

Pondérations économiques des indices de sélection dans le cas des populations de race mixte ⁽¹⁾

J. M. ELSEN

*Station de génétique quantitative et appliquée,
Centre national de Recherches zootechniques, I.N.R.A.,
78350 Jouy-en-Josas*

Résumé

Dans le cas de populations où les descendants d'un même reproducteur ne sont pas tous destinés au même type de production (filles traites et descendants abattus d'un taureau de race mixte, par exemple), ce reproducteur doit être améliorateur pour chacun des objectifs. Le critère pour calculer les pondérations économiques proposé par SOLLER et coll. (1966) et Mc CLINTOCK et CUNNINGHAM (1974) est repris ici : l'indice de sélection optimum est celui qui maximise les recettes actualisées dues à la sélection, ce qui revient à donner d'autant plus de poids à un caractère qu'il sera exprimé plus tôt et par un plus grand nombre de descendants du reproducteur.

Un algorithme, basé sur le modèle de HILL (1974) et ELSEN et MOCQUOT (1974) est proposé pour estimer ces pondérations. L'influence de divers facteurs sur ces pondérations est étudiée : durée d'actualisation, structure de diffusion du progrès génétique, mortalités entre classes d'âge... On montre que l'expression algébrique des pondérations se simplifie considérablement quand les recettes sont calculées sur une durée infinie.

Introduction

La théorie classique des indices de sélection consiste à classer les candidats sur une combinaison linéaire de leurs valeurs génétiques estimées ($I = \sum a_i \bar{G}_i$), les coefficients de pondération a_i étant généralement définis comme l'accroissement marginal de la valeur H de l'animal pour un accroissement d'une unité de la variable G_i (HAZEL, 1943). Ceci suppose que la valeur H de l'animal soit unique ce qui implique que tous les descendants du reproducteur seront utilisés pour le même type de production.

Dans le cas des populations de vaches traites de race mixte, les reproducteurs

(¹) Cet article a fait l'objet d'une communication lors du séminaire sur « L'estimation de la valeur génétique des reproducteurs » organisé à Toulouse les 6 et 7 octobre 1976 par le Département de Génétique Animale de l'I.N.R.A. (France).

doivent transmettre d'une part une bonne aptitude laitière (H1) à leurs filles, d'autre part un bon potentiel de croissance (H2) à leurs descendants abattus (veaux mâles et génisses non retenues pour le renouvellement). SOLLER et coll. (1966) et McCLINTOCK et CUNNINGHAM (1974) proposent alors de choisir les reproducteurs sur l'indice I qui maximise les recettes dues à la sélection. Ils estiment ces dernières proportionnelles à $N_1 \cdot \Delta H_1 + N_2 \cdot \Delta H_2$ où N_1 et N_2 sont les nombres d'expressions par insémination des supériorités ΔH_1 et ΔH_2 des reproducteurs pour le lait et la viande. SOLLER et coll. (1966) évaluent ces nombres d'expressions à partir de l'ensemble des réponses directes et corrélées par insémination et tiennent compte, par le biais d'un coefficient d'actualisation, de l'intervalle de temps séparant chacune de ces réponses de la mise en place de l'insémination. McCLINTOCK et CUNNINGHAM (1974) négligent les réponses corrélées mais introduisent dans la pondération le phénomène de transmission du progrès génétique d'une génération à la suivante.

Cette solution, qui revient à définir la valeur d'un animal par la fonction $H = N_1 \cdot H_1 + N_2 \cdot H_2$, peut être généralisée à toute population où plusieurs objectifs de production sont poursuivis (lait, viande, production de laine par exemple) : on définira dans ce cas autant de valeurs génétiques $H_p(g_1, g_2, \dots, g_{n_p})$, fonction de composantes génétiques élémentaires, que d'objectifs de production et la valeur globale d'un animal sera $H = \sum_p N_p \cdot H_p$.

Nous voulons montrer ici comment le modèle de prévision développé par HILL (1974) et ELSÉN et MOCQUOT (1974) peut être utilisé pour rendre compte du phénomène de transmission du progrès génétique et, par là, pour déterminer de façon relativement simple les coefficients de pondération de l'indice.

I. — Méthode

Pour des raisons de cohérence avec le modèle de HILL (1974) et ELSÉN et MOCQUOT (1974) nous redéfinirons ici les nombres N_p comme les nombres totaux d'expressions des supériorités des reproducteurs retenus une année donnée. Les nombres d'expressions par insémination sont alors simplement N_p/RD où R est le nombre de candidats à la sélection retenus pour la reproduction cette année et D le nombre de descendants par reproducteur.

Dans ces conditions, les nombres N_p peuvent être estimés par la démarche que nous allons décrire.

a) La population est subdivisée en plusieurs cohortes homogènes vis-à-vis de la sélection : une cohorte est en particulier caractérisée par l'âge et le sexe des animaux la composant, ainsi que par le type de production auquel ils sont destinés (vaches en lactation, animaux en âge d'être abattus par exemple).

Les nombres N_p sont alors donnés par :

$$N_p = \sum_{t=0}^T \sum_{i=1}^C \text{Epit. Fpit.} \quad (1)$$

où i est l'indice de cohorte.

Epit le nombre d'animaux de la cohorte i qui réalisent, à la date t , le $p^{\text{ième}}$ type de production.

Fpit la part de la supériorité des reproducteurs sélectionnés que l'on retrouve chez les animaux de la cohorte i à la date t .

En notant $E\phi t = \begin{bmatrix} E\phi_{1t} \\ E\phi_{2t} \\ \vdots \\ \vdots \\ E\phi_{ct} \end{bmatrix}$ et $F\phi t = \begin{bmatrix} F\phi_{1t} \\ F\phi_{2t} \\ \vdots \\ \vdots \\ F\phi_{ct} \end{bmatrix}$, $N\phi$ devient :

$$N\phi = \sum_{t=0}^T {}^t r E\phi t \cdot F\phi t \tag{2}$$

${}^t r E\phi t$ étant le transposé de $E\phi t$.

On peut introduire un coefficient d'actualisation $\left(\frac{1}{1+r}\right)$, et :

$$N\phi = \sum_{t=0}^T \left(\frac{1}{1+r}\right)^t {}^t r E\phi t \cdot F\phi t \tag{3}$$

b) Le vecteur $F\phi t$ peut être estimé par récurrence puisque :

$$F\phi t = D \cdot F\phi_{t-1} \tag{4}$$

avec D une matrice de transition dont chaque élément d_{ij} représente la part moyenne du génotype des animaux appartenant à la cohorte i en t qui est identique au génotype des individus appartenant à la cohorte j en $t-1$.

En effet, si $F\phi_{t-1}$ est donné, la part $F\phi_{it}$ de la supériorité des reproducteurs sélectionnés l'année 0 que l'on retrouve chez les animaux de la cohorte i en t s'écrit :

$$F\phi_{it} = \sum_{j=1}^c d_{ij} F\phi_{jt-1} \tag{5}$$

On a alors $F\phi t = D^t F\phi_0$ (6)

et $F\phi_0$, vecteur initial des $F\phi_{io}$ est un vecteur dont tous les éléments sont nuls, sauf le terme correspondant à la cohorte des candidats à la sélection, qui est égal à 1.

Si on suppose que les effectifs $E\phi_{it}$ ne varient pas dans le temps, ce qui implique que la population soit en équilibre démographique, (3) devient :

$$N\phi = {}^t r E\phi \sum_{t=0}^T \left(\frac{1}{1+r}\right)^t D^t F\phi_0 \tag{7}$$

où, en notant $\Theta = \frac{1}{1+r}$

$$N\phi = {}^t r E\phi (I - \Theta D)^{T+1} (I - \Theta D)^{-1} F\phi_0 \tag{8}$$

et pour T infini

$$N\phi = {}^t r E\phi (I - \Theta D)^{-1} F\phi_0 \tag{9}$$

II. — Application

A titre d'illustration, nous allons montrer comment aborder, avec cette méthode, le problème étudié par McCLINTOCK et CUNNINGHAM (1974). Ces auteurs ont traité le cas d'une population de V vaches de race mixte dont une fraction K est croisée avec des mâles de race à viande. Les taureaux de la race mixte sont

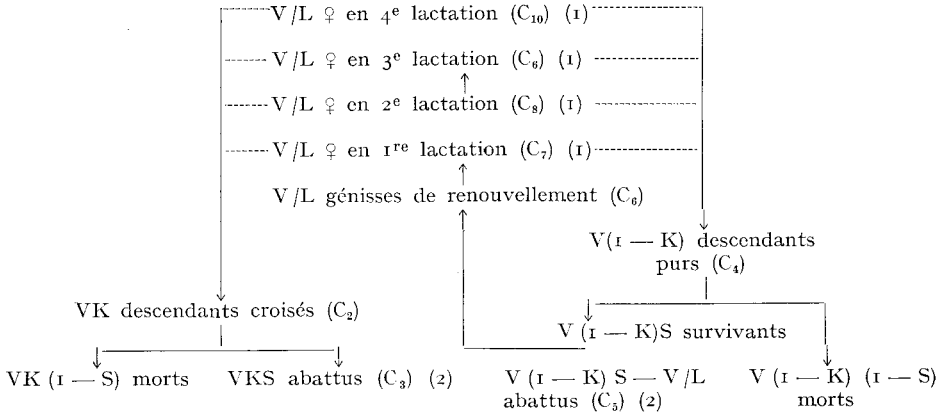


FIG. 1. — Effectifs des différentes cohortes C d'animaux.
Number of animals in the different groups C_i.

- (1) cohortes d'animaux exprimant le potentiel laitier.
- (2) cohortes d'animaux exprimant le potentiel en viande.
- V : nombre total de femelles traitées.
- L : nombre de lactations par femelle.
- S : taux de survie.
- K : taux de croisement.

	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈	C ₉	C ₁₀	
D =	0	0	0	0	0	0	0	0	0	0	C ₁
	0	0	0	0	0	1/2L	1/2L	1/2L	1/2L	0	C ₂
	0	1	0	0	0	0	0	0	0	0	C ₃
	1/2	0	0	0	0	1/2L	1/2L	1/2L	1/2L	0	C ₄
	0	0	0	1	0	0	0	0	0	0	C ₅
	0	0	0	1	0	0	0	0	0	0	C ₆
	0	0	0	0	0	1	0	0	0	0	C ₇
	0	0	0	0	0	0	1	0	0	0	C ₈
	0	0	0	0	0	0	0	1	0	0	C ₉
	0	0	0	0	0	0	0	0	1	0	C ₁₀

FIG. 2. — Matrice de transition D dans le cas où les vaches font quatre lactations.
Transition matrix when cows have four lactations (L = 4).

- Chaque élément d_{ij} de la matrice représente la part du génotype des individus de la cohorte C_i qui est identique à celui des individus de la cohorte C_j.
- C₁ reproducteurs mâles de race mixte.
 - C₂ descendants croisés de 0 à 1 an.
 - C₃ descendants croisés de 1 à 2 ans.
 - C₄ descendants purs de 0 à 1 an.
 - C₅ descendants purs abattus (1 à 2 ans).
 - C₆ femelles pures retenues pour le renouvellement (1 à 2 ans).
 - C₇ femelles traitées de 2-3-ans.
 - C₈ femelles traitées de 3-4 ans.
 - C₉ femelles traitées de 4-5 ans.
 - C₁₀ femelles traitées de 5-6 ans.

sélectionnés à la fois sur leurs potentiels en lait et en viande exprimés respectivement par leurs descendantes traites au cours de L lactations et leurs descendants abattus, purs ou croisés (fig. 1). McCLINTOCK et CUNNINGHAM (1974) ont négligé dans l'estimation des nombres d'expressions $N\phi$ (ici N_1 et N_2), la transmission de la supériorité des reproducteurs par les voies pères-fils et mères-fils. Dans ces conditions, si on suppose que ces mâles ne sont utilisés pour la reproduction qu'une année, la matrice D prend la forme simple décrite figure 2 dans le cas où $L = 4$ et les vecteurs F_{10} et F_{20} sont les vecteurs d'éléments $F\phi_{10} = 1$ et $F\phi_{i0} = 0$ si $i \neq 1$. En fait, les auteurs raisonnaient en gain par insémination de race pure, ce qui revient à diviser F_{10} et F_{20} par $V(1 - K)$.

Enfin, les effectifs tE_1 et tE_2 sont, d'après la figure 1 :

$${}^tE_1 = (0, 0, 0, 0, 0, 0, V/L, V/L, V/L, V/L)$$

$${}^tE_2 = (0, 0, VKS, 0, V(1 - K)S - V/L, 0, 0, 0, 0, 0).$$

Il est alors possible de calculer les nombres d'expressions N_1 et N_2 , et on retrouve les résultats obtenus par McCLINTOCK et CUNNINGHAM (1974) (tableau I).

TABLEAU I

Évolution des nombres d'expressions annuels et cumulés avec le temps
Evolution with time of the numbers of expressions per year and cumulated

Année	N_{1t}	N_1	N_{2t}	N_2
1	0	0	0	0
2	0	0	0	0
3	0	0	$(S - \frac{1}{L(1-K)}) / 2$	0,225
4	$\frac{1}{2} \times \frac{1}{L(1-K)}$	0,085	0	0,225
5	$\frac{1}{2} \times \frac{1}{L(1-K)}$	0,163	$\frac{SL-1}{L(1-K)} \times \frac{1}{4L}$	0,249
6	$(\frac{1}{4L} + \frac{1}{2}) \times \frac{1}{L(1-K)}$	0,242	$\frac{SL-1}{L(1-K)} \times \frac{1}{4L}$	0,270
7	$(\frac{1}{2} + \frac{1}{2L}) \times \frac{1}{L(1-K)}$	0,322	$\frac{SL-1}{L(1-K)} \times (\frac{1}{8L^2} + \frac{1}{4L})$	0,291
8	$(\frac{1}{8L^2} + \frac{3}{4L}) \times \frac{1}{L(1-K)}$	0,345	$\frac{SL-1}{L(1-K)} \times (\frac{1}{4L^2} + \frac{1}{4L})$	0,313
9	$(\frac{3}{8L^2} + \frac{1}{L}) \times \frac{1}{L(1-K)}$	0,374	$\frac{SL-1}{L(1-K)} \times (\frac{1}{10L^3} + \frac{3}{8L^2})$	0,320
10	$(\frac{1}{10L^3} + \frac{3}{4L^2} + \frac{3}{4L}) \times \frac{1}{L(1-K)}$	0,397	$\frac{SL-1}{L(1-K)} \times (\frac{3}{10L^3} + \frac{1}{2L^2})$	0,328

N_{1t}, N_{2t} : nombres d'expressions du génotype lait et viande l'année t .

N_1, N_2 : coefficients de pondération du lait et de la viande ($N_1 = \sum_t \theta^t N_{1t}$ et $N_2 = \sum_t \theta^t N_{2t}$).

$\theta = \frac{1}{1+r}$ est le coefficient d'actualisation.

L : nombre de lactations par femelle ($L = 4$).

K : taux de croisement ($K = 0$).

S : taux de survie ($S = 0,85$).

r : taux d'actualisation ($r = 0,10$).

III. — Intérêt de la méthode

La méthode qui vient d'être proposée, présente deux avantages par rapport à la démarche adoptée par McCLINTOCK et CUNNINGHAM (1974) : possibilités de tenir compte d'un plus grand nombre de facteurs et simplicité des calculs.

a) *Prise en compte de facteurs nouveaux*

α) On peut en premier lieu intégrer dans le modèle, sans difficultés supplémentaires, les voies de transmission pères-fils et mères-fils négligées par McCLINTOCK et CUNNINGHAM (1974).

Dans ce cas, la matrice de transition D peut être décomposée en quatre sous matrices.

$$D = \begin{array}{cc|cc} & \text{Pères} & & \text{Mères} & \\ \hline & D_1 & & D_2 & \text{Fils} \\ \hline & D_3 & & D_4 & \text{Filles} \end{array}$$

Chacune de ces sous matrices correspondant à une des voies. Il n'y a alors aucune raison pour que les nombres d'expressions N_i calculés dans ce cas soient les mêmes que ceux qui avaient été obtenus précédemment (fig. 3), le fait de négli-

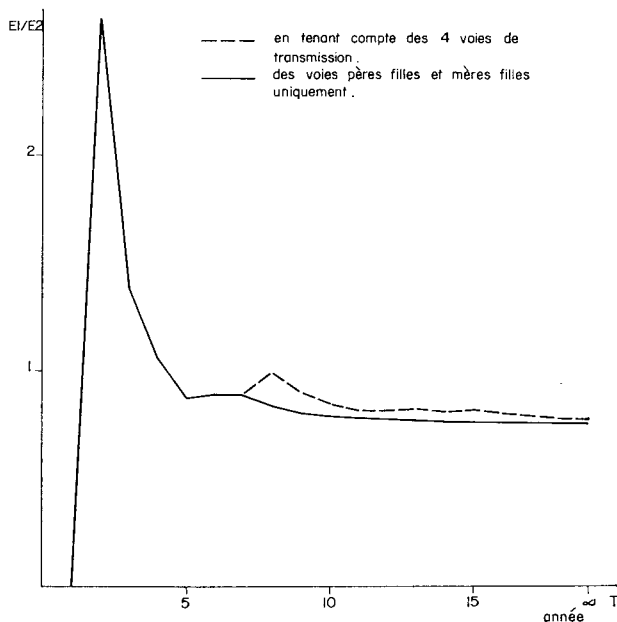


FIG. 3. — *Rapport des pondérations lait/viande en fonction de la durée d'actualisation*
($L = 4, k = 0, S = 0,85, r = 0,10$).

Ratio of milk/meat weightings in function of the length of the accounting period.

ger les voies pères-fils et mères-fils revenant à utiliser comme matrice de transition D' telle que :

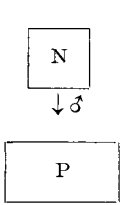
$$D' = \left[\begin{array}{c|c} 0 & 0 \\ \hline D_3 & D_4 \end{array} \right]$$

On peut cependant montrer, dans un cas assez général, que le rapport des pondérations N_1/N_2 , calculé à l'infini, est le même dans les deux cas (Annexe 1).

β) On peut également tenir compte des taux d'élimination (mortalité et réforme) qui interviennent inévitablement d'une classe d'âge à la suivante : il suffit pour cela de remplacer les éléments $1/2 L$ des lignes de la matrice D correspondant aux animaux âgés de 0 à 1 an par $pi/2$ où pi est la probabilité qu'un de ces animaux soit fils d'une mère d'âge i .

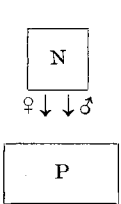
γ) Enfin, cette méthode permet de traiter le cas de populations de structure plus complexe. A titre d'exemple, nous avons comparé deux situations de diffusion du progrès génétique chez le porc. Il s'agit ici d'une sélection sur la prolificité de la mère (H1) et la valeur individuelle en croissance et adiposité de la carcasse (H2). La sélection est pratiquée dans un noyau (N) et le progrès génétique ainsi obtenu est transmis à un étage de production (P.). Cette transmission se fait par l'utilisation, comme reproducteurs de l'étage P, de mâles (schéma 1), ou de

Schéma 1



N					P					
0	1/2	0	1/4	1/4	0	0	0	0	0	N
1	0	0	0	0	0	0	0	0	0	
0	1/2	0	1/4	1/4	0	0	0	0	0	
0	0	1	0	0	0	0	0	0	0	
0	0	0	1	0	0	0	0	0	0	P
1	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	1/4	1/4	0	1/6	1/6	
0	0	0	0	0	0	0	1	0	0	
0	0	0	0	0	0	0	0	1	0	
0	0	0	0	0	0	0	0	0	1	
0	0	0	0	0	0	0	0	0	0	

Schéma 2.



N					P					
0	1/2	0	1/4	1/4	0	0	0	0	0	N
1	0	0	0	0	0	0	0	0	0	
0	1/2	0	1/4	1/4	0	0	0	0	0	
0	0	1	0	0	0	0	0	0	0	
0	0	0	1	0	0	0	0	0	0	P
1	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	1	0	0	0	0	
0	0	1	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	1	0	0	
0	0	0	0	0	0	0	0	1	0	
0	0	0	0	0	1/4	1/4	1/6	1/6	1/6	
0	0	0	0	0	0	0	0	0	0	

FIG. 4. — Matrices de transition correspondant aux deux schémas de sélection comparés.

Transition matrix for the two compared selection schemes.

N : noyau de sélection; P : étage de production.

mâles et de femelles (schéma 2) fils des reproducteurs sélectionnés (fig. 4). On a supposé que les candidats sont sélectionnés à 6 mois et que, dans le noyau, les mâles font une période (6 mois) de saillie (2 dans l'étage de production) et les femelles deux mises bas (3 dans P.). Les matrices de transition correspondant aux deux schémas sont données figure 4.

On peut alors estimer les rapports $(N_1/N_2) \sigma$ et $(N_1/N_2) \varphi$ des nombres d'expressions des supériorités ΔH_1 et ΔH_2 des mâles et des femelles sélectionnés dans les noyaux. L'essentiel des animaux appartenant à l'étage de production, on a négligé, dans l'estimation de ces rapports, la part due au noyau de sélection.

On constate alors (fig. 5) que ces rapports varient considérablement avec le schéma et la durée d'actualisation, et dans une moindre mesure, le sexe des candidats à la sélection. On notera en particulier que l'utilisation de femelles nées dans le noyau comme reproductrices de l'étage de production permet une mise en valeur beaucoup plus rapide de l'effort de sélection sur la prolificité et favorise donc ce caractère par rapport à la croissance.

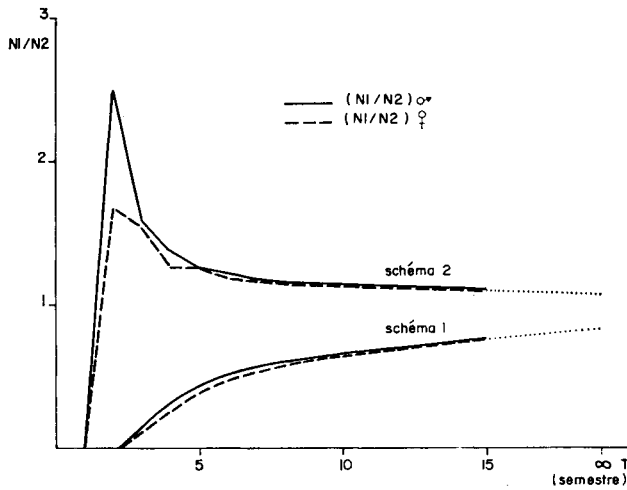


FIG. 5. — Rapports (N_1/N_2) des nombres d'expressions de la supériorité génétique des reproducteurs pour la prolificité (ΔH_1) et pour la valeur en croissance et adiposité de la carcasse (ΔH_2).

(Ratios (N_1/N_2) of the numbers of expressions of the genetic superiority of breeding animals for prolificacy (ΔH_1) and for growth and carcass fatness (ΔH_2 .)

b) Simplicité des calculs

Elle est liée à l'emploi de la notation matricielle et est d'autant plus évidente que le nombre de cohortes et la durée d'actualisation sont plus grands. Pour une durée infinie, on est d'ailleurs ramené à une seule inversion de matrice (relation 9).

On peut même montrer (Annexe 2) que, dans certains cas, le rapport N_1/N_2 s'écrit simplement, pour une durée infinie :

$$N_1 = \frac{\sum_{i=1}^c E_{1i} \Theta^{a(i)}}{\sum_{i=1}^c E_{2i} \Theta^{a(i)}}$$

où $a(i)$ est l'âge des animaux de la cohorte i .

Ce rapport est celui des nombres « actualisés » d'animaux exprimant les deux caractères dans la population à un instant donné.

L'estimation des nombres $N\phi$ pour T infini semble d'ailleurs une bonne approximation de leur valeur pour une durée finie, au moins si le taux et la durée d'actualisation ne sont pas trop faibles (CUNNINGHAM et JOAN RYAN, 1975).

Discussion — Conclusion

Le calcul des nombres d'expressions N_i tel qu'il a été proposé ici a pour but d'estimer les recettes dues à la sélection dont la maximisation est proposée par SOLLER et coll. (1966) et McCLINTOCK et CUNNINGHAM (1974) comme le critère de choix de l'indice de sélection des taureaux de race mixte. Deux remarques nous semblent cependant fondamentales. D'une part les recettes ne sont pas toujours simplement linéaires en $H\phi$: elles seront quadratiques si un des groupes de caractères est la productivité numérique — ce que nous avons d'ailleurs négligé plus haut —, elles prendront une forme algébrique plus complexe si la sélection se fait en deux étapes... D'autre part, on retrouve ici les problèmes classiques posés par le choix des pondérations des indices. Ainsi, le critère sous-jacent (maximisation des recettes dues à la sélection) ne prend pas en compte le problème de la répartition des fruits de l'effort de sélection entre les agents économiques concernés par celle-ci (sélectionneurs — producteurs — consommateurs) ce qui revient à supposer une répartition équitable entre eux. De même les prix de certains des produits dont on veut améliorer l'efficacité de la production peuvent être artificiels et la maximisation des recettes ne conduira pas nécessairement à la situation optimale vis-à-vis de l'équilibre entre quantités de produits offerts.

Enfin, de même que les poids économiques calculés classiquement, les valeurs ainsi obtenues seront toujours approximatives. En particulier, l'hypothèse d'indépendance entre la valeur économique d'une unité d'un caractère et la valeur moyenne de celui-ci n'est plus justifiée dans ce cas puisque les recettes sont estimées sur plusieurs années après le choix des reproducteurs. De même, il est évident que du fait des aléas tant génétiques que démographiques et économiques, la prévision est très incertaine pour les années les plus éloignées. Quoiqu'il en soit, la méthode proposée est certainement plus précise qu'une estimation intuitive des pondérations et a le mérite de leur donner une forme mathématique simple. Elle peut être un des éléments du choix des reproducteurs.

Reçu pour publication en février 1977.

Remerciements

Nous remercions J. C. Mocquot pour la lecture critique du manuscrit.

Summary

Economic weighting of selection indexes for mixed breed populations

In populations where the progeny of each animal are not assigned to the same type of production (e.g. milked daughters and slaughtered progeny for the sire of a dual purpose breed), animals selected for breeding must improve all types of production. The criterion suggested by SOLLER *and al.* (1966) and Mc CLINTOCK and CUNNINGHAM (1974) has been used in this study: the optimum selection index is the index maximizing the discounted economic return from

selection, i.e. the sooner a trait is expressed and by a larger number of progeny, the more weight it receives.

An algorithm, based on the model proposed by HILL (1974) and ELSÉN and MOCQUOT (1974) is used to estimate these weightings. The influence of several factors on these weightings is investigated: length of the accounting period, breeding structure, mortalities for each age class... It is shown that the algebraic formula that gives the weightings is much simpler when returns are computed for an infinite length of time.

Références bibliographiques

- CUNNINGHAM E. P., JOAN RYAN, 1975. A note on the effect of the discount rate and length of the accounting period on the economic value of genetic improvement in cattle populations. *Anim. Prod.*, **21**, 77-80.
- ELSÉN J. M., MOCQUOT J. C., 1974. Méthode de prévision de l'évolution du niveau génétique d'une population soumise à une opération de sélection et dont les générations se chevauchent. *Bull. techn. Dép. Génét. anim.* (Inst. nat. Rech. agron. Fr.), **17**, 30-54.
- HAZEL L. N., 1943. The genetic basis for constructing selection indexes. *Genetics*, **28**, 476-490.
- HILL W. G., 1974. Prediction and evaluation of response to selection with overlapping generations. *Anim. Prod.*, **18**, 117-140.
- MC CLINTOCK A. E., CUNNINGHAM E. P., 1974. Selection in dual purpose cattle populations: defining the breeding objective. *Anim. Prod.*, **18**, 237-247.
- SOLLER M., BAR ANAN R., PASTERNAK M., 1966. Selection of dairy cattle for growth rate and milk production. *Anim. Prod.*, **8**, 109-119.

Annexe I

Nous allons montrer ici que, dans un cas assez général, il est équivalent de tenir compte ou non des voies pères-fils et mères-fils dans l'estimation du rapport des nombre d'expression N_1/N_2 calculé pour une durée d'actualisation infinie.

Ce cas est celui où la matrice D est de la forme

$$D = \left[\begin{array}{ccc|ccc} d_{11} & d_{12} & \dots & d_{1n} & d_{1n+1} & \dots & d_{1n+m} \\ & \text{I} & & & & & \\ & & & \text{I} & & & \\ \hline & & & & d_{n+1,n+1} & & d_{n+1,n+m} \\ d_{n+1,1} & d_{n+1,2} & & d_{n+1,n} & \text{I} & & \\ & & & & & & \text{I} \end{array} \right] = \begin{bmatrix} D_1 & D_2 \\ D_3 & D_4 \end{bmatrix}$$

Les lignes 1 à n et n + 1 à n + m correspondant respectivement aux voies pères-descendants et mères-descendants; les colonnes 1 à n et n + 1 à n + m, respectivement aux voies parents-fils et parents-filles.

D est telle que :

$$\sum_{j=1}^n d_{1j} = \sum_{j=n+1}^{n+m} d_{1j} = \sum_{j=1}^n d_{n+1,j} = \sum_{j=n+1}^{n+m} d_{n+1,j} = I/2$$

$$d_{i,i-1} = I \text{ si } i \neq 1 \text{ et } n+1$$

$$d_{ij} = 0 \text{ sinon}$$

Comme nous l'avons dit plus haut, négliger les voies parents-fils revient à utiliser comme matrice de transition la matrice :

$$D' = \begin{bmatrix} 0 & 0 \\ D_3 & D_4 \end{bmatrix}$$

Pour une durée d'actualisation infinie, les nombres d'expressions s'écrivent selon que l'on tient compte ou non des voies parents-fils.

$$Np = {}^t r E p (I - \Theta D)^{-1} F p_0$$

$$N'p = {}^t r E' p (I - \Theta D')^{-1} F' p_0$$

où $E p$ est un vecteur dont les n premiers éléments sont nuls.

$E p = (0, E' p)$, $E' p$ étant de dimension m .

$F p_0$ est un vecteur dont le 1^{er} élément, correspondant à la cohorte de animaux sélectionnés, est égal à 1, tous les autres étant nuls :

$$F p_0 = (F' p_0, 0), F' p_0 \text{ étant de dimension } n, F' p_0 = (0, 0 \dots 1 \dots 0).$$

On peut inverser les matrices $I - \Theta D$ et $I - \Theta D'$ par blocs.

$$(I - \Theta D')^{-1} = \left[\begin{array}{c|c} I_1 & 0 \\ \hline (I_4 - D_4)^{-1} \Theta D_3 & (I_4 - \Theta D_4)^{-1} \end{array} \right]$$

$$(I - \Theta D)^{-1} = \left[\begin{array}{c|c} I_1 - \Theta D_1 - \Theta D_2 (I_4 - \Theta D_4)^{-1} \Theta D_3^{-1} = S_1 & \\ \hline (I_4 - \Theta D_4)^{-1} \Theta D_3 \cdot S_1 = S_3 & \\ \hline & (I_1 - \Theta D_1)^{-1} \Theta D_2 S_4 = S_2 \\ \hline & I_4 - \Theta D_4 - \Theta D_3 (I_1 - \Theta D_1)^{-1} \Theta D_2^{-1} = S_4 \end{array} \right]$$

où I_1 et I_4 sont les matrices unités de mêmes dimensions que D_1 et D_4 .

On est donc ramené à comparer

$$\frac{N_1'}{N_2'} = \frac{{}^t r E' (I_4 - \Theta D_4)^{-1} \Theta D_3 F'_{10}}{{}^t r E' (I_4 - \Theta D_4)^{-1} \Theta D_3 F'_{20}}$$

et

$$\frac{N_1}{N_2} = \frac{{}^t r E (I_4 - \Theta D_4)^{-1} \Theta D_3 S_1 F_{10}}{{}^t r E (I_4 - \Theta D_4)^{-1} \Theta D_3 S_1 F_{20}}$$

en notant que, dans ce cas, $F'_{10} = F'_{20}$

Si on appelle c_{ij} les éléments de $(I_4 - \Theta D_4)^{-1}$, ∂_{1j} ceux de D_3 et s_{ij} de S_1 , la matrice $(I_4 - \Theta D_4)^{-1} D_3$ est alors de la forme :

$$\left[\begin{array}{ccc} c_{11} \cdot \partial_{11} & c_{11} \cdot \partial_{12} & \dots \cdot c_{11} \cdot \partial_{1n} \\ c_{21} \cdot \partial_{11} & c_{21} \cdot \partial_{12} & \dots \cdot c_{21} \cdot \partial_{1n} \\ \vdots & \vdots & \vdots \\ c_{m1} \cdot \partial_{11} & c_{m1} \cdot \partial_{12} & \dots \cdot c_{m1} \cdot \partial_{1n} \end{array} \right]$$

et la matrice $(I_4 - \Theta D_4)^{-1} D_3 S_1$:

$$\left[\begin{array}{ccc} c_{11} \sum \partial_{1k} s_{k1} & c_{11} \sum \partial_{1k} s_{k2} & \dots \cdot c_{11} \sum \partial_{1k} s_{kn} \\ c_{21} \sum \partial_{1k} s_{k1} & c_{21} \sum \partial_{1k} s_{k2} & \dots \cdot c_{21} \sum \partial_{1k} s_{kn} \\ \vdots & \vdots & \vdots \\ c_{m1} \sum \partial_{1k} s_{k1} & c_{m1} \sum \partial_{1k} s_{k2} & \dots \cdot c_{m1} \sum \partial_{1k} s_{kn} \end{array} \right]$$

On trouve alors que :

$$N\phi = \left(\sum_{i=1}^m E'_{pi} \cdot c_{i1} \right) \left(\sum_{k=1}^m \partial_{1k} \cdot s_{kl} \cdot \Theta \right)$$

et

$$N'\phi = \left(\sum_{i=1}^m E'_{pi} \cdot c_{i1} \right) \cdot \partial_{1l}$$

Donc $N_1/N_2 = N'_1/N'_2$ et seuls les éléments de $(I_4 - \Theta D_4)^{-1}$ interviennent dans le rapport des nombres d'expressions.

Annexe II

Toujours dans le cas de la matrice D envisagée dans l'annexe I, on peut montrer que ce rapport N_1/N_2 est indépendant des éléments de cette matrice.

La matrice $(I_4 - \Theta D_4)$ s'écrit :

$$I_4 - \Theta D_4 = \begin{bmatrix} I - \Theta d_{n+1, n+1} - \Theta d_{n+1, n+2} \dots \dots - \Theta d_{n+1, n+m-1} - \Theta d_{n+1, n+m} & & & & \\ -\Theta & I & & & 0 & 0 \\ 0 & -\Theta & & & 0 & 0 \\ \cdot & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ 0 & 0 & & -\Theta & & I \end{bmatrix}$$

On peut l'inverser par blocs :

$$(I_4 - \Theta D_4) = \begin{bmatrix} E & G \\ F & H \end{bmatrix} \text{ où } E \text{ est l'élément } I - \Theta d_{n+1, n+1}$$

$$(I_4 - \Theta D_4)^{-1} = \begin{bmatrix} S & U \\ T & V \end{bmatrix} \text{ avec}$$

$$S = (E - G \cdot H^{-1} F)^{-1} \text{ et}$$

$$T = -H^{-1} F (E - G H^{-1} F)^{-1} = H^{-1} F S.$$

Il suffit en effet de connaître les valeurs de S et T puisque seule la première colonne de $(I_4 - \Theta D_4)^{-1}$ intervient dans le rapport N_1/N_2 .

La matrice H qui s'écrit :

$$H = \begin{bmatrix} I & 0 & 0 & \dots & 0 & 0 \\ -\Theta & I & 0 & \dots & 0 & 0 \\ 0 & -\Theta & I & \dots & 0 & 0 \\ \cdot & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ 0 & 0 & 0 & -\Theta & 0 & \end{bmatrix}$$

a pour inverse

$$H^{-1} = \begin{bmatrix} I & 0 & 0 & \dots & 0 & 0 \\ \Theta & I & 0 & \dots & 0 & 0 \\ \Theta^2 & \Theta & I & \dots & 0 & 0 \\ \Theta^{m-2} & \Theta^{m-3} & \Theta^{m-4} & \dots & \Theta & I \end{bmatrix}$$

Dès lors s s'écrit :

$$S = [I - \Theta d_{n+1, n+1} - (-\Theta d_{n+1, n+2}, \dots, -\Theta d_{n+1, m+m}) \begin{bmatrix} -\Theta \\ -\Theta^2 \\ \vdots \\ -\Theta^{m-1} \end{bmatrix}]^{-1}$$

$$S = \frac{I}{I - \sum_{j=1}^m j d_{n+1, n+j}} = \frac{I}{\Delta}$$

On en déduit T :

$$T = \begin{bmatrix} \Theta \\ \Delta \\ \Theta^2 \\ \Delta \\ \vdots \\ \Theta^{m-1} \\ \Delta \end{bmatrix}$$

Les termes c_{j1} de la première colonne de $(I_4 - \Theta D_4)^{-1}$ sont donc :

$$c_{j1} = \frac{\Theta^{j-1}}{\Delta}, \text{ et le rapport } \frac{N_I}{N_2} = \frac{\sum_{j=1}^m E'_{1j} c_{j1}}{\sum_{j=1}^m E'_{2j} c_{j1}}$$

et

$$\frac{N_I}{N_2} = \frac{\sum_{j=1}^m \Theta^{j-1} E'_{1j}}{\sum_{j=1}^m \Theta^{j-1} E'_{2j}}$$

qui est le rapport des nombres actualisés pour leur âge d'animaux réalisant chacun des types de production à un instant donné.

