

J.- M. ELSÉN

INRA Station d'Amélioration Génétique des Animaux BP 27 31326 Castanet-Tolosan Cedex

La gestion des populations

De l'optimisation au progrès génétique réalisé dans les schémas de sélection

Résumé. Un schéma de sélection est une opération complexe qui fait intervenir, par cycles successifs, l'évaluation des valeurs génétiques des reproducteurs ou des candidats à la reproduction, le tri des meilleurs d'entre eux et leurs accouplements. Des modèles mathématiques ont été mis au point pour en décrire le fonctionnement. Ces modèles sont des outils d'aide à la décision pour les sélectionneurs et pour l'état qui les subventionne. A partir d'un exemple concret, nous décrivons ici la structure et l'utilisation de ces modèles. Quelques indications sont également données sur les méthodes d'estimation a posteriori du gain génétique réalisé grâce à ces schémas de sélection.

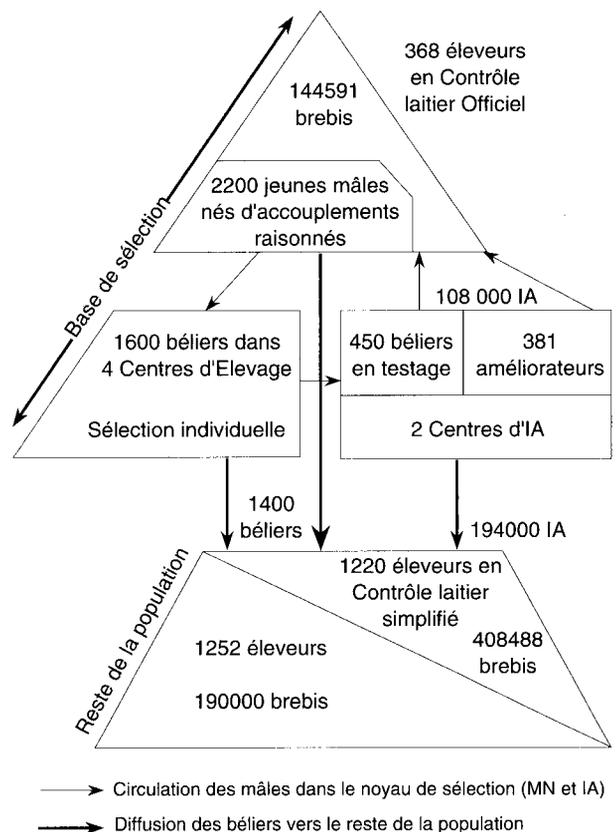
Dans un schéma de sélection, les estimations des valeurs génétiques (index de sélection) sont utilisées pour classer les animaux, les meilleurs sont retenus comme reproducteurs et accouplés, générant ainsi de nouvelles cohortes de candidats à la sélection qui seront à leur tour évalués. L'organisation de la mesure des performances - nécessaire à l'évaluation génétique - les modalités de tri et les plans d'accouplement sont autant de procédures dont la définition et l'enchaînement caractérisent le schéma de sélection. Toutes ces étapes peuvent être optimisées dans leur déroulement afin de maximiser l'efficacité de l'amélioration génétique. Nous allons exposer ici comment modéliser un plan de sélection et le progrès qu'il génère, et comment utiliser le modèle pour chercher les conditions de sélection optimales. Nous montrerons aussi ce que ces plans de sélection donnent effectivement comme progrès et comment ce progrès peut être estimé. Le propos sera illustré par le schéma d'amélioration de la race ovine de Lacaune sur la production de lait.

1 / Exemple de plan de sélection : le schéma de sélection Lacaune lait

Le plan de sélection tel qu'il s'est déroulé en 1989 est présenté sur la figure 1. La population est divisée en deux ensembles : un noyau, où est généré le progrès génétique, qui comporte 144 551 brebis soumises au contrôle laitier intensif, et un étage commercial de 605 098 femelles qui profite du progrès génétique en achetant des reproducteurs (ou de la semence) au noyau. Dans ce noyau, la sélection porte essentiellement sur la production laitière (quantité et qualité) et s'opère de la façon suivante :

Sur 2200 mâles nés d'accouplements entre les meilleurs géniteurs, 1600 sont amenés dans des centres d'élevage à l'âge de un mois, y sont contrôlés

Figure 1. Le schéma de sélection sur la production de lait en race ovine de Lacaune.



sur leur conformation au standard de la race et leur développement.

Parmi les 1600, 450 jeunes béliers sont mis en testage sur descendance : chacun d'entre eux est accouplé, en majorité par insémination artificielle, à des brebis du noyau de sélection, et procréé un effectif d'environ 35 filles qui sont conservées au moins jusqu'à la fin de leur première lactation pour évaluer leur père.

Parmi les 450 béliers testés, 150 environ sont considérés comme améliorateurs et largement utilisés en IA sur les brebis du noyau. Au total, 75 % des brebis contrôlées sont inséminées.

La diffusion du progrès génétique vers les troupeaux commerciaux s'opère à la fois par la vente de reproducteurs (1400 jeunes mâles/an) et par l'insémination artificielle avec de la semence de béliers du noyau (134000 IA).

La description de ce schéma met donc en évidence les trois composantes évoquées plus haut : contrôle de performance / sélection des reproducteurs / accouplements.

Contrôle de performance : après une première phase en centre d'élevage il s'opère ici pour l'essentiel en ferme, ce qui permet de distinguer les élevages créateurs des élevages utilisateurs du progrès génétique. Dans d'autres circonstances, des contrôles pourront être réalisés dans des stations spécialisées où l'environnement est mieux contrôlé (stations de contrôle individuel ou de la descendance).

Sélection des reproducteurs : nous avons vu qu'elle se déroule pour les mâles en plusieurs étapes (au moins deux : après la sortie du centre d'élevage et après le testage sur descendance). C'est une situation très courante qui correspond à une forme d'optimalité dans laquelle le statut de reproducteur de bon niveau génétique est remis en cause de façon continue. Dans d'autre cas le choix peut être définitif, en une étape. Un autre élément important à retenir est le fait d'une sélection sur descendance, qui a d'ailleurs tendance, dans les schémas du type Lacaune laitier, à s'orienter vers une sélection sur l'ensemble des informations disponibles grâce au "modèle animal". Dans d'autres cas (croissance en station) la sélection peut ne porter que sur la performance propre du candidat.

Accouplement : nous repérons ici au moins trois types d'accouplements : les accouplements de testage destinés à procréer des filles pour l'évaluation génétique des béliers, les accouplements entre reproducteurs élites qui vont fournir les cohortes de jeunes mâles à tester, et les accouplements de diffusion vers les élevages commerciaux.

Le schéma présenté montre aussi le caractère dynamique du processus. Cette dimension temporelle intervient à deux niveaux : dans la succession des événements subis par une cohorte d'animaux (mesure, sélection, reproduction) - on parlera d'un cycle de sélection -, et dans la diffusion progressive des effets d'un cycle de sélection sous la forme d'une augmentation de la valeur moyenne des animaux de la population.

2 / Modélisation et optimisation de plans de sélection

2.1 / Modélisation

L'objectif est de prévoir l'évolution de l'"état" de la population qui serait soumise au schéma de sélection caractérisé par ses trois composantes. Celles-ci seront décrites à la fois dans leur structure générale (contrôle individuel ou de la descendance, sélection en une ou plusieurs étapes ...) et dans le détail de leur organisation (effectif d'animaux contrôlés, taux de sélection...) par une série de relations entre variables aboutissant à une fonction objectif.

a / Variables

Pour caractériser l'état de la population, il est nécessaire de la diviser en classes ($c=1, \dots, C$), ou sous-populations, caractérisées chacune par une ou des variables d'état. La définition des classes varie d'un schéma de sélection à l'autre, et d'un modèle à l'autre, mais respecte toujours la règle que tous les animaux appartenant à la même classe sont soumis aux mêmes décisions (de contrôle, sélection ou accouplement). Ainsi, les individus d'une même classe sont souvent de même sexe et âge, alors que les femelles de trois ans soumises au contrôle laitier pourront être distinguées des femelles de trois ans des élevages commerciaux.

Les variables d'état associées à chaque classe les caractérisent sur le plan génétique. Il peut s'agir le plus simplement (et le plus souvent) de l'espérance de la valeur génétique des individus de la classe, mais dans certains modèles, notamment dans le cas de la sélection en plusieurs étapes, l'état d'une classe est décrit par la fonction de densité de la valeur génétique. D'autres indicateurs sont parfois utilisés qui décrivent l'appareillement entre les membres de la population : taux moyen de consanguinité, probabilités d'origine des gènes ...

Distinguons dès à présent les cas où le processus de sélection se répète de façon invariable d'une cohorte à la suivante, des cas où ce processus évolue au cours du temps. Dans la première situation, les classes seront définies indépendamment du temps (on parlera de la classe des mâles de un an en testage), alors que dans le second les individus d'une classe seront contemporains (on parlera de la classe des mâles de un an en testage en 1993). Le second type de modèle a été utilisé notamment pour décrire l'évolution de la variabilité génétique. Nous nous limiterons maintenant au premier type, le plus fréquent.

Les valeurs prises par les variables d'état évoluent selon la dynamique du processus de sélection et le modèle doit rendre compte de cette évolution. En fait deux types de modèles ont été développés. Les premiers estiment le progrès génétique asymptotique et ne considèrent que les effets à long terme de la sélection. Dans ce cas, les classes sont définies à un niveau très global, par exemple les quatre classes du modèle de Rendel et Robertson (1950) : pères à fils, pères à filles, mères à fils, mères à filles. Les seconds modèles, plus récents, décrivent l'évolution des valeurs génétiques pas à pas (selon une unité de temps telle que l'année), considérant ainsi les effets à court et long terme de la sélection (Elsen et Mocquot 1974, Hill 1974). Dans ce cas, la définition des classes

est plus fine, distinguant en particulier les animaux dont les âges sont distants de plus d'une unité de temps. L'équivalence asymptotique de ces deux approches a été démontrée.

Le modèle fait intervenir des variables élémentaires parmi lesquelles nous distinguons des paramètres qui sont imposés par la "nature" (au sens de la théorie des jeux), des variables de décision qui décrivent le détail des composantes des schémas de sélection et des variables internes au modèle qui relient paramètres et variables de décision aux variables d'état.

Les paramètres sont répartis en quatre groupes (tableau 1). Les variables de décision sont pour l'essentiel des effectifs d'animaux (ou des rapports d'effectifs) et des durées d'utilisation des reproducteurs (tableau 2).

Les variables internes sont des quantités communément utilisées en génétique quantitative. Ce sont des intermédiaires entre les paramètres et variables de décision d'une part, les variables d'état et la fonction objectif d'autre part. Ainsi sont définies :

- l'intervalle de génération L, ou âge moyen des parents quand naissent leur descendants ;
- le coefficient de détermination r_{IA} d'un indice de sélection, ou corrélation entre la valeur génétique et son estimateur ;
- la différentielle de sélection S, ou l'écart phénotypique (au sens large : ce peut être un écart de valeur

d'indice de sélection) entre les sélectionnés et les retenus.

b / Relations entre variables

Elles sont de deux types : contraintes dues à la limitation des disponibilités et relations fonctionnelles entre variables de différents niveaux.

Dans l'organisation du schéma de sélection certaines contraintes imposent de faire des choix entre la précision de l'estimation des valeurs génétiques, l'intensité du tri des reproducteurs, l'ampleur de leur diffusion et le maintien de la variabilité génétique sur le long terme. Le modèle doit rendre compte de ces limitations.

Les contraintes sont essentiellement de trois natures : démographique, biologique ou organisationnelle. Selon les cas elles se traduisent sous la forme d'équations ou d'inéquations. Le tableau 3 donne les plus courantes d'entre elles.

Les relations fonctionnelles décrivent comment variables de décision et paramètres sont liés pour former les variables internes et, au-delà, l'évolution des variables d'état. Par exemple, nous aurons pour le coefficient de détermination r_{IA} :

$$r_{IA}^2 = \frac{1/4 \cdot n \cdot h^2}{1 + 1/4 \cdot (n - 1) \cdot h^2}$$

Groupe de paramètres			
Démographiques	Génétiques	Zootecniques	Economiques
mortalité fertilité femelle prolificité mortinatalité sex-ratio taille de la population F	héritabilité h^2 corrélation perte due à la consanguinité	moyenne de départ variance totale taille des élevages pourcentage d'IA nombre d'éjaculats/mâle	coût des mesures coût d'entretien/mâle coût de calcul marge du produit

Tableau 1.
Paramètres des modèles de schéma de sélection.

Variable	Symbole
Taux de sélection	q
Nombre de mâles élites	E
Nombre total de mâles sélectionnés	S
Nombre de mâles en station de contrôle individuel	Ts
Nombre de mâles en testage sur descendance	Td
Nombre de descendants/mâle en testage	n
Pourcentage de femelles contrôlées	c
Part de la population contrôlée en accouplements de testage	u
Nombre de descendants/mâle sélectionné	d
Durée d'utilisation des mâles	l

Tableau 2.
Exemple de variables de décision des modèles de schéma de sélection.

Type de contrainte	Formulation
Certaines variables sont des fractions : les taux de sélection le taux de testage le taux de contrôle	$0 \leq q \leq 1$ $0 \leq u \leq 1$ $0 \leq c \leq 1$
Le nombre total de femelles est donné La capacité de testage est fixée par u La capacité de diffusion des mâles est limitée Les mâles testés sont fils d'élite	$F.c = F.c.u + F.c.(1-u)$ $F.c.u = Td.n$ $F.c.(1-u) = S.d.1$ $Ts = k.E$

Tableau 3.
Contraintes des modèles.

et pour la différentielle de sélection S :

$$S = \frac{\exp(-x^2/2) \cdot \sigma_i}{\sqrt{2\pi} \cdot q}$$

(x étant le seuil de sélection et σ_i l'écart type de l'indice de sélection)

En prenant pour exemple le modèle déjà évoqué de Rendel et Robertson, quatre variables d'état seront attachées aux quatre classes définies : père à fils MM, père à filles MF, mère à fils FM, mère à filles FF. Ces variables d'état prendront des valeurs successives données par le système d'équations suivant :

$$MM_t = 1/2(MM_{t-1} + FM_{t-1}) + \Delta_{MM}$$

$$MF_t = 1/2(MM_{t-1} + FM_{t-1}) + \Delta_{MF}$$

$$FM_t = 1/2(MF_{t-1} + FF_{t-1}) + \Delta_{FM}$$

$$FF_t = 1/2(MF_{t-1} + FF_{t-1}) + \Delta_{FF}$$

$$MM_t = MM_{t-1} - L_{MM} \cdot \Delta g$$

$$MF_t = MF_{t-1} - L_{MF} \cdot \Delta g$$

$$FM_t = FM_{t-1} - L_{FM} \cdot \Delta g$$

$$FF_t = FF_{t-1} - L_{FF} \cdot \Delta g$$

où l'indice t est celui de la génération, $\Delta_i = S_i \cdot \sigma_i^2$ l'espérance de la valeur génétique des reproducteurs sélectionnés pour la classe i, L_i l'intervalle de génération pour la classe i et Δg le progrès génétique annuel. Le progrès génétique annuel se déduit du système d'équations précédent par la formule :

$$\Delta g = \frac{\Delta_{MM} + \Delta_{MF} + \Delta_{FM} + \Delta_{FF}}{L_{MM} + L_{MF} + L_{FM} + L_{FF}}$$

Dans le cas de modèles pas à pas, l'ensemble des variables d'état (espérances des valeurs génétiques de chaque classe d'individus) est regroupé dans un vecteur X_t , et l'évolution du système s'écrit, comme ci-dessus, par un système d'équations du type $X_t = D \cdot X_{t-1} + \Delta$, où Δ est un vecteur de supériorités et D une matrice de transition.

2.2 / Optimisation

Le modèle permet d'évaluer une fonction objectif qu'il faut optimiser.

a / Fonction objectif

Elle peut être de nature technique ou économique.

Critères techniques

Le progrès génétique annuel, tel que défini ci-dessus, est évidemment le premier critère de choix des schémas de sélection. Il concerne d'abord l'évolution du caractère (ou de l'ensemble des caractères) faisant l'objet de la sélection, et, dans certains cas, celle de caractères corrélés.

L'écart entre les valeurs génétiques moyennes du noyau de sélection et des élevages commerciaux, et de façon générale entre sous populations, est un élément important puisque, à long terme, même si toutes les strates évoluent à la même vitesse, un écart subsiste, qui est d'autant plus important que le nombre de reproducteurs issus du noyau est relativement faible et leur sélection moins intense (Bichard 1971).

L'augmentation du taux de consanguinité est critique dans le cas de petites populations. Cependant, en l'absence de théorie claire sur l'effet de la consanguinité sur les caractères phénotypiques, les auteurs se contentent de donner pour chaque plan de sélection

étudié le couple de valeurs progrès génétique annuel x augmentation du taux de consanguinité.

La diminution de la variabilité génétique, très liée au critère précédent, est un élément qui a pris beaucoup d'importance avec le développement des techniques modernes de reproduction (superovulation et transfert d'embryons) (Nicholas et Smith 1983, Colleau 1985).

Critères économiques

La sélection est une activité coûteuse dont les résultats techniques sont obtenus progressivement. Poutous et Vissac en 1962 ont été les premiers à considérer cette dimension économique. Le thème a été largement étudié avec la publication de nombreux articles entre 1970 et 1980. Le principe général est de mettre en balance coûts et recettes (Elsen et Mocquot 1974).

Les coûts concernent surtout le contrôle de performance et l'entretien des candidats à la sélection qui ne seront pas gardés pour la reproduction. Au démarrage d'un schéma de sélection, il faut aussi envisager des investissements tels qu'une station de contrôle des performances. Il s'agit également des coûts de calcul des indices de sélection, des coûts de contrôle de filiation et des coûts liés à l'utilisation de techniques de reproduction qui n'auraient pas de raison d'être utilisées en l'absence d'un schéma de sélection (transfert d'embryon, insémination intra utérine). Tous ces coûts sont supportés par le "sélectionneur" (selon les cas, la firme, l'UPRA ou l'unité de sélection) éventuellement aidé par l'état. Ils interviennent dès la mise en place du schéma de sélection.

Les recettes sont au contraire perçues plus tardivement (au fur et à mesure que les générations expriment le progrès génétique) et par l'ensemble des éleveurs dont le cheptel est amélioré. Dans tous les modèles, il est en effet supposé que les recettes sont proportionnelles au supplément de valeur des animaux exprimant le progrès génétique par rapport à une situation sans sélection, à leur effectif N (éventuellement un vecteur d'effectif), au nombre E d'expressions du caractère par animal (nombre de lactations, de mise bas, de tontes) et à la marge/unité de caractère, m. Le supplément de valeur des animaux est une fonction croissante du temps, (t. Δg à la date t dans les modèles asymptotiques, X_t dans les modèles pas à pas), alors que les autres composants sont fixes. Les animaux exprimant le progrès génétique sur un caractère sont très généralement un sous-ensemble de la population totale : les vaches en lactation pour la production de lait, les mères pour la prolificité, les jeunes pour la vitesse de croissance...

Les recettes à la date t s'expriment donc comme $R_t = (N \cdot X_t) \cdot E \cdot m$ ou t. $\Delta g \cdot E \cdot m$

Les séries de coûts Ct et de recettes Rt sont actualisées (taux r) et sommées, les coûts et recettes globales, qui s'écrivent :

$$C = \sum_{t=1}^T C_t / (1+r)^t \quad \text{et} \quad R = \sum_{t=1}^T R_t / (1+r)^t$$

étant comparés; pour un horizon T, dans un critère unique qui peut être le bénéfice actualisé (R-C), le taux de rentabilité (ro tel que R = C) ou le temps de récupération (To tel que R = C).

b / Méthode d'optimisation

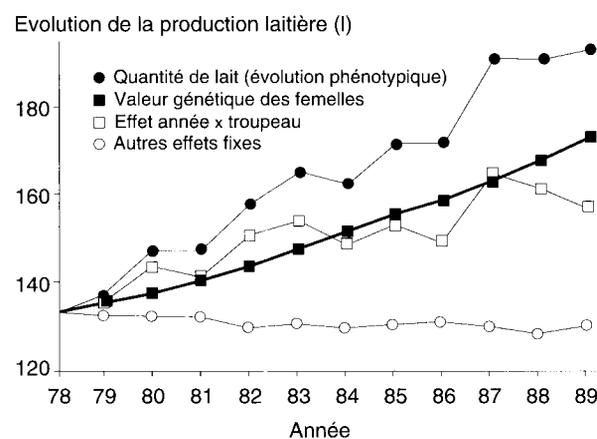
Quelle que soit la fonction objectif retenue, il s'agit finalement d'un problème de maximisation d'une fonction mathématique sous contraintes. Diverses méthodes existent pour résoudre ce problème (méthode du simplexe, du gradient réduit...). Cependant les auteurs ont le plus souvent travaillé par quadrillage, évaluant la fonction objectif sur un ensemble de cas possibles. Cette démarche a l'inconvénient de sa lenteur (certaines publications font état de dizaines de milliers de situations explorées) mais présente l'intérêt de fournir une vision globale du comportement de la fonction objectif dans l'espace de décision. Il est d'ailleurs le plus souvent noté que la fonction objectif varie peu dans de larges zones autour de l'optimum, ce qui donne beaucoup de liberté dans le choix du plan de sélection. Enfin, la sensibilité de la réponse à un schéma de sélection, à la variation de paramètres instables (prix) ou mal connus (corrélations génétiques), est également un élément important à considérer dans le choix final.

3 / Retour à l'exemple du schéma de sélection Lacaune lait : du modèle à la réalité

Le schéma de sélection de la race Lacaune a été modélisé à plusieurs reprises pour en optimiser le fonctionnement, rechercher les meilleures conditions d'utilisation de l'insémination artificielle et évaluer l'intérêt du transfert d'embryon. Le dernier modèle réalisé était basé sur des valeurs des variables de décision (nombre de mâles élites, d'insémination artificielle...) correspondant au schéma de sélection tel qu'il fonctionnait ces dix dernières années. Ce modèle aboutissait à une prévision de progrès génétique de 5,9 kg de lait/an.

Des calculs récents, à partir de l'étude par un modèle animal de un million de lactations réalisées

Figure 2. Progrès génétique et phénotypique donné par le schéma de sélection sur la production de lait en race Lacaune.



de 1978 à 1989 ont estimé le progrès réalisé à 3,6 kg lait/an (figure 2) (Barillet *et al* 1992).

Cette différence importante entre prédiction et réalisation est un phénomène courant qui montre les limites des modèles utilisés. Elle est probablement due en grande partie à l'existence de "pertes de char-

ge" souvent mal considérées dans les modèles : de nombreux animaux sont écartés de la reproduction pour des raisons sanitaires, de non conformité au standard de la race ou de mauvaise libido, ce qui diminue les pressions de sélection possibles. Par ailleurs, les performances qui servent à l'indexation des reproducteurs sont soumises à l'effet de nombreux facteurs de variations (troupeau, année...) qui sont rarement pris en compte dans les modèles de schéma de sélection. Enfin, un grand nombre d'hypothèses relatives à l'équilibre du plan de sélection (nombre de descendants par reproducteurs) et à l'invariance des paramètres et des variables de décision au cours du temps ne sont pas respectées dans la réalité.

Ainsi, il faut compléter les modèles de schémas de sélection par des études a posteriori du déroulement et de l'efficacité de la sélection réalisée. Le déroulement sera décrit par un ensemble d'indicateurs portant sur les pertes de charges relatives à la sélection, la qualité des contrôles de performances, la cohérence des effectifs des différentes classes d'animaux... Ceci a conduit, en ce qui concerne l'espèce ovine, à la réalisation de "tableaux de bord" des schémas de sélection qui fourniront de façon régulière aux sélectionneurs des éléments d'appréciation de la qualité de leur travail. L'efficacité technique de la sélection est estimée par le progrès génétique réalisé. Dans le passé, divers protocoles ou techniques statistiques avaient été proposées pour cette estimation : lignée témoin (avec possibilité d'utilisation de sperme ou d'embryons congelés) et répétition de génotypes (plusieurs mesures/individus, répétition d'accouplements, de pères). La méthode de référence est maintenant l'utilisation du modèle animal, par exemple en calculant les moyennes intra classe intra année des effets individuels.

Conclusion

Les schémas de sélection sont capables de fournir une augmentation annuelle de productivité de 2 % environ. Cette augmentation est obtenue à la suite de cycles de sélection enchaînant contrôle, tri et mise à la reproduction. Ces opérations sont coûteuses et méritent d'être optimisées. Les modèles mathématiques sont un outil d'aide à la décision qui devrait être utilisé plus souvent dans le choix des plans de sélection. Même s'ils ont tendance à surestimer systématiquement le progrès génétique, on peut supposer que le classement relatif des plans possibles restera le plus souvent respecté.

Références bibliographiques

- Barillet F., Boichard D., Barbat A., Astruc J.M., Bonaiti B., 1992. Use of an animal model for genetic evaluation of the Lacaune dairy sheep. *Livest. Prod. Sci.*, 31.
- Richard M., 1971. Dissemination of genetic improvement through a livestock industry. *Anim. Prod.*, 13, 401-411.
- Colleau J.J., 1985. Efficacité génétique du transfert d'embryons dans les noyaux de sélection chez les bovins laitiers. *Génét. Sél. Evol.*, 17, 499 - 538.

Elsen J.M., Mocquot J.C., 1974. Recherches pour une rationalisation technique des schémas de sélection des bovins et des ovins. Bulletin Technique du Département de Génétique Animale, 17, INRA, Paris.

Elsen J.M., 1980. Diffusion du progrès génétique dans les populations avec générations imbriquées : quelques propriétés d'un modèle de prévision. Ann. Génét. Sél. Anim., 12, 49 - 80.

Hill W.G., 1974. Prediction and evaluation of response to selection with overlapping generations. Anim. Prod. 18, 117 - 140.

Nicholas F.W., Smith C., 1983. Increased rates of genetic rates of genetic changes in dairy cattle by embryo transfer and splitting. Anim. Prod., 36, 341 - 353.

Rendel J.M., Robertson A., 1950. Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. J. Genet., 50, 1 - 8.