

Estimation de la valeur génétique des reproducteurs dans les cas d'incertitude sur les apparentements

II - Utilisation des durées de gestation pour le calcul des probabilités de filiation

J.M. ELSÉN, J.P. POIVEY

Station d'Amélioration Génétique des Animaux, I.N.R.A.,
B.P. 27, F 31326 Castanet-Tolosan Cedex

Résumé

Pour tenir compte des incertitudes sur les paternités dans l'expression des indices de sélection, il faut d'abord estimer la loi *a priori* (avant mesure des phénotypes) des probabilités de filiations. Si la distribution des durées de gestation est connue, le contrôle des périodes de monte des reproducteurs et des dates de naissance des produits permet cette estimation. Nous avons étudié cette possibilité dans le contexte de l'élevage bovin tropical où la reproduction s'opère sur une période étendue de l'année. Dans ce cas, on peut préconiser un système de reproduction se caractérisant par une succession de périodes de monte de durée fixée d , où un seul mâle est présent dans le troupeau, séparées les unes des autres par des intervalles de durée r . Parallèlement, on aura une succession de périodes de contrôle de naissance de durée a , séparées par des intervalles de durée s , telles que tous les individus nés au cours d'une même période seront supposés demi-frères. Le choix des valeurs, d , r , a et s pour une maximisation du progrès génétique est décrit. On montre, sur un exemple, que ce choix peut s'opérer plus facilement sur le coefficient de corrélation moyen entre la valeur génétique du père et la moyenne des valeurs phénotypiques de ses ascendants présumés que sur l'espérance du progrès génétique.

Mots clés : Indexation, erreur de filiation, durée de gestation.

Summary

*Prediction of genetic merit in the case of uncertain paternity
II - Using gestation length for computing probabilities of descent*

In order to take into account errors about paternity in the expression of selection indexes, the *a priori* distribution (before obtaining phenotypic values) of probabilities of descent must be estimated. If the distribution of gestation length is known, the recording of mating periods of sires and of birth dates allows this estimation. We have studied this possibility in the tropical cattle breeding context where reproduction goes on over a long period in the year. In this case, we

propose a reproduction system consisting of a series of mating periods having fixed length, d say, during which only one male is used in the herd, separated from each other by intervals of length r . There will be a corresponding series of periods of birth recording of length a separated by intervals of length s , such that individuals born during the same period are supposed to be half-sibs. The choice of d , r , a and s so as to maximise genetic progress is described. It is seen, in an example, that this choice may be more easily made on the averaged correlation coefficient between the genetic value of the sire and the mean phenotypic value of his supposed progeny than on expected genetic gain.

Key words : Selection index, uncertain paternity, gestation length.

I. Introduction

Il est fréquent, dans les populations où la reproduction se fait en monte naturelle, que les accouplements ne soient pas tous connus et qu'il y ait donc une incertitude sur les paternités. Dans ces situations, il faut rechercher, pour estimer la valeur génétique des candidats à la sélection, des critères tenant compte des différents liens de parenté possibles entre ces candidats et les individus qui sont supposés leur être apparentés et dont les phénotypes sont utilisés comme prédicteurs.

Dans une première partie (POIVEY & ELSEN, 1984), nous avons étudié la forme algébrique de tels indices de sélection. On a pu montrer que l'espérance conditionnelle, qui est le meilleur critère au sens de la maximisation du progrès génétique, est en général d'expression très lourde alors que la régression linéaire, beaucoup plus simple à utiliser, ne possède pas les propriétés optimales du point de vue du gain génétique. Cependant, quelques simulations d'opérations de sélection nous ont indiqué que l'utilisation de ces 2 critères (espérance conditionnelle et régression linéaire) donnent des résultats très peu différents, le progrès génétique obtenu avec la régression linéaire n'étant jamais inférieur à 98 p. 100 du gain maximum. Bien que ces simulations n'aient porté que sur quelques exemples, les résultats nous semblent suffisamment convergents pour que nous nous limitons ici à l'utilisation de la régression linéaire comme critère de sélection.

Pour appliquer de tels indices, il faut connaître la loi *a priori* (avant mesure des phénotypes) des probabilités de filiations. Une méthode simple consiste à contrôler les périodes de monte des reproducteurs et observer ensuite les dates de naissance des produits. Les probabilités de filiation s'en déduisent, connaissant la distribution des durées de gestation. Cette méthode peut ainsi être mise en place dans les élevages ovins où on a une lutte principale des brebis avec un ou des béliers identifiés puis une repasse où les accouplements ne sont pas contrôlés. On pourra dans ce cas attribuer, avec une probabilité connue, les paternités aux(x) bélier(s) de la lutte principale.

Plus généralement, on peut préconiser cette méthode dans les situations où la reproduction s'étale dans le temps en une succession de périodes telles qu'il n'y a aucune ambiguïté sur les accouplements au cours d'une même période. Nous l'envisagerons à titre d'illustration pour les bovins élevés en milieu tropical. Nous montrons sur cet exemple comment concevoir la mise en place de ce procédé, en particulier dans le choix des dates extrêmes d'observation des naissances entre lesquelles tout individu né est supposé être fils du même mâle et éventuellement de la durée des intervalles de temps séparant 2 périodes successives de monte.

II. Définitions et hypothèses démographiques

A. Description du protocole d'accouplements

On se place dans la situation d'un ranch où une partie des femelles est dissociée en L lots de F vaches. Au cours de l'année chacun de ces lots est mis à la reproduction et contrôlé selon le calendrier suivant (fig. 1) : M mâles lui sont attribués successivement et de façon non chevauchante (il y a donc $L.M$ taureaux au total). La période de monte s'étale en tout sur D jours. Le premier mâle (A_1) (*) est présent des jours 1 à d_1^r , le deuxième (A_2), des jours d_2^r à d_2^r , etc. Pour des raisons pratiques, on ne considérera que les cas où les intervalles de temps du protocole sont fixes et indépendants des mâles : $d_j^r - d_{j-1}^r + 1 = d$ et $d_{j+1}^r - d_j^r - 1 = r$ (pour $j = 1$ à M).

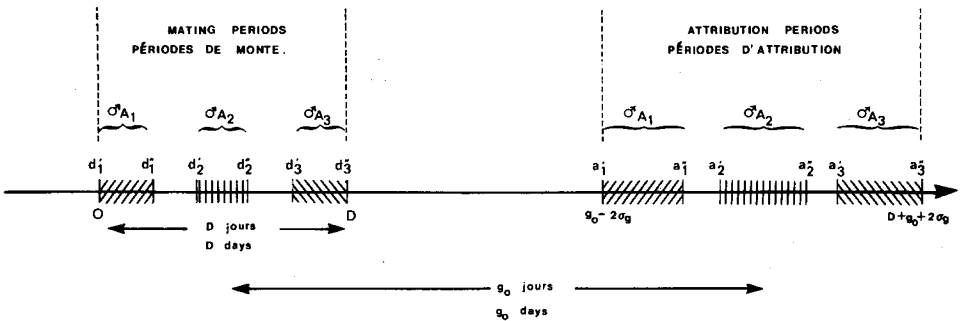


FIG. 1

Calendrier de monte et d'attribution des descendants.

Planning of the mating and attribution periods.

On connaît la durée de gestation moyenne g_0 et on décide de considérer comme fils de A_j tout individu né autour de la date $g_0 + (d_j^r + d_{j+1}^r)/2 = d_j^r + g_0 + d/2$, soit entre les jours a_j^i et a_j^r (inclus). En fait, la durée de gestation est aléatoire, et il est possible que certains animaux conçus avant la date d_j^r ou après d_j^r naissent au cours de cette période d'attribution $[a_j^i, a_j^r]$.

A l'inverse, certains des descendants du mâle A_j naîtront avant a_j^i ou après a_j^r . En fait, on s'est toujours placé dans des cas simples en supposant :

- que les périodes sans attribution (de a_j^r à a_{j+1}^i) sont d'égales durées ($a_{j+1}^i - a_j^r - 1 = s$)
- que $a_1^i = g_0 - 2\sigma_g$ où σ_g est l'écart-type de la durée de gestation
- que $a_M^r = D + g_0 + 2\sigma_g$
- que $a_j^r - a_j^i = 1 = a$ pour $j = 1$ et M .

(*) Dans la suite de l'exposé, sauf indication contraire, on ne mentionne pas l'indice 1 du lot. En toute rigueur, il faudrait par exemple écrire A_{11} et non A_1 , et en général A_{j1} , a_{j1} , d_{j1} , etc.

B. Probabilités des événements observés

Si on considère une des F femelles mises à la reproduction, 2 événements (f_i et m_j) la caractérisent : elle a été fécondée par $i^{\text{ème}}$ mâle (f_i), son descendant est attribué au $j^{\text{ème}}$ mâle (m_j). On remarque que, pour que les f_i forment un système complet d'événements, il faut leur adjoindre f_0 : « la femelle reste vide ».

On définit par ailleurs :

q_{ij} la probabilité conjointe de f_i et m_j

q_{ij} la probabilité de f_i conditionnellement à m_j ,

telle que $q_{vj} = \frac{q_{ij}}{\sum_i q_{ij}}$

q_0 la probabilité de f_0 .

Soit n_j le nombre de descendants nés entre a_j^i et a_j^v , donc attribués au mâle A_j . Nous notons S_{k/n_j} l'événement constitué par la réalisation des apparentements entre le mâle A_j et ces n_j descendants. Cet événement, qui est conditionnel au nombre n_j , sera qualifié de situation de parenté conformément à POIVEY & ELSEN (1984).

La loi de distribution de la situation s'écrit :

$$\text{Prob}(S_{k/n_j}) = \text{Prob}(n_{1j}, n_{2j}, \dots, n_{Mj}/n_j) = \frac{n_j!}{n_{1j}! \cdot n_{2j}! \cdot \dots \cdot n_{Mj}!} \cdot q_{1j}^{n_{1j}} \cdot q_{2j}^{n_{2j}} \cdot \dots \cdot q_{Mj}^{n_{Mj}}$$

n_{ij} étant le nombre de femelles, dans S_{k/n_j} , qui sont fécondées par A_i et dont les descendants sont attribués à A_j ($\sum_i n_{ij} = n_j$). Les probabilités élémentaires q_{ij} s'écrivent :

$$q_{ij} = \sum_{t_i = d_i}^{d_i'} \text{prob}(f(t_i)) \cdot \sum_{t_j = a_j}^{a_j^v} \text{prob}(m(t_j)/f(t_i))$$

où $f(t_i)$ et $m(t_j)$ sont les événements : la femelle a été fécondée à la date t_i et elle a mis bas à la date t_j .

• $\text{Prob}(m(t_j)/f(t_i))$ dépend de la loi de distribution de la durée de gestation. Nous l'avons supposée normale $N(g_0, \sigma_g)$. Ceci implique la symétrie de la distribution, hypothèse que nous avons retenue, bien qu'elle ne soit pas rigoureusement vérifiée.

Dans ces conditions on a :

$$\sum_{t_j = a_j}^{a_j^v} \text{prob}(m(t_j)/f(t_i)) = \pi \left(\frac{a_j^v - t_i - g_0}{\sigma_g} \right) - \pi \left(\frac{a_j^i - t_i - g_0}{\sigma_g} \right)$$

• $\text{Prob}(f(t_i))$ dépend :

- de la loi de la $r^{\text{ème}}$ venue en chaleur au cours de la période de monte. Soit $Q_r(t)$ la probabilité de cet événement à la date t ;
- du calendrier de monte, tel que $\epsilon(t) = 1$ si un mâle est présent en t , 0 sinon ;
- de la fertilité instantanée. Soit $\varphi(t)$ la probabilité qu'une femelle en chaleur en t soit fécondée si un mâle est présent à cette date ;
- de la durée du cycle œstral δ que nous avons supposée fixée.

Avec ces notations on a :

$Q_1(t)$ donné *a priori*

$$Q_2(t + \delta) = Q_1(t) \cdot [1 - \varphi(t) \cdot \epsilon(t)] \text{ pour } t \in [0, D - \delta]$$

$$Q_3(t + 2\delta) = Q_2(t + \delta) \cdot [1 - \varphi(t + \delta) \cdot \epsilon(t + \delta)] \text{ pour } t \in [0, D - 2\delta]$$

etc.

$$\text{et donc Prob}(f(t_i)) = \sum_r Q_r(t_i) \cdot \varphi(t_i) \cdot \epsilon(t_i)$$

la sommation étant étendue sur les valeurs définies des $Q_r(t_i)$.

C. Objectif

Il faut choisir les durées d , r , a et s . En ce qui concerne la longueur de la période de monte d , on notera que son accroissement permet d'augmenter le nombre de descendants par mâle, donc la précision de sa connaissance, mais en contrepartie diminue le nombre de mâles testés par unité de temps, donc les possibilités de sélection. Les descendants sont d'autant mieux discriminés que la durée r des intervalles est grande mais, ici aussi, ce surcroît de précision va à l'encontre des possibilités de sélection. Enfin, on conçoit que plus la période de mesure (a) est longue, plus nombreux seront les descendants attribués au mâle concerné, mais avec de plus en plus de risques d'erreur.

En toute rigueur, le choix des durées devrait donc s'opérer globalement, compte tenu des oppositions évoquées et en prenant en considération les contraintes économiques ou zootechniques qui pèsent sur le système de reproduction : il ne serait pas concevable, par exemple, de changer de mâle tous les jours, même si du point de vue du gain génétique cette solution était optimale.

III. Modèles et critères génétiques

A. Indice de sélection utilisé

1. Expression générale

Nous nous plaçons donc dans la situation où les animaux sont classés sur la régression linéaire de leur valeur génétique en les phénotypes de leurs descendants supposés. Au moment de l'indexation du mâle A_j , on dispose de 2 informations : le nombre n_j de « descendants » nés entre a_j^i et a_j^j et leurs valeurs phénotypiques. Pour simplifier la présentation, nous supposons :

- 1) que les dates de naissance ne sont considérées qu'à travers leur appartenance aux différentes périodes d'observation des mises bas,
- 2) que seule, la valeur moyenne des n_j phénotypes est connue,
- 3) que l'information apportée par les moyennes des descendants attribués aux autres mâles est ignorée dans l'estimation de la valeur génétique de A_j .

Dans ces conditions, reprenant des notations similaires à celles de POIVEY & ELSÉN (1984), la régression linéaire s'inscrit, pour le mâle A_j :

$$\bar{y}_j = \frac{C_{k_j/n_j}}{V_{k_j/n_j}} \cdot \bar{x}_j = \frac{\sum_{k_j} \text{Prob}(S_{k_j/n_j}) \cdot C_{k_j/n_j}}{\sum_{k_j} \text{Prob}(S_{k_j/n_j}) \cdot V_{k_j/n_j}} \cdot \bar{x}_j \quad (1)$$

où \bar{x}_j est la moyenne des valeurs phénotypiques des n_j descendants.

2. Expressions des variances et covariances

C_{k_j/n_j} est la covariance entre la valeur génétique du mâle A_j et la moyenne des n_j valeurs phénotypiques sous l'hypothèse S_{k_j/n_j} .

$$C_{k_j/n_j} = \frac{n_{ij} \cdot \sigma_a^2}{2 \cdot n_j}$$

et V_{k_j/n_j} est la variance de cette moyenne, dans les mêmes conditions.

$$V_{k_j/n_j} = \frac{n_j \cdot \sigma_p^2 + \left(\sum_{i=1}^M n_{ij} \cdot (n_{ij} - 1) \right) \sigma_a^2 / 4}{n_j^2}$$

σ_p^2 étant la variance totale, σ_a^2 la variance génétique additive, telles que $h^2 = \sigma_a^2 / \sigma_p^2$ est l'héritabilité. Sous ces hypothèses, la variance V_n et la covariance C_n qui interviennent dans l'expression (1) de \bar{y}_j sont :

$$C_{n_j} = \frac{\sigma_a^2}{2 \cdot n_j} \cdot \sum_{k_j} n_{ij} \cdot \text{Prob}(S_{k_j/n_j}) = \frac{1}{2} \cdot h^2 \cdot \sigma_p^2 \cdot q_{ij}$$

et

$$V_{n_j} = \frac{1 + 1/4 \cdot (n_j - 1) \cdot h^2 \cdot \left(\sum_{i=1}^M q_{ij}^2 \right)}{n_j} \cdot \sigma_p^2$$

puisque

$$\begin{aligned} & \sum_{k_j} n_{ij} \cdot (n_{ij} - 1) \cdot \text{Prob}(S_{k_j/n_j}) \\ &= \sum_{n_{ij}=0}^{n_j} n_{ij} \cdot (n_{ij} - 1) \cdot \frac{n_j!}{n_{ij}! (n_j - n_{ij})!} \cdot q_{ij}^{n_{ij}} \cdot (1 - q_{ij})^{n_j - n_{ij}} \\ &= (n_j \cdot q_{ij} \cdot (1 - q_{ij}) + (n_j \cdot q_{ij})^2) - n_j \cdot q_{ij} \\ &= n_j \cdot (n_j - 1) \cdot q_{ij}^2 \end{aligned}$$

l'indice de sélection est donc simplement :

$$\bar{y}_j = \frac{1/2 \cdot n_j \cdot h^2 \cdot q_{ij}}{1 + 1/4 \cdot (n_j - 1) \cdot h^2 \cdot \sum_i q_{ij}^2} \cdot \bar{x}_j$$

B. Méthodes et critères de comparaison des protocoles d'accouplement

1. Généralités

Comme l'ont discuté GOFFINET & ELSÉN (1984), les contraintes pour le renouvellement des reproducteurs sont diverses, les plus couramment étudiées étant :

— H_1 : retenir un nombre fixé R de reproducteurs. Il suffit dans ce cas de trier les candidats sur leurs indices \bar{y}_j et de garder la fraction supérieure $\alpha = R/L.M.$ L'espérance du gain génétique sera notée $\hat{\Lambda}$.

— H_2 : retenir en espérance R reproducteurs, le taux de sélection α étant une moyenne sur une infinité d'opérations de sélection. Dans cette situation on peut, conformément à la présentation de COCHRAN (1951) se fixer un seuil λ tel que tout mâle dont l'indice de sélection \bar{y}_j lui est inférieur est éliminé.

Le seuil λ est déterminé par le choix du taux de sélection moyen α . Si une telle procédure est appliquée dans le cas où la loi conjointe de la valeur génétique du mâle et de la moyenne des descendants, conditionnée par la réalisation de S_{k_j/n_j} , est binormale, le gain génétique espéré, défini ici comme l'espérance de la valeur génétique des retenus, sera (POIVEY & ELSÉN, 1984) :

$$\bar{\Lambda} = \frac{1}{L.M} \cdot \sum_j \bar{\Lambda}_j \quad (2)$$

avec

$$\bar{\Lambda}_j = \frac{1}{\alpha} \cdot \sum_{n_j} \text{Prob}(n_j) \cdot \sum_{k_j} \text{Prob}(S_{k_j/n_j}) \cdot Z_{k_j/n_j} \cdot C_{k_j/n_j} \cdot \frac{C_{n_j}}{V_{n_j}} \quad (3)$$

où

$$Z_{k_j/n_j} = \frac{1}{\sqrt{2\pi} \cdot V_{k_j/n_j}} \cdot \frac{V_{n_j}}{C_{n_j}} \cdot \exp \left[-\frac{1}{2} \cdot \frac{\lambda^2 \cdot V_{n_j}^2}{C_{n_j}^2 \cdot V_{k_j/n_j}} \right] \quad (4)$$

$$\alpha = \frac{1}{L.M} \cdot \sum_j \alpha_j \quad (5)$$

$$\alpha_j = \sum_{n_j} \text{Prob}(n_j) \cdot \sum_{k_j} (S_{k_j/n_j}) \cdot \left[1 - \pi \left(\frac{\lambda \cdot V_{n_j}}{C_{n_j} \sqrt{V_{k_j/n_j}}} \right) \right] \quad (6)$$

$\text{Prob}(n_j)$ est la probabilité que n_j descendants naissent entre a_j^- et a_j^+ ; π est la fonction de répartition de la loi normale.

En pratique, par la multiplicité des situations possibles (S_{k_j/n_j}), l'espérance du progrès génétique sous H_2 n'est pas calculable. Le cas de la contrainte H_1 est encore plus complexe puisque l'estimation du gain génétique y fait appel aux statistiques d'ordre.

Nous préconisons donc d'évaluer ces gains par simulation.

2. Description des simulations

Les paramètres biologiques (héritabilité h^2 , durée du cycle δ , lois $\varphi(t)$ et $Q_r(t)$) et le protocole d'indexation (effectifs M , L et F , calendriers de monte et d'attribution)

étant définis, on peut simuler NSIM fois le déroulement de la sélection, en utilisant les lois précédemment déterminées (q_{ij} et $q_0 = \text{prob}(f_0)$). La $s^{\text{ème}}$ simulation se déroule ainsi :

- pour chacun des L.M pères, on génère sa valeur génétique $y_{ji}(s)$ dans une $N(0, 1/4 h^2)$;
- pour chacune des F femelles du lot 1 ($1 = 1 \dots L$)
- = on génère une valeur phénotypique résiduelle de son produit dans une $N(0, 1-1/4 h^2)$
- = on détermine à partir des q_{ij} (et de q_0) le père qui la féconde (s'il y en a un) et celui qui lui est attribué comme conjoint.

Les indices $\bar{y}_{ji}(s)$ sont alors calculés et stockés ainsi que les $y_{ji}(s)$. Par ailleurs on enregistre la valeur génétique moyenne \bar{y}_s des R premiers des L.M mâles triés sur $\bar{y}_{ji}(s)$.

Après NSIM simulations, on peut estimer :

- d'une part l'espérance du gain sous la contrainte H_1 , par $\hat{\Lambda} = \sum_s \bar{y}_s / \text{NSIM}$
- d'autre part le seuil λ et la valeur génétique moyenne $\bar{\Lambda}$ correspondant au taux de sélection $\alpha = R/L.M$ sous H_2 .

3. Critères de comparaison des protocoles

Pour toutes hypothèses sur les valeurs des paramètres et des effectifs, on doit déterminer les meilleurs calendriers de monte et d'attribution. Plusieurs critères peuvent être considérés à cet effet. Les espérances de gain dont nous venons de décrire la méthode de calcul seront estimées pour différents taux de sélection α . On les comparera également aux gains maximum (sous H_1 et sous H_2) qu'il serait possible d'obtenir si les paternités étaient parfaitement connues soit en conservant les mêmes calendriers, soit en se plaçant dans le cas où il n'y a pas d'intervalles entre les périodes de monte ($r = 0$), donc où la fertilité totale est maximum.

En contrepartie à sa simplicité, l'utilisation des simulations exige un grand nombre de répétitions pour atteindre un niveau de précision suffisant. Il serait donc souhaitable de disposer d'un critère de comparaison qui soit plus simple à calculer que les gains $\hat{\Lambda}$ et $\bar{\Lambda}$ tout en donnant à peu près la même classification.

Nous avons étudié dans cette optique une approximation du coefficient de corrélation entre la valeur génétique d'un mâle et la moyenne des valeurs phénotypiques de ses descendants supposés :

$$C = \sqrt{\frac{1}{M} \cdot \sum_{j=1}^M \frac{1/4 \cdot E(n_j) \cdot h^2 \cdot q_{jj}}{1 + 1/4 \cdot E(n_j - 1) \cdot h^2 \cdot (\sum_i q_{ij}^2)}} \quad (7)$$

$$E(n_j) = F \cdot \sum_i q_{ij}$$

Comme cela a été évoqué dans le paragraphe II, le choix du protocole d'indexation devrait reposer sur un critère global tenant compte des possibilités de sélection (nombre

de mâles testés), de diffusion (nombre de mâles retenus), et des coûts. Cette approche globale ne peut être que relative à une structure particulière définie par les besoins annuels en reproducteurs, la capacité de testage, etc. Nous nous contentons ici de mettre en balance gains ($\hat{\Lambda}$, $\bar{\Lambda}$) ou corrélations (ρ) et fertilité moyenne ($1 - q_0$) qui reflète le nombre de descendants produits et donc à la fois la diffusion possible et le manque à gagner dû à l'absence de père dans le troupeau pendant les intervalles entre périodes de monte.

IV. Illustration

A. Cas étudiés

Nous nous sommes placés dans la situation où le ranch est formé de $L = 5$ lots de $F = 40$ vaches. Dans chacun de ces lots, on teste successivement $M = 2$ taureaux, soit un total de 10 mâles par an, sur un caractère d'héritabilité moyenne ($h^2 = 0,3$).

Les paramètres de reproduction ont été choisis compte tenu de considérations sur l'élevage en milieu tropical. Ainsi, LANDAIS *et al.* (1980) ont montré l'importance des dates des mises bas précédentes pour la venue en chaleur des femelles. Il paraît ainsi possible d'envisager une répartition des premières chaleurs uniforme sur toute la saison de monte ($Q_1(t) = 1/D$). Nous avons supposé que celle-ci est de 120 jours, ce qui est assez limité pour que la fécondité instantanée y soit constante ($\varphi(t) = \varphi$).

Dans ces conditions, en supposant $\varphi = 0,5$, on trouve une fécondité annuelle de 82 p. 100 s'il y a toujours un mâle dans le troupeau ($r = 0$).

La distribution de la durée de gestation est centrée sur $g_s = 285$ jours, avec un écart-type de 8 jours, ce qui est une situation moyenne et proche des valeurs données par RALAMBOFIRINGA (1978), étudiant la physiologie de la reproduction chez des femelles bovines de race *N'Dama*. Enfin, on a considéré une durée δ du cycle de 21 jours. Sous cet ensemble d'hypothèses, nous avons calculé les différents critères définis plus haut pour plusieurs valeurs de la durée de non-attribution ($s = 0, 5, 15$ et 20 jours) et de l'intervalle de temps séparant la présence des 2 mâles ($r = 0, 10, 20, 30, 40, 50, 60, 80$ et 100 jours).

Les espérances des progrès génétiques ont été calculées à partir de $NSIM = 12\ 000$ simulations pour les 4 premières valeurs de r , $NSIM = 8\ 000$ pour les 3 suivantes, $NSIM = 4\ 000$ pour les dernières (les nombres de simulations étant augmentés dans les zones où les différences entre progrès sont faibles).

B. Résultats

Les résultats obtenus sous l'hypothèse H_1 étant en tous points identiques à ceux concernant H_2 , nous ne présenterons que ces derniers.

1. Gains génétiques absolus $\bar{\Lambda}$

Les espérances des gains génétiques, en supposant un taux de sélection α de 50 p. 100 sont reportées dans le tableau 1. On doit remarquer en premier lieu la grande

stabilité des gains quand les variables r et s changent. Il est pratiquement indifférent de laisser un intervalle de 0 à 30 jours entre la sortie du premier mâle et l'entrée du second dans le troupeau, et de ne pas attribuer de filiations pendant une durée de 0 à 15 jours. Ce dernier résultat est encore plus évident pour les grandes valeurs de r .

TABLEAU 1

Espérances des gains génétiques $\bar{\Lambda}$ en fonction des durées (en jours) des intervalles r entre périodes de saillies et s entre attributions. (Taux de sélection $\alpha = 50$ p. 100).

Expected genetic gain $\bar{\Lambda}$ as a function of the lengths of the intervals r between mating periods and s between attribution périods. (Selection rate $\alpha = 50$ p. 100).

r	s	0	5	10	15	20
	0	.142	.141	.142	.139	.137
10	.143	.144	.142	.141	.136	
20	.143	.143	.142	.142	.139	
30	.140	.141	.139	.142	.144	
40	.137	.140	.138	.140	.137	
50	.137	.134	.134	.130	.136	
60	.125	.129	.132	.130	.126	
80	.117	.114	.112	.113	.107	
100	.088	.087	.088	.090	.086	

S'il existe une certaine fluctuation des résultats avec r et s , manifestement due à l'imprécision de la méthode, on constate toutefois, *primo* que le gain maximum est obtenu pour r et s proches de 0, *secundo* que pour un intervalle entre attribution s fixé, il existe une valeur optimale de la durée r , qui augmente avec s .

En conclusion, sur ce seul critère de l'espérance du gain, on peut donc préconiser de laisser des intervalles de l'ordre de la semaine tant pour les saillies que pour les attributions.

2. Gains relatifs aux cas de certitude sur les filiations

Le tableau 2 donne les gains $\bar{\Lambda}$ relativement aux espérances des gains maximum $\bar{\Lambda}_m$ (filiation certaine, mêmes calendriers) d'une part, du gain maximum maximorum $\bar{\Lambda}_M$ (filiation certaine, r et s nuls) d'autre part. Les résultats les plus marquants sont :

- que pour les faibles intervalles r (moins de 20 jours), la diminution du gain génétique due à l'incertitude est très faible par rapport à $\bar{\Lambda}_m$, faible par rapport à $\bar{\Lambda}_M$;

• que cette dernière est importante si on augmente beaucoup l'intervalle entre saillies r : la quasi certitude sur les apparentements qu'apporte cette option ne compense pas la forte baisse des effectifs de descendants attribués ;

• qu'au contraire, les différences entre $\bar{\Lambda}$ et $\bar{\Lambda}_m$ s'estompent si r augmente ;

• que les gains relatifs dépendent plus de r que de s dans la gamme de variation étudiée.

TABLEAU 2

Gain $\bar{\Lambda}$ en pourcentage des gains obtenus quand les filiations sont connues :
 $\bar{\Lambda}_m$, au-dessus de la diagonale, pour les mêmes valeurs r et s . $\bar{\Lambda}_M$, au-dessous, pour r et s nuls.
 (Taux de sélection $\alpha = 50$ p. 100 ; r et s en jours).

Gain $\bar{\Lambda}$ as a percentage of that with known filiations : $\bar{\Lambda}_m$, above the diagonal,
 for the same values of r and s . $\bar{\Lambda}_M$ under, for $r = s = 0$.
 (Selection rate $\alpha = 50$ p. 100 ; r and s are in days).

r	s	0	5	10	15	20
0		96	98	98	98	99
	96		95	96	94	92
10		99	100	99	99	100
	96	97		96	95	92
20		99	99	100	100	100
	97	96		96	96	94
30		100	100	100	100	100
	95	95		94	95	97
40		100	100	100	100	100
	93	94		93	94	92
50		100	100	100	100	100
	93	90		90	88	91
60		100	100	100	100	100
	85	86		89	88	85
80		100	100	100	100	100
	79	77		75	76	72
100		100	100	100	100	100
	59	60		59	60	58

En conclusion, ces résultats confirment l'intérêt de limiter les durées des intervalles r et s à moins de 10 jours.

3. Prise en compte de la fertilité

La fertilité moyenne $1 - q_0$ est indépendante de l'intervalle s et diminue si r augmente (tabl. 3). La perte est sensible si r excède 20 jours, ce qui confirme l'intérêt de limiter r et s .

TABLEAU 3

Fertilité $1 - q_0$ selon l'intervalle entre périodes de saillies r .
Fertility $1 - q_0$ as a function of the interval r between mating periods.

r	0	10	20	30	40	50	60	80	100
$1 - q_0$.552	.541	.532	.517	.498	.466	.434	.349	.187

La conclusion générale de cet exemple est qu'on gagne peu à vouloir préciser les paternités. Ceci est lié à la longueur de la période totale de monte $D = 120$ jours, qui couvre 15 écarts-types de durée de gestation. Même pour de très faibles (relativement à D) valeurs de r , on a déjà une bonne certitude sur les filiations.

C. Critères simplifiés

Avant de comparer les gains $\bar{\Lambda}$ et les corrélations ρ , nous devons justifier, *a posteriori*, le choix de $\alpha = 50$ p. 100.

1. Influence du taux de sélection

On a constaté en effet que les classements des protocoles (définis par r et s) en terme de gains génétiques ne sont pas modifiés si on change le taux de sélection α . A titre d'exemple, on trouvera sur la figure 2 le gain espéré $\bar{\Lambda}$ pour divers protocoles et diverses valeurs du taux α . Naturellement, les différences entre gains espérés diminuent si α augmente (pour $\alpha = 100$ p. 100, le gain est nul quel que soit le protocole).

2. Comparaison des gains et corrélations

La corrélation entre ces critères, calculée sur l'ensemble des données, est très élevée : 0,991. Si on se limite au sous-échantillon des 4 premières valeurs de r , où les gains génétiques sont tous estimés à partir du même nombre de simulations, elle reste forte : 0,674. Il est difficile cependant de faire la part entre le manque réel de liaison et l'incertitude due à la méthode de calcul des gains.

En conclusion, on peut donc proposer d'utiliser comme premier critère la corrélation ρ qui est très simple à calculer, et, éventuellement, pour les protocoles les meilleurs, d'effectuer un 2^e tri sur le gain génétique.

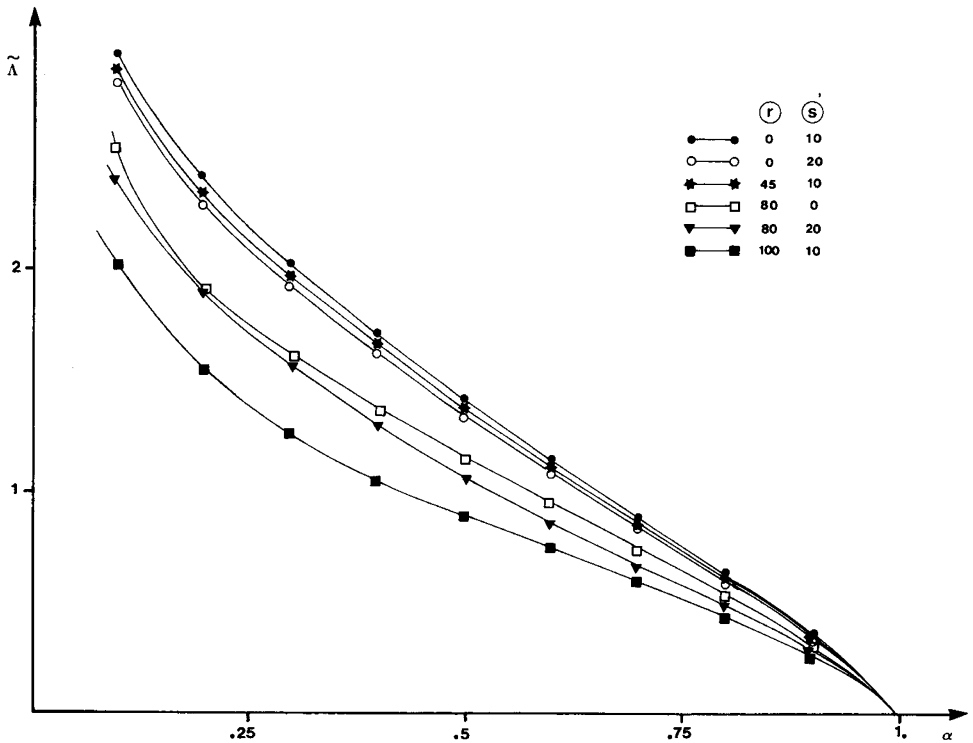


FIG. 2

Gain génétique espéré en fonction du taux de sélection pour quelques valeurs des intervalles r et s .
 Expected genetic gain as a function of the selection rate for some values of the intervals r and s .

3. Exemple

A titre d'illustration, on trouvera sur le tableau 4 les valeurs de ρ en fonction des intervalles r et s pour les mêmes valeurs de paramètres.

On constate que les tendances sont les mêmes que précédemment. Les protocoles optimaux sont ceux pour lesquels r et s sont nuls.

V. Discussion, conclusions

Notre objectif était de proposer une démarche pour l'utilisation des informations apportées par les durées de gestation afin d'accroître la précision de l'indexation. Le fait d'avoir raisonné sur un exemple, obligatoirement de portée limitée, restreint la généralité des résultats. La démarche est cependant répétable. Elle consiste :

- 1) à comparer différents calendriers de monte et d'attribution ;

TABLEAU 4

Corrélation ρ en fonction des durées (en jours) des intervalles r entre périodes de saillies et s entre attributions.

Correlations ρ as a function of length (days) of the intervals r between mating periods and s between attribution periods.

r	s	0	5	10	15	20
0		.670	.662	.653	.644	.634
10		.666	.662	.657	.650	.643
20		.661	.659	.656	.652	.647
30		.652	.651	.650	.648	.646
40		.640	.640	.640	.640	.638
50		.623	.623	.623	.623	.623
60		.606	.606	.606	.606	.606
80		.555	.555	.555	.555	.555
100		.449	.449	.449	.449	.449

2) à calculer les probabilités *a priori* q_{ij} de fertilisation par le i^{me} mâle et attribution au j^{me} mâle, à partir de ces calendriers et d'hypothèses sur les probabilités de premières venues en chaleur, fertilité et durée de gestation ;

3) à simuler le progrès génétique obtenu pour chaque protocole envisagé, en supposant une sélection sur la régression linéaire de la valeur génétique d'un mâle sur celle de ses descendants attribués ;

4) éventuellement à utiliser comme critère de comparaison simplifié la corrélation moyenne.

Dans l'exemple traité, plusieurs hypothèses ont probablement influencé notablement les résultats :

- on a supposé que seule la moyenne des filles attribuées à un mâle servait à son indexation. Il est certain que la prise en compte des dates de naissance individuelles réduirait fortement l'intérêt des saillies et des mises bas intervenant dans les parties extrêmes des périodes de monte et d'attribution, et les durées optimales des intervalles entre ces périodes devraient augmenter. Cependant, la prise en compte de ces informations nécessite l'inversion de la matrice des covariances entre femelles dont les éléments hors diagonaux sont non nuls. Une solution intermédiaire pourrait être de raisonner sur les moyennes des filles nées sur des périodes de durées moyennes telles que la semaine ;

- on a raisonné en terme d'espérance des retenus, considérant aussi la fertilité moyenne. Ces critères ne sont que des éléments de décision, la sanction finale étant plus globale (elle doit faire intervenir les intervalles de génération et les possibilités de diffusion des reproducteurs) et de nature économique (coût de la mesure, recettes liées au protocole choisi) ;

- les hypothèses physiologiques sont importantes. En particulier il faudrait confirmer l'uniformité de la distribution des dates de premières venues en chaleur. Il serait également opportun de prendre en compte la répétabilité de la fertilité, et son évolution au cours de la saison de monte, alors que dans l'exemple, on a supposé cette fertilité invariable ;

- nous n'avons pas considéré dans notre étude les influences saisonnières qui marquent les performances des produits naissant à des périodes différentes. A cet égard, on pourrait faire l'hypothèse que ces effets sont connus *a priori* et que nos données peuvent être corrigées de façon à pouvoir appliquer directement notre méthodologie établie au cours de la première partie. En revanche, si l'on admet, avec le souci d'une plus grande généralité, que des effets fixes doivent être estimés simultanément aux valeurs génétiques des pères, il serait alors nécessaire d'élargir l'étude méthodologique ;

- enfin, rappelons que la régression linéaire n'est qu'une approximation de l'espérance conditionnelle.

Reçu le 12 décembre 1984.

Accepté le 26 septembre 1985.

Références bibliographiques

- COCHRAN W.G., 1951. Improvement by means of selection. *Proceedings of the second Berkeley Symposium on Mathematical Statistics and Probability*, 1951. Neyman J. (ed.), 449-470, University of California Press, Berkeley.
- GOFFINET B., ELSÉN J.M., 1984. Critère optimal de sélection : quelques résultats généraux. *Génét. Sél. Evol.*, **16**, 307-318.
- LANDAIS E., POIVEY J.P., SEITZ J.L., 1980. Recherches sur la reproduction du cheptel taurin sédentaire du Nord de la Côte-d'Ivoire. Utilisation des intervalles entre les vélages ; aspects méthodologiques et premiers résultats. *Rev. Elev. Med. Vet. Pays Trop.*, **33**, 193-204.
- POIVEY J.P., ELSÉN J.M., 1984. Estimation de la valeur génétique des reproducteurs dans les cas d'incertitude sur les apparentements. I - Formulation des indices de sélection. *Génét. Sél. Evol.*, **16**, 445-454.
- RALAMBOFIRINGA A., 1978. Note sur les manifestations du cycle œstral et sur la reproduction des femelles *N'Dama*. *Rev. Elev. Med. Vet. Pays Trop.*, **31**, 91-94.