progeny recorded during 3 years.

In addition, it will be noticed that a certain range of variation is allowed for these numbers of progeny without affecting the accuracy too much as shown by results in table 1 indicating for the same total number N the two combinations (n, m) providing 90 p. 100 of maximum accuracy. In particular, it is possible to substantially reduce the number of progeny m derived from AI (naturally by increasing n per contra) without reducing the accuracy too much. Thus, for example with 2 sires to be evaluated in a herd, 65 progeny divided into $m = 19$ and $2n = 46$ are needed in order to get an accuracy (CD) of 0.4. Recording again a total of 65 progeny, it is possible theoretically to adopt a distribution into $m = 7$ and $2n = 58$ without reducing the accuracy by more than 10 p. 100. It is important to emphasize this possibility because in practice obtaining a sufficiently high rate of AI appears to be one of the most limiting factors for the application of such an evaluation system.
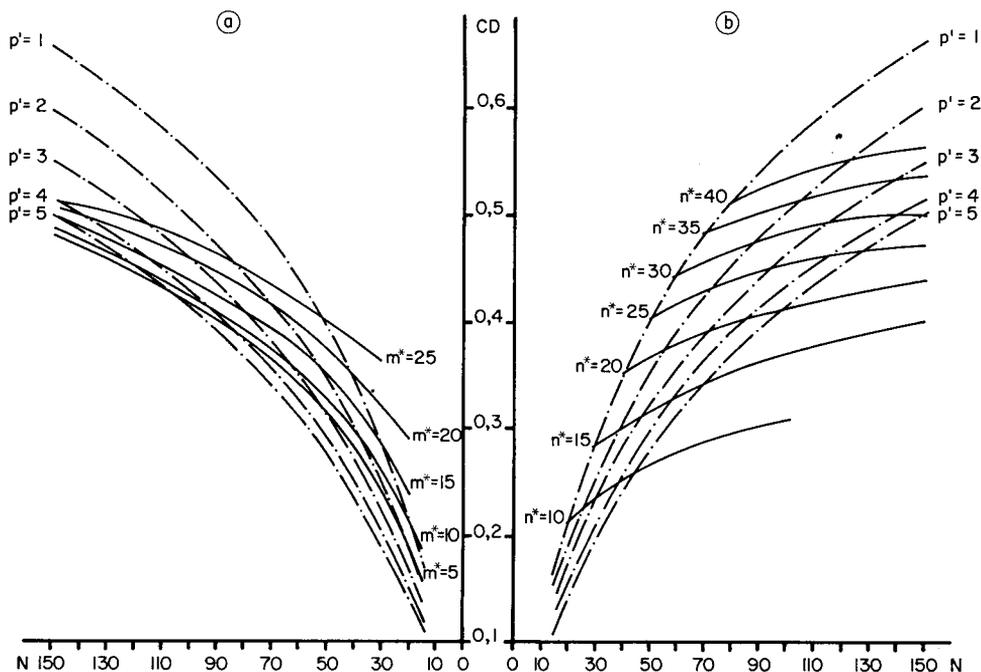
FIG. 1. — *Coefficient of determination (CD) as a function of herd size (N), number of NS bulls (p') and progeny numbers of the AI reference bull (m) and per NS bull (n) (respectively)*

a) curves iso-m^* ($\lambda = \sigma^2_e / \sigma^2_s = 19$)
b) curves iso-n^*

Discussion

A. — *Comparison with previous results*

Contrary to the american data (NIELSEN, 1974) the number of progeny (m) required for the reference bull decreases when the number p' of NS sires increases or, in other words, when the size (N) of herd increases. The following explanation can be given on account of the model adopted: when p' increases, the amount of information useful for a given sire and which comes from contemporaneous calves of the herd derived from NS (i.e. intra block information) also increases. As reasoning is based on a fixed total accuracy, the part of information needed from AI progeny (i.e. inter block information) decreases accordingly.

B. — *Utilization of several AI reference sires*

A design using only one reference sire has the advantage of being simple. Moreover, it is theoretically ideal, at least in the absence of a marked sire x herd interaction. For a concrete estimation of the magnitude of this interaction and for

giving a choice to the breeders rendering the programm more attractive, it is necessary to suggest several AI reference sires.

Thus, in some associations several reference sires are available for the breeders who are required to use two of them in their herd, one being chosen freely by the breeder. NIELSEN (1974) clearly showed the value of such a measure where the total number of reference sires available is limited to 4 in comparison with an unplanned program where the overall number of AI sires (10 in his study) and the number used in each herd (5 to 10) are much higher.

In any case, it is very important to create a minimum of spatial and temporal links between these AI sires. With respect to the former, the examination of the particular case where 2 AI sires are available for the breeders shows that a large latitude of utilization is possible in the distribution of progeny of these two sires provided the standards of offspring are followed (see appendix 2).

Concerning connections between years, it may be suggested like the Beef Improvement Federation to use each AI sire at least two years in the program so as to calculate the prediction from a common fixed genetic basis. Besides, within herds, repetitions of one same AI sire at least should be made from one year to another, particularly when obliged to wait several years before reaching the numbers of progeny necessary to evaluate a sire.

C. — *Effect of considering a common c^2 environmental effect*

With the aim of chosing sires from different herds and consequently for a potential use of these bulls in different herds, it is important, not only to correct accurately for the effects of herds, but also to take into consideration that, even adjusted, the offspring of the same sire reared in the same herd are generally more alike than when kept in different herds. This effect called common c^2 environmental effect implies in model (1) that the residual variables e_{ijkl} are no longer independent. Various structures of covariances may exist. That generally adopted (the most simple but not necessarily the most realistic one) corresponds to a common environmental effect within a herd year jk as shown in the following model:

$$e_{ijk} = c_{ijk} + \varepsilon_{ijkl} \qquad\qquad (11)$$

where:

— the c_{ijk} are independent with expectation zero and homogeneous variance σ^2_c,

— the ε_{ijk} are independent among themselves, independent of c_{ijk} and have variance σ^2_ε.

The system of equations giving the predictions of s_i in this new model, after absorption of μ, h_{jk} and c_{ijk} can easily be deduced from that of the previous model (see 2 and 3) by replacing (HENDERSON, 1974).

$$n_{ijk} \text{ by } \frac{\beta n_{ijk}}{n_{ijk} + \beta}$$

where

$$\beta = \frac{\sigma^2_\varepsilon}{\sigma^2_c}$$

This means that the number of progeny in a herd-year is replaced by half the harmonic mean between this number and β, β becoming the upper limit of this mean.

Accordingly if (n^*, m^*) are the annual standards deduced from model 1, the corresponding numbers (n^*_c, m^*_c) in the presence of c^2 effects in (11) are obtained by

$$n^*_c = \frac{\beta n^*}{\beta - n^*} \quad \text{and} \quad m^*_c = \frac{\beta m^*}{\beta - m^*}$$

If assuming like some authors (NIELSEN, 1974) that σ^2_c may be of the same magnitude as σ^2_s, it appears that its effect on the accuracy of $\hat{s}_i$ is far from negligible especially in herds using 1 or 2 bulls (Table 1). Thus, for these herds, it is necessary to increase the progeny groups recorded by 70 to 80 p. 100 if one desires the same accuracy (CD = 0.40) as previously and assuming that σ^2_c equals 5 p. 100 of the variance σ^2_y. For the evaluation of one NS sire, the standards (n, m) change from (25, 25) to (45, 45); with two sires in a herd (40, 30) are needed per sire instead of (23, 19). c^2 effects probably exist in French beef herds in particular because of birth grouping, feeding and choice of dams. However, its magnitude has not been determined due to lack of sufficient exchanges of bulls. It is probably necessary to distinguish with respect to this, selection from commercial herds where more homogeneous management between paternal offsprings is expected. As sufficiently accurate estimates of c^2 effects are lacking at the present time, it is necessary, on the one hand to demand a rigorous design in order to limit them as much as possible and, on the other hand, not to fix standards which are to low.

D. — *Sampling of sires*

a) *Existence of subpopulations*

The model used assumes that all sires including reference sires are randomly sampled from the same population. This hypothesis may appear unrealistic.

As the objective is to compare NS sires, it may be thought that introducing a group effect for those sires does not appreciably modify the results obtained here. In fact, the number of NS sires being very high relative to the number of reference sires, the estimate of this group effect will be little different from the general mean and of similar accuracy. With different objectives, if in particular the aim is to compare NS sires to certain reference sires, a specific study of the accuracy and of the progeny numbers should be undertaken although, in this case, the value of grouping would be also questionable on account of the large number of progeny per reference sire.

b) *Distribution of sires according to herds*

The differences between mean genetic values of sires used in different herds are largely due to a sampling of a very small number of bulls used per herd-year. It may be that these differences are also affected by a choice of sires within herds. According to HENDERSON (1973) and FIMLAND (1975) the bias due to this type of selection is eliminated by considering herds as fixed like in model 1.

Besides, if this sire selection is related to the herd effect h, the model should be changed because of this correlation between s and h effects.   It might be assumed that if the NS sires are produced and used in the same herd, a relationship will appear between s and h effects since h includes the genetic level of stock females.   However, a rather large number of breeders buy bulls from outside and the criteria of their choice may be rather different from those considered in the evaluation so that the relationship (h, s) is much less obvious than assumed *a priori*.

E. — *Value of criterion* R^2 — *Conclusion*

The first element to be discussed is the value of choosing the criterion R^2 to plan this selection program.   This criterion is in fact well adapted to provide an overall appreciation of the validity of ranking sires relative to their true breeding values.   If we are interested in particular comparisons between sires, we must take into account not only the variances but also the covariances between prediction errors.

The covariance relative to 2 sires of the same herd j can be obtained from formula (7) i.e.

$$\left[\frac{\tau_j}{n_{0j} + \lambda} \cdot \frac{1 - \theta_j}{p'_j \theta_j} + \frac{\tau^2_j}{\lambda(\sum_i p'_j \tau_j + 1)} \right] \sigma^2_e$$

For 2 sires used in different herds j and j', this covariance is

$$\frac{\tau_j \cdot \tau_j'}{\lambda(\sum_j p'_j \tau_j + 1)} \sigma^2_e$$

On account of the low value of this term (see appendix 1), the variance of prediction errors of the difference between these 2 sires is practically equal to the sum of approximate variances of prediction errors (see 8).   Reasoning at constant level of CD for all sires leads to a constant variance of error for the difference between 2 sires of different herds.   Thus for $h^2 = 0.2$ and the progeny numbers defined for CD = 0.4, this variance equals $6.32 \times 10^{-2}\sigma^2_e$ (i.e. a standard deviation of 0.25 σ_e). As regards the error committed when comparing 2 sires within the same herd, it is naturally lower.   In our model, this variance is $2 \sigma^2/(n + \lambda)$; for the optimum progeny groups of 18 and 23 given in table 1, its value then varies between 4.8 and $5.4 \times 10^{-2}\sigma^2_e$.

Furthermore as well demonstrated by ROBERTSON (1957) the determination of the size of progeny groups should result not only from consideration of R^2 but also from the parameters R, i (selection intensity) and L (generation length) which are influencing the expected genetic improvement.

The use which will be made of the indices for selection purposes will determine the efficiency of such an evaluation system as we know that the power of decision is not centralized but rather atomized among different categories of breeders. In addition as the means of recording are limited (size of herds, number of cows bred per year to one bull, AI rate practised) obtaining a certain level of accuracy may imply progeny testing over several years.   Here we have to find a compromise again between accuracy and generation length.

The procedure adopted here consisting fixing progeny group size for a given level of accuracy appears to be incomplete.   Its main purpose is to show clearly

to the breeders'associations the constraints of applying such a system and accordingly to avoid an anarchic setting up of the latter.

More generally, the model used in this study may seem excessively simple. However, it seems to us to provide a concrete basis for an initial discussion with breeders of the possibilities of using such a system under French conditions.

Acknowledgements

Résumé

Effectifs de descendants à contrôler en vue de l'évaluation des taureaux de monte naturelle à partir de taureaux de connexion diffusés par I.A.

L'utilisation du contrôle de performances bouchères en ferme en vue de l'évaluation des taureaux de monte naturelle peut être un moyen pour améliorer l'efficacité du choix des taureaux de race à viande. La validité du classement d'un grand nombre de taureaux effectuant la monte dans des troupeaux différents reste toutefois limitée tant que des connexions systématiques entre ceux-ci ne seront pas réalisées.

Cette étude établit les effectifs de descendants à contrôler dans un système où ces liens résultent de la large diffusion de taureaux d'insémination artificielle. Les effectifs ont été calculés pour un niveau donné de précision des indices et optimisés compte tenu du nombre de taureaux de monte naturelle utilisés en supposant au départ, l'existence d'un seul taureau d'IA de connexion. Ainsi, pour un coefficient de détermination de 0,4 calculé sur la base d'un coefficient d'héritabilité de 0,20, il faut contrôler dans le troupeau les effectifs n (descendants par taureau de monte naturelle) et m (descendants du taureau d'IA de connexion) suivants : $(n = 25; m = 25)$, $(n = 23; m = 19)$ et $(n = 21; m = 12)$ selon que l'on veut qualifier un, deux ou trois taureaux par troupeau respectivement.

Les incidences de différents facteurs tels que le nombre de taureaux de connexion, les effets de milieu commun et le mode d'échantillonnage des taureaux sont abordés dans la discussion.

References

ANONYMOUS, 1972. *Guidelines for uniform beef improvement programs. Program aid 1020.* Extension service. USDA, 70 p.

FIMLAND E. A., 1975. Estimation of sire's breeding value. IV. Maximum liklihood derivation of the predictor and the estimator of a general mixed model given some type of selection of a correlated random variable. *Z. Tierzüchtg Züchtgs Biol.*, **92**, 176-187.

HENDERSON C. R., 1973a. Sire evaluation and genetic trends. In *Proc. Anim. Breed. Genet. Symp.*, Blacksburg, Virginia, p. 10-41. *American Society of Animal Science*, Champaign, Illinois.

HENDERSON C. R., 1974. General flexibility of linear model techniques for sire evaluation. *J. Dairy Sci.*, **57**, 963-972.

MOLINUEVO H. A., VISSAC B., 1972. Variabilité génétique de la croissance avant sevrage dans les races *Charolais* et *Limousine*. *Ann. Génét. Sél. anim.*, **4**, 423-444.

NIELSEN M. K., 1974. *An examination of analysis procedure for national beef sire evaluation programs.* Unpublished Ph. D. thesis, Iowa State University, Ames, Iowa, 84 p.

ROBERTSON A., 1957. Optimum group sire in progeny testing and family selection. *Biometrics*, **4**, 442-450.

SPIKE P. L., FREEMAN A. E., 1977. Effect of genetic differences among herds on accuracy of selection and expected genetic change. *J. Dairy Sci.*, **60**, 967-974.

Appendix

I. — *Calculation of the inverse of matrix A*

By partitioning A^{-1} in the same way as A,

$$A^{-1} = \begin{bmatrix} W_{11} & & & & & V_1 \\ & \ddots & & & & \vdots \\ & & W_{jj} & W_{jj'} & & V_j \\ & & & \ddots & & \vdots \\ & & W_{j'j} & & \ddots & \\ & & & & W_{qq} & V_q \\ \hline V'_1 & \cdots & V'_j & \cdots & V'_q & v_r \end{bmatrix}$$

and by writing that $AA^{-1} = I$, we get

$$W_{jj} = A^{-1}{}_j + v_r \cdot A^{-1}{}_j \cdot R_j \cdot R'_j \cdot A^{-1}{}_j$$

$$V_r = \frac{1}{u_r - \sum\limits_{j=1}^{q} R'_j \cdot A^{-1}{}_j \cdot R_j}$$

we notice that A_j may be written as

$$A_j = (n_{j0} + \lambda)I_{p'j} + u'_j J_{p'j}$$

$J_{p'j}$ denoting the square matrix of size $p'j$ in which all elements are one.

Hence, $A^{-1}{}_j$ is

$$A^{-1}{}_j = \frac{1}{n_{j0} + \lambda}\left(I_{p'j} - \frac{u'_j}{\lambda - u''_j} J_{p'j}\right)$$

In addition,

$$R'_j . R_j = p'_j u''^2_j$$
$$R'_j . J_{p'_j} R_j = p'^2_j u''^2_j$$
$$R_j . R_j = u''^2_j . J_{p'_j}$$

Then,

$$v_r = \cfrac{1}{\lambda \left(\sum_j \cfrac{p'_j . u''_j}{\lambda - u''_j} + 1 \right)}$$

$$A^{-1}_j . R_j . R'_j . A^{-1}_j = \frac{u''^2}{(\lambda - u''_j)^2} J_{p'_j}$$

Let

$$\tau_j = -\frac{u''_j}{\lambda - u''_j} = \frac{\theta_j n_{j0}}{\theta_j n_{j0} + \lambda}$$

The blocks W_{jj} in which we are interested may be written as follows:

$$W_{jj} = \frac{1}{n_{j0} + \lambda} \left[I_{p'_j} + \frac{1 - \theta_j}{p'_j \theta_j} \tau_j J_{p'_j} \right] + \frac{\tau^2_j}{\lambda (\sum_j p'_j \tau_j + 1)} J_{p'_j}$$

The term $X = \dfrac{\tau^2_j}{\lambda (\sum_j p'_j \tau_j + 1)}$ can be neglected in practice, in most cases.

Thus, provided the total number of bulls involved in the program is at least 85, the overestimation of CD resulting from neglecting X is a maximum of 10^{-2} when basing the calculations on values of m (optimum and the lowest suboptimum) given in table 1 for CD = 0.4 and $h^2 = 0.2$ and assuming variable proportions of herds using one, two or three bulls [(a) 0.40; 0.35; 0.25 or (b) 0.10; 0.50; 0.40].

2. — Study of the design involving 2 references sires

In the case of this design, we consider only herds having one or two bulls to be evaluated; herds will then be denoted type 1 or type 2. In the herds of type 1, breeders will have the choice between the two reference sires. We suggest that L_1 breeders will use one sire and L_2 the other and that they also apply the previously defined, progeny group standards. The L_3 breeders having two NS bulls to be tested, should use the two reference sires according to the standards that are to be established.

It will be assumed that in each herd of type 2, the different NS sires have the same number of progeny and the AI rate is constant from one year to another. Let:

— n_1 (n_2 respectively) be the total number of offspring of a NS sire in a herd of type 1 (type 2 respectively),

— m_1 be the total number of offspring of the reference bull used in each herd of type 1,

— m_2 and m_3 be the number of offspring per herd of each of the two reference bulls used in herds of type 2.

so that
$$N_1 = n_1 + m_1$$
$$N_2 = 2n_2 + m \qquad \text{with} \qquad m = m_2 + m_3$$

By regrouping, on the one hand the NS bulls and, on the other, the AI bulls we obtain a partition of A in the form

$$A = \left[\begin{array}{c|c} M & R \\ \hline R' & R_0 \end{array} \right]$$

Blocks M, R and R' may be partitioned themselves according to L_1, L_2 and L_3 herds. A becomes then

$$A = \left[\begin{array}{ccc|c} A_1 & & \underline{0} & R_1 \\ & A_2 & & R_2 \\ \underline{0} & & A_3 & R_3 \\ \hline R'_1 & R'_2 & R'_3 & R_0 \end{array} \right]$$

The constitutive submatrices of A have the following characteristics:

- A_1 (L_1, L_1) and A_2 (L_2, L_2) have the form uI

$$\text{with} \qquad u = \frac{n_1 m_1}{N_1} + \lambda$$

- $A_3(2L_3,\ 2L_3) = \left[\begin{array}{cc} \left\{ \begin{array}{cc} v & w \\ w & v \end{array} \right\} & \underline{0} \\ \underline{0} & \left\{ \begin{array}{cc} v & w \\ w & v \end{array} \right\} \end{array} \right]$

with $v = n_2 \left(1 - \dfrac{n_2}{N_2} \right) + \lambda$

$w = - \dfrac{n_2^2}{N_2}$

- $R_1(L_1,\ 2) = \left[\begin{array}{cc} r_1 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ r_1 & 0 \end{array} \right]$ \qquad with \quad $r_1 = - \dfrac{n_1 m_1}{N_1}$

Likewise $R_2(L_2,\ 2) = \left[\begin{array}{cc} 0 & r_1 \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & r_1 \end{array} \right]$

- $R_3(2L_3,\ 2) = \left[\begin{array}{cc} r_2 & r_3 \\ \vdots & \vdots \\ \vdots & \vdots \\ r_2 & r_3 \end{array} \right]$

with $r_2 = - n_2 \dfrac{m_2}{N_2}$

$r_3 = - n_2 \dfrac{m_3}{N_2}$

- $R_0 (2, 2) = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$

with $a = L_1 m_1 \left(1 - \dfrac{m_1}{N_1}\right) + L_3 m_2 \left(\left(1 - \dfrac{m_2}{N_2}\right) + \lambda\right)$

$b = - L_3 \dfrac{m_2 m_3}{N_2}$

$c = L_2 m_1 \left(1 - \dfrac{m_1}{N_1}\right) + L_3 m_3 \left(1 - \dfrac{m_3}{N_2}\right) + \lambda.$

We partition A^{-1} in the same way as A, i.e.

$$A^{-1} = \begin{bmatrix} W_{11} & W_{12} & W_{13} & T_1 \\ W'_{12} & W_{22} & W_{23} & T_2 \\ W'_{13} & W'_{23} & W_{33} & T_3 \\ \hline T'_1 & T'_2 & T'_3 & T_0 \end{bmatrix}$$

Were are only interested in W_{33} since it is this matrix which is used to calculate the accuracy of the evaluation of the bulls used in herds of type 2.

As $AA^{-1} = I$, then

$W_{33} = A^{-1}_3 + A^{-1}_3 . R_3 (R_0 - R'.M^{-1}.R.)^{-1}.R'_3.A^{-1}_3$

Let :

$$\begin{bmatrix} x & y \\ y & x \end{bmatrix} = \begin{bmatrix} v & w \\ w & v \end{bmatrix}^{-1} = \frac{1}{n_2 + \lambda}\left[I + \frac{n^2_2}{n_2 m + \lambda N_2} J\right]$$

$$R_0 - R'.M^{-1}.R = \begin{bmatrix} a - L_1 \dfrac{r_1^2}{u} - L_3(x + y)r^2_2; & b - L_3(x + y)r_2 r_3 \\ b - L_3(x + y)r_2 r_3 & ; \quad c - \dfrac{L_2 r_1^2}{u} - L_3(x + y)r_3^2 \end{bmatrix}$$

Let, e, f and g be the terms of the inverse of this matrix

$$\begin{bmatrix} e & g \\ g & f \end{bmatrix} = (R_0 - R'.M^{-1}.R)^{-1}$$

we get

$W_{33} = A^{-1}_3 + (x + y)^2(r^2_2 e + r^2_3 f + 2r_2 r_3 g)J_{2L_3}$

and the diagonal term h of W_{33} is

$h = x + (x + y)^2(r^2_2 e + r^2_3 f + 2r_2 r_3 g)$

The second term of h i.e $h' = h - x$ proves generally to be negligible in most practical cases.

For instance, if (unfavourable case) $L_1 = 20$; $L_2 = 25$; $L_3 = 5$; $\lambda = 19$; $n = 5$ to 50; h' varies between 1.5×10^{-4} and 6×10^{-4} i.e an effect on CD of 3×10^{-3} to 10^{-2}.

This term h' will therefore be neglected, hence

$$h = x$$

and

$$CD = \frac{n_2}{n_2 + \lambda}\left(1 - \frac{\lambda n_2}{n_2 m + \lambda N_2}\right)$$

The expression found is similar to that of the scheme based on one reference sire; the numbers of offspring m_2 and m_3 from the 2 reference sires act through their sum and the accuracy does not depend in many situations on the values of L_1, L_2 and L_3.