

Application de la théorie des indices de sélection à des caractères répétés. Exemple de la sélection sur la prolificité chez le porc

M. TARTAR et G. BOLET

avec la collaboration technique de Marie-Laure LE PAIH,
Marie-Reine PERRETANT et D. TANGUY

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

Résumé

Pour des caractères répétés (production laitière, prolificité...), l'indice de sélection (I_m) classiquement utilisé est basé sur la moyenne des n observations et ne dépend que de l'héritabilité et la répétabilité du caractère ; nous lui comparons un indice combiné (I_c) où chaque observation est pondérée par un coefficient calculé à partir de la théorie de HAZEL, prenant en compte les corrélations génétiques (r_G) entre répétitions, éventuellement inférieures à 1. Pour un objectif de sélection H constant, nous faisons varier le nombre d'observations dans l'indice (1 à 5), l'héritabilité (0,10 et 0,60) et les corrélations génétiques r_G (de 0 à 1 par pas de 0,1). Pour une même valeur de ces paramètres, la précision (corrélation entre H et I) et l'efficacité (espérance de progrès génétique sur H par unité d'intensité de sélection) de I_m et I_c sont pratiquement identiques ; la perte de précision de I_m par rapport à I_c est au maximum de 9 p. 100 ; elle est d'autant plus élevée que r_G est faible et que le nombre de mesures augmente. Le gain de précision et d'efficacité de I_m ou I_c lié à la multiplication des mesures est d'autant plus important que r_G est élevé ; il est plus important pour $h^2 = 0,10$ que pour $h^2 = 0,60$. L'hypothèse d'identité génétique entre observations ($r_G = 1$), si elle est fautive, conduit à surestimer la précision et l'efficacité de la sélection ; cette surestimation est d'autant plus importante que le nombre de mesures est élevé ; elle varie également avec l'héritabilité. Ainsi, dans le cas de la sélection sur la prolificité chez le porc ($h^2 = 0,10$), la surestimation de l'espérance de progrès génétique peut aller jusqu'à 300 p. 100 ; le cas de ce caractère est discuté en fonction des estimations de r_G obtenues dans la littérature, elles sont variables mais inférieures à 1, ce qui permet d'expliquer en partie le désaccord entre les estimations théoriques et les résultats pratiques des expériences de sélection.

Mots clés : Indice de sélection, corrélation génétique, prolificité, porc.

Summary

*Application of the theory of selection indices to repeated characters.
Example of the selection on prolificacy in pigs*

The selection index (I_m) which is classically used for repeated characteristics (e.g. milk production, prolificacy) is based on an average of n measurements and depends only on the coefficients of heritability (h^2) and repeatability (r) of the trait concerned. As an alternative, a combined index (I_c) is suggested in which each observation is weighted by a coefficient calculated according to HAZEL's

theory ; this takes into account genetic correlations (r_G) between repeated performances. For a desired breeding goal (H), this index is compared to the conventional one by varying the number of observations made (1 to 5), the values of h^2 (0.1 to 0.6) and r_G (0 to 1 in steps 0.1). For a given value of these parameters, the accuracy (correlation between H and I) and the efficiency (expected genetic improvement on H by selecting on I_m and I_c respectively) of I_m and I_c are almost the same. The greatest loss of accuracy of I_m relative to I_c is 9 p. 100 ; the difference being more marked as r_G is lower and n greater. The accuracy and efficiency of both I_m and I_c are improved with increasing n and this is amplified by increasing r_G and decreasing h^2 values. If the hypothesis of $r_G = 1$ is wrong, the accuracy and the efficiency of selection may be overestimated ; this overestimation is all the larger as the number of measures is higher ; it also varies with heritability. For example, in the case of selection on prolificacy in pigs ($h^2 = .10$), overestimation of the expected genetic progress may be 300 p. 100. This character is discussed according to estimations of r_G made in the literature ; these estimations are variable but lower than 1 and this may partly explain the discrepancies between theoretical expectations and the actual results obtained in selection experiments.

Key words : Selection index, genetic correlation, prolificacy, pig.

I. Introduction

Le principe de la sélection individuelle est de prédire la valeur génétique d'un individu à partir d'un critère, ou indice de sélection I, rassemblant les informations sur sa valeur phénotypique. Si H est l'objectif de sélection, ensemble des valeurs génétiques à prédire, l'efficacité de la sélection peut se mesurer grâce à R_{HI} , coefficient de corrélation entre H et I. HAZEL (1943) et ROUVIER (1977) ont élaboré les formules d'indices optimaux permettant de maximiser cette corrélation ; 2 cas classiques se présentent :

— Lorsque la sélection porte sur un seul caractère, l'objectif H est la valeur génétique G pour ce caractère, et l'indice optimal élaboré à partir de n observations successives sur l'individu est :

$$I = \frac{nh^2}{1 + (n - 1) r} P_n \quad (1)$$

où h^2 est l'héritabilité du caractère, r sa répétabilité, et P_n la moyenne des n observations, en écart à la moyenne de la population. Cet indice est proportionnel à la moyenne des valeurs phénotypiques mesurées ;

— Lorsque la sélection porte sur plusieurs caractères, l'objectif de sélection H peut s'exprimer sous la forme d'une combinaison linéaire des m valeurs génétiques G_k à prédire, affectées d'une pondération a_k liée à leur importance relative. L'indice de sélection est une combinaison linéaire des n valeurs phénotypiques mesurées, affectées de coefficients b_i :

$$I = \sum_{i=1}^n b_i P_i \quad (2)$$

Ces 2 cas s'appliquent, *a priori*, à des situations totalement différentes. Notamment, l'utilisation de l'indice de la formule (1) suppose que la corrélation génétique entre les performances successives est égale à 1, c'est-à-dire qu'elles sont l'expression d'un seul génotype G. Or un certain nombre de travaux, portant par exemple sur les différentes lactations chez la vache (FREEMAN, 1960 ; BARKER & ROBERTSON, 1966) ou sur la taille

des portées chez la truie (OLLIVIER & BOLET, 1981 ; BOLET & FELGINES, 1981 ; JOHANSSON & KENNEDY, 1982 ; VIDOVIC, 1982) ont fait apparaître des corrélations génétiques entre performances successives inférieures à 1.

Dans ce cas, l'hypothèse d'identité génétique entre répétitions est remise en cause, et l'indice de la formule (1), élaboré pour un objectif unique G, n'est plus optimal pour l'objectif multiple H. Cette nouvelle approche soulève 2 problèmes :

— calculer la précision et l'efficacité de l'indice de la formule (1), par rapport à cet objectif H ;

— formuler, en fonction de ce même objectif H, le véritable indice optimal à partir de la formule (2), et comparer sa précision et son efficacité à celles du premier.

Notre objectif est donc d'étudier ces 2 aspects du problème dans une gamme de situations théoriques et, à titre d'exemple, d'envisager une application à la sélection sur la prolificité chez le porc.

II. Méthodologie

Nous étudions et nous comparons l'efficacité et la précision de la sélection basée, pour un même objectif et avec les mêmes hypothèses génétiques, sur 2 types d'indice individuel :

— L'indice de la formule (1), que nous appellerons indice moyen ou I_m est, à un coefficient près, la somme des n observations individuelles :

$$I_m = \sum_{i=1}^n P_i$$

— L'indice de la formule (2), que nous appellerons indice combiné ou I_c , est celui calculé, selon la théorie de HAZEL (1943) reprise par ROUVIER (1977) à partir des pondérations des composantes de l'objectif de sélection, des matrices de variances et covariances génétiques et phénotypiques entre les composantes de l'objectif et de l'indice :

$$I_c = \sum_{i=1}^n b_i P_i \text{ (ou, sous forme matricielle, } I_c = \mathbf{bP})$$

Soit l'objectif de sélection $H = \mathbf{aG}$ où \mathbf{a} est le vecteur des pondérations et \mathbf{G} le vecteur des valeurs génétiques, I_c , meilleur estimateur de H est tel que :

$$I_c = \hat{H} = \mathbf{a S}_{gp} \mathbf{S}_{pp}^{-1} \mathbf{P} = \mathbf{bP}$$

(\mathbf{X} = transposée de la matrice \mathbf{X} ; \mathbf{X}^{-1} = inverse de la matrice \mathbf{X}).

\mathbf{P} est le vecteur des valeurs phénotypiques mesurées, \mathbf{S}_{gp} et \mathbf{S}_{pp} sont avec \mathbf{S}_{gp} , les matrices de variances et covariances des vecteurs \mathbf{G} et \mathbf{P} .

La précision de l'indice de sélection est mesurée par R_{HI} , coefficient de corrélation entre H et I :

$$R_{HI} = \frac{\text{Cov}(H, I)}{\sigma_H \sigma_I}$$

L'efficacité est mesurée par $E(\Delta H)$, espérance du progrès génétique sur l'objectif de sélection par génération :

$$E(\Delta H) = i \cdot R_{HI} \sigma_H$$

où i est l'intensité de sélection.

Ces 2 paramètres, permettant de comparer I_m et I_c , dépendent de l'ensemble des vecteurs et matrices \mathbf{a} , \mathbf{G} , \mathbf{P} , \mathbf{S}_{gg} , \mathbf{S}_{pp} et \mathbf{S}_{gp} définies précédemment (ROUVIER, 1977).

$$\text{Cov}(H, I_c) = {}^t \mathbf{a} \mathbf{S}_{gp} \mathbf{S}_{pp}^{-1} {}^t \mathbf{S}_{gp} \mathbf{a}$$

$$\text{Cov}(H, I_m) = {}^t \mathbf{a} \mathbf{S}_{gp} \mathbf{1} \quad (\mathbf{1} : \text{matrice } (n,1) \text{ unité})$$

$$\sigma_{I_c}^2 = {}^t \mathbf{a} \mathbf{S}_{gg} \mathbf{a}$$

$$\sigma_{I_c}^2 = {}^t \mathbf{b} \mathbf{S}_{pp} \mathbf{b} = {}^t \mathbf{a} \mathbf{S}_{gp} \mathbf{S}_{pp}^{-1} {}^t \mathbf{S}_{gp} \mathbf{a}$$

$$\sigma_{I_m}^2 = {}^t \mathbf{1} \mathbf{S}_{pp} \mathbf{1}$$

Lorsque le nombre d'observations est égal au nombre de valeurs génétiques à prédire, nous calculons également la précision et l'efficacité de l'indice « de base » I_b (WILLIAMS, 1962), dont les coefficients sont directement les pondérations de l'objectif de sélection :

$$I_b = \sum_{i=1}^m a_i P_i = {}^t \mathbf{a} \mathbf{P}$$

$$\text{Cov}(H, I_b) = {}^t \mathbf{a} \mathbf{S}_{gp} \mathbf{a}$$

$$\sigma_{I_b}^2 = {}^t \mathbf{a} \mathbf{S}_{pp} \mathbf{a}$$

Hypothèses :

— Vecteurs \mathbf{a} , \mathbf{G} , \mathbf{P} :

Le caractère sélectionné est exprimé en unités d'écart-type, le nombre de répétitions prises en compte dans l'indice varie de 1 à 5 ; par contre, l'objectif de sélection est constant pour tous les calculs, c'est la valeur génétique de l'individu pour 5 répétitions, affectées de pondérations que nous avons fixées, à titre d'exemple, en nous inspirant de la pyramide des effectifs des truies dans le cheptel français en fonction de leur numéro de mise bas :

$$H = G_1 + 0,79G_2 + 0,63G_3 + 0,45G_4 + 0,32G_5$$

— Matrice \mathbf{S}_{gg} :

La variance génétique σ_g^2 du caractère est supposée identique quelle que soit la répétition. La matrice \mathbf{S}_{gg} peut donc s'exprimer sous la forme $\sigma_g^2 \times \mathbf{S}'_{gg}$ ou $h^2 \times \mathbf{S}'_{gg}$ (puisque la variance phénotypique est égale à 1). \mathbf{S}'_{gg} est une matrice symétrique dont la diagonale prend la valeur 1 et dont les autres termes sont les corrélations génétiques r_G entre les répétitions 2 à 2. Nous lui avons donné la forme suivante :

$$\mathbf{S}'_{gg} = \begin{pmatrix} 1 & x & x-0,1 & x-0,2 & x-0,3 \\ & 1 & x+0,1 & x & x-0,1 \\ & & 1 & x+0,1 & x \\ & & & 1 & x+0,1 \\ & & & & 1 \end{pmatrix} \quad (\text{transposée})$$

Cette formulation est inspirée des résultats de BOLET & FELGINES (1981) ; elle suppose que la corrélation génétique entre observations adjacentes est supérieure à celle entre répétitions non consécutives, et que les corrélations incluant la première sont plus faibles que les autres. Nous avons fait varier x de 0 à 1 ; les corrélations déduites de la valeur de x ne peuvent toutefois pas être négatives ni, évidemment, supérieures à 1. D'autre part, l'héritabilité prend une valeur faible ($h^2 = 0,10$) et une élevée ($h^2 = 0,60$).

— Matrice S_{pp} :

Puisque l'écart-type phénotypique est égal à 1, S_{pp} est une matrice dont la diagonale est 1 et les autres termes sont les corrélations phénotypiques r_p ; elles se déduisent des corrélations génétiques r_G et environnementales r_E par la formule classique :

$$\begin{aligned} r_{p_{ij}} &= r_{G_{ij}} \sqrt{h_i^2 h_j^2} + r_{E_{ij}} \sqrt{(1 - h_i^2)(1 - h_j^2)} \\ r_{p_{ij}} &= r_{G_{ij}} h^2 + r_{E_{ij}} (1 - h^2) \quad (\text{puisqu'on a } h_i^2 = h_j^2) \end{aligned}$$

$r_{E_{ij}}$ est supposée constante et nous avons fixé sa valeur à 0,08 (d'après BOLET & FELGINES, 1981).

Pour $h^2 = 0,10$ d'une part, $h^2 = 0,60$ d'autre part, nous avons donc calculé et comparé R_{HI_m} , R_{HI_c} , $E(\Delta H)$ en utilisant I_m et I_c , en faisant varier x dans S'_{eg} de 0 à 1. Les calculs ont été réalisés en APL.

III. Résultats

Pour les mêmes paramètres génétiques, et pour le même nombre de mesures, le rapport de la précision des 2 types d'indices $\frac{R_{HI_m}}{R_{HI_c}}$ est égal au rapport de leur efficacité

$\frac{E(\Delta H)_m}{E(\Delta H)_c}$, puisque $E(\Delta H) = iR_{HI}\sigma_H$. Les valeurs de ces rapports apparaissent dans le tableau 1, pour x variant de 0 à 1, et le nombre de mesures variant de 2 à 5 (pour 1 mesure, ce rapport est égal à 1). On voit que ce rapport varie de 0,91 à 1, qu'il est d'autant plus faible que les paramètres génétiques sont faibles et que le nombre de mesures est élevé. Il faut remarquer que, lorsque toutes les corrélations génétiques sont égales à 1, nous sommes ramenés à l'hypothèse d'identité génétique et I_m et I_c sont égaux. Ces rapports étant, dans la plupart des cas, très proches de 1, nous ne nous intéresserons, dans la suite de cet article, qu'à l'indice combiné I_c , les conclusions à tirer sur l'indice moyen I_m étant identiques.

Les figures 1 ($h^2 = 0,10$) et 2 ($h^2 = 0,60$) présentent l'évolution de la précision de l'indice en fonction des corrélations génétiques et du nombre de mesures. On voit que le gain de précision lié à la multiplication des mesures est d'autant plus important que les corrélations génétiques sont élevées. Les figures 3 et 4 permettent de comparer les résultats pour $h^2 = 0,10$ et $h^2 = 0,60$, dans un nombre de situations limité à 5 :

TABLEAU 1

Rapport de la précision et de l'efficacité de l'indice moyen sur celles de l'indice combiné pour les héritabilités 0,10 et 0,60.

Ratio of accuracy and efficiency of the average index on those of the combined index with heritabilities .10 and .60.

x	r _G minimum	r _G maximum	Nombre de mesures (number of records)							
			2		3		4		5	
			h ² 0,10	h ² 0,60	h ² 0,10	h ² 0,60	h ² 0,10	h ² 0,60	h ² 0,10	h ² 0,60
0,00	0,0	0,0	0,992	0,993	0,979	0,981	0,951	0,958	0,913	0,926
0,00	0,0	0,1	0,996	0,997	0,991	0,989	0,971	0,970	0,935	0,937
0,10	0,0	0,2	1,00	1,00	0,996	0,995	0,984	0,980	0,952	0,949
0,20	0,0	0,3	1,00	0,999	0,998	0,999	0,988	0,986	0,964	0,960
0,30	0,0	0,4	0,999	0,998	0,999	1,00	0,994	0,992	0,974	0,969
0,40	0,1	0,5	0,999	0,998	1,00	1,00	0,996	0,995	0,984	0,978
0,50	0,2	0,6	0,999	0,998	1,00	1,00	0,998	0,997	0,990	0,984
0,60	0,3	0,7	0,999	0,998	0,999	1,00	0,999	0,998	0,993	0,988
0,70	0,4	0,8	0,999	0,997	0,999	0,999	0,999	0,999	0,996	0,992
0,80	0,5	0,9	0,999	0,997	0,999	0,999	0,999	1,00	0,997	0,994
0,90	0,6	1,0	0,999	0,996	0,999	0,999	0,999	1,00	0,998	0,996
1,00	0,7	1,0	1,00	0,997	1,00	0,999	1,00	1,00	1,00	0,999
1,00	0,8	1,0	1,00	0,999	1,00	0,999	1,00	1,00	1,00	0,999
1,00	0,9	1,0	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
1,00	1,0	1,0	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

x : Coefficient de la matrice S' _{gg} (coefficient of the matrix S' _{gg}).

r_G : Corrélation génétique entre mesures (genetic correlation between measures).

- indépendance génétique entre les mesures (x = 0) ;
- identité génétique entre les mesures (x = 1) ;
- corrélation faible entre les mesures (x = 0,2) ;
- corrélation moyenne entre les mesures (x = 0,5) ;
- corrélation forte entre les mesures (x = 0,8).

Si l'on prend pour base 100 R_{HI} (fig. 3) et E (ΔH) (fig. 4) pour 1 mesure et x = 0, on voit que le gain relatif de précision et d'efficacité quand le nombre de mesures augmente est plus élevé pour h² = 0,10 que pour h² = 0,60.

Pour 5 mesures, nous comparons également la précision de I_b à celles de I_c et I_m, à partir des rapports $\frac{R_{HIb}}{R_{HIc}}$ et $\frac{R_{HIb}}{R_{HIc}}$ (tabl. 2). L'indice combiné est toujours le plus précis, l'indice « de base » I_b est légèrement plus précis que I_m pour les corrélations génétiques faibles, moins précis pour les plus élevées.

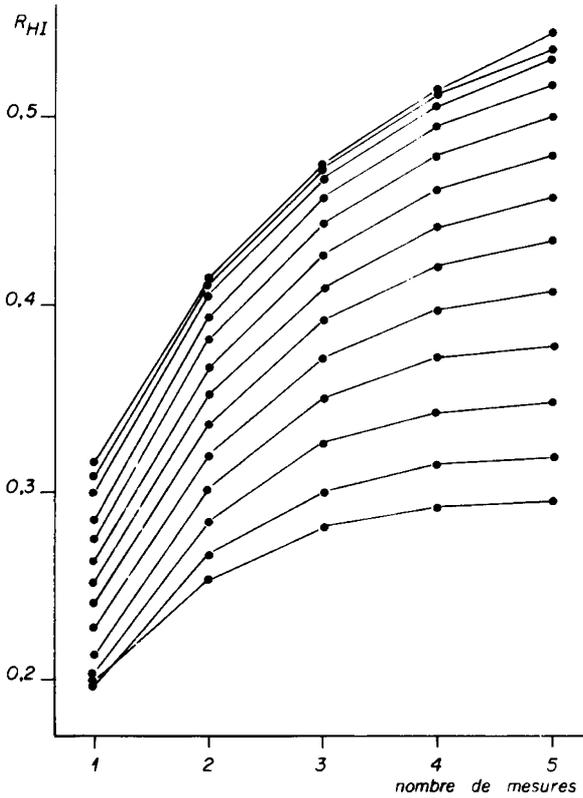


FIG. 1

Précision de l'indice combiné en fonction du nombre de mesures dans l'indice et de x , pour $h^2 = 0,10$.

Accuracy of the combined index according to the number of measurements in the index and to x with $h^2 = .10$.

h^2 : hérabilité de chaque mesure (*heritability of each measurement*).

x coefficient de S_{gg} varie de 0 pour la courbe la plus basse jusqu'à 1 pour la plus haute (voir tabl. 1) (*coefficient of S_{gg} , varies from 0 for the lowest curve to 1 for the highest curve, see table 1*).

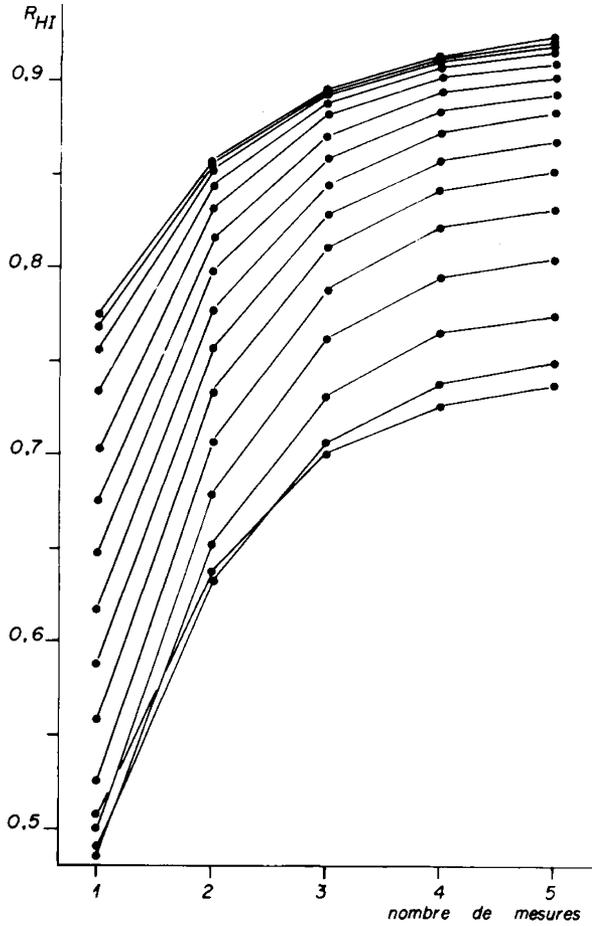


FIG. 2

*Précision de l'indice combiné. Id. figure 1, $h^2 = 0,60$.
Accuracy of the combined index. Id. fig. 1, $h^2 = .60$.*

TABLEAU 2

Rapport de la précision de l'indice « de base » sur celles de l'indice moyen et de l'indice combiné pour 5 mesures, pour les héritabilités 0,10 et 0,60.

Ratio of accuracy of the base index on those of the average index and of the combined index for 5 measures, with heritabilities .10 and .60.

x	r _G minimum	r _G maximum	R _{HI_b} /R _{HI_m}		R _{HI_b} /R _{HI_c}	
			h ² 0,10	h ² 0,60	h ² 0,10	h ² 0,60
0,00	0,0	0,0	1,088	1,079	0,994	0,999
0,00	0,0	0,1	1,067	1,067	0,997	0,999
0,10	0,0	0,2	1,046	1,053	0,996	0,999
0,20	0,0	0,3	1,029	1,041	0,992	0,999
0,30	0,0	0,4	1,015	1,031	0,988	0,998
0,40	0,1	0,5	1,001	1,020	0,985	0,998
0,50	0,2	0,6	0,992	1,014	0,982	0,998
0,60	0,3	0,7	0,985	1,009	0,979	0,997
0,70	0,4	0,8	0,980	1,005	0,976	0,997
0,80	0,5	0,9	0,976	1,002	0,973	0,996
0,90	0,6	1,0	0,973	1,000	0,971	0,996
1,00	0,7	1,0	0,972	0,998	0,971	0,996
1,00	0,8	1,0	0,970	0,996	0,970	0,995
1,00	0,9	1,0	0,968	0,994	0,968	0,994
1,00	1,0	1,0	0,967	0,993	0,967	0,993

x : Coefficient de la matrice S' _{gg} (coefficient of the matrix S' _{gg}).

r_G : Corrélacion génétique entre mesures (genetic correlation between measures).

IV. Discussion

A. Comparaison des deux types d'indices I_m et I_c

Dans toute cette étude, nous comparons les indices pour un nombre de mesures identique. Le problème du biais lié à la comparaison d'individus connus sur une quantité d'informations variable ne rentre pas dans notre propos.

A paramètres génétiques identiques, la précision et l'efficacité de l'indice moyen et de l'indice combiné pour l'objectif H sont pratiquement identiques. Il faut cependant noter que nous avons formulé un certain nombre d'hypothèses restrictives qui ne permettent pas de généraliser cette conclusion ; notamment, l'héritabilité est identique quelle que soit la répétition, les corrélations génétiques ne sont jamais négatives et l'écart-type de toutes les répétitions est égal à 1. Il faut également signaler que les pondérations économiques de l'objectif de sélection sont assez arbitraires ; toutefois, PEASE *et al.* (1967) ont montré que leur imprécision a relativement peu d'influence sur l'efficacité de l'indice. Par ailleurs, le but de cette étude est son application aux mesures répétées,

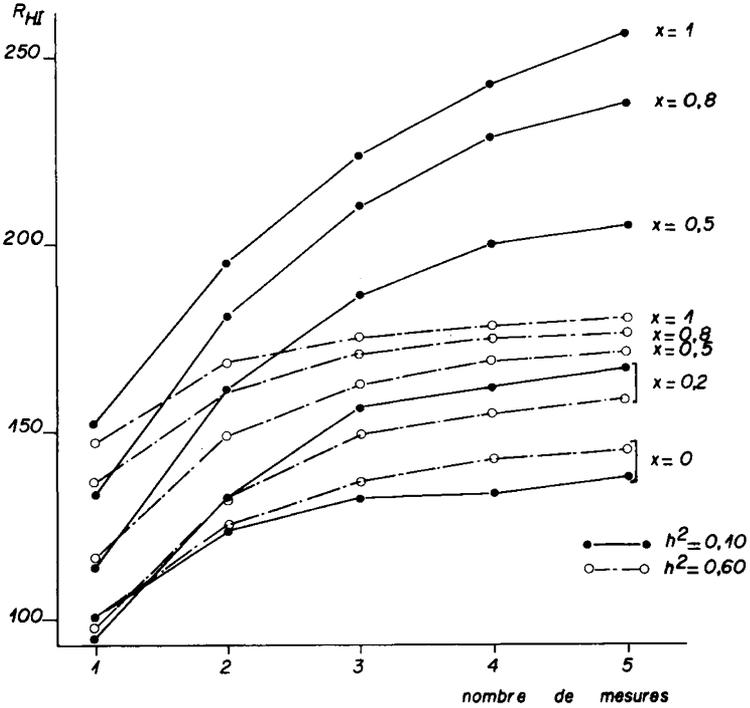


FIG. 3

Accroissement de la précision de l'indice combiné en fonction du nombre de mesures et de x . La base 100 est le cas de 1 mesure avec $x = 0$.

Increment of the accuracy of the combined index according to the number of measurements and to x . The base 100 is the case when is 1 measurement and $x = 0$.

et les hypothèses que nous avons prises sont proches de celles que l'on peut rencontrer, par exemple pour la production laitière chez les bovins (RENDEL *et al.*, 1957 ; BARKER & ROBERTSON, 1966 ; MAJALA & HANNA, 1974) ou la prolificité chez le porc (BOLET & FELGINES, 1981 ; BOLET & LEGAULT, 1982). On peut donc admettre, en première approche, que l'indice moyen est un outil de sélection convenable, et les conclusions que nous allons apporter sur l'indice combiné lui sont applicables.

B. Analyse de l'indice « combiné »

1. - Sous l'hypothèse d'identité génétique des mesures successives, l'indice combiné est égal à l'indice moyen et l'on retrouve les résultats classiques : la répétition des mesures présente d'autant plus d'intérêt que l'héritabilité est faible : en passant de 1 à 5 mesures, la précision et l'efficacité de l'indice s'accroissent de 72 p. 100 quand $h^2 = 0,10$ et 19 p. 100 quand $h^2 = 0,60$.

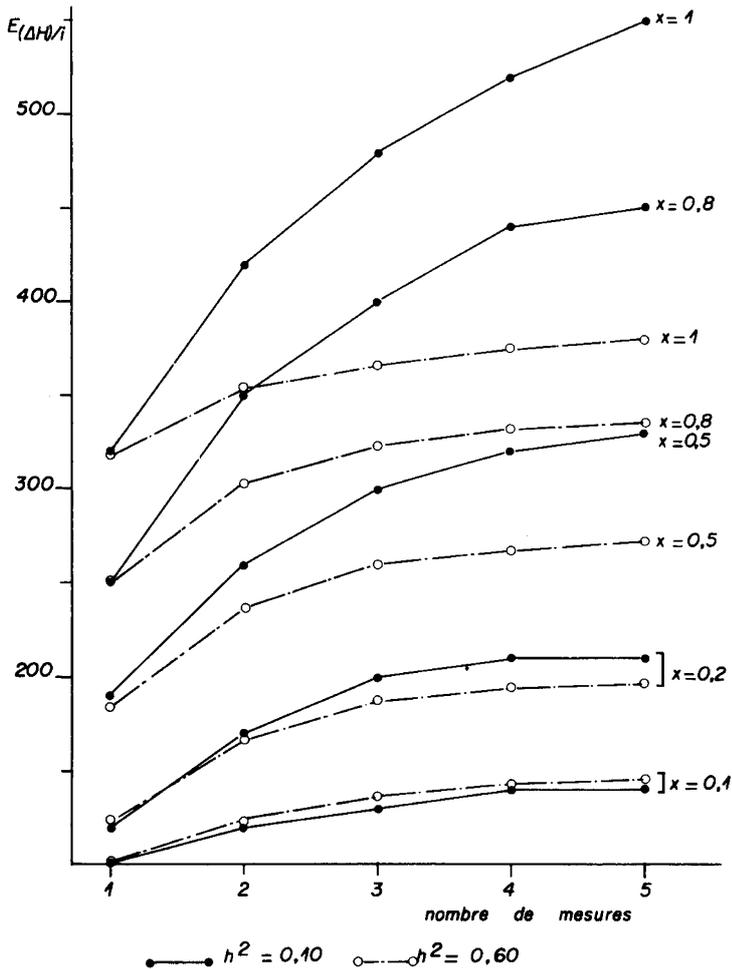


FIG. 4

Accroissement de l'espérance de progrès génétique pour H en fonction du nombre de mesures et de x . La base 100 est le cas de 1 mesure avec $x = 0$.
 Increment of the expected genetic progress on the breeding goal H according to the number of measurements and to x . The base 100 is the case when there is 1 measurement and $x = 0$.

2. - Pour un nombre de mesures donné, les figures 3 et 4 permettent de voir l'influence des variations de paramètres génétiques sur la précision et l'efficacité de l'indice de sélection optimum. Ces 2 paramètres évoluent de façon similaire, l'espérance du progrès génétique étant proportionnelle à la précision de l'indice ; leur augmentation en fonction de S_{gg} dépend du nombre de mesures et de l'héritabilité. A titre d'exemple, le tableau 3 présente leur accroissement quand on passe de l'hypothèse d'indépendance ($x = 0$) à celle d'identité génétique ($x = 1$). Quand le nombre de mesures augmente, le rapport augmente pour $h^2 = 0,10$ et diminue pour $h^2 = 0,60$. Ainsi, si l'on admet

TABLEAU 3

Rapport de la précision et de l'efficacité de l'indice combiné pour $r_G = 1$ sur celles pour $r_G = 0$, en fonction du nombre de mesures et de l'héritabilité h^2 .

Ratio of the accuracy and the efficiency of the combined index with $r_G = 1$ on those with $r_G = 0$, according to the number of measures and the heritability h^2 .

Nombre de mesures (Number of records)	$h^2 = 0,10$		$h^2 = 0,60$	
	(R_{HI})	E	(R_{HI})	E
1	1,53	3,19	1,53	3,19
2	1,62	3,39	1,35	2,81
3	1,71	3,59	1,27	2,66
4	1,81	3,78	1,25	2,62
5	1,91	3,98	1,25	2,62

(R_{HI}) : Rapport de R_{HI} pour $r_G = 1$ sur R_{HI} pour $r_G = 0$.
(Ratio of R_{HI} with $r_G = 1$ on R_{HI} with $r_G = 0$.)

E : Rapport de E (ΔH) pour $r_G = 1$ sur E (ΔH) pour $r_G = 0$.
(ratio of E (ΔH) with $r_G = 1$ on E (ΔH) with $r_G = 0$)

que l'hypothèse d'identité génétique est une surestimation, elle conduit également à surestimer l'efficacité de la sélection : pour une faible héritabilité, cette surestimation sera d'autant plus forte que le nombre de mesures sera élevé ; pour une forte héritabilité, ce sera l'inverse. On voit donc que, comme les variations de corrélations génétiques, les variations du nombre de mesures n'ont pas la même influence suivant l'héritabilité du caractère. Ceci peut s'expliquer en partie par le mode de calcul de la corrélation phénotypique entre mesures : avec $h^2 = 0,60$, elle augmente fortement quand x augmente (elle passe de 0,03 à 0,63), alors qu'elle varie moins pour $h^2 = 0,10$ (elle passe de 0,07 à 0,17) ; cet accroissement peut alors masquer en partie l'influence des seules variations génétiques.

3. - De ces 2 approches des paragraphes 1 et 2 ressort une conclusion commune : une mauvaise estimation des corrélations génétiques influence notre appréciation de l'intérêt de la multiplication des mesures et de l'efficacité attendue de la sélection, mais elle apparaît d'autant plus grave que l'héritabilité des caractères est faible. Aussi une attention particulière doit-elle être portée aux erreurs d'estimation des paramètres génétiques et phénotypiques, des estimations erronées compromettant l'efficacité réelle de l'indice utilisé. SALES & HILL (1976) ont par exemple montré que, pour une sélection individuelle sur un caractère, l'adjonction d'un second caractère apportant une information sur le premier pouvait diminuer le progrès génétique faute d'une précision suffisante dans l'estimation des paramètres. Nos résultats montrent que, lorsqu'on n'est pas en mesure d'estimer les paramètres génétiques avec précision, il n'est pas opportun d'élaborer un indice combiné optimal pour l'objectif ; l'indice moyen n'amène qu'une perte de précision faible ; l'indice « de base », affecté directement des pondérations économiques, semble pratiquement aussi précis, notamment pour les caractères à forte héritabilité

(tabl. 2). Il faut toutefois signaler que, si l'efficacité de ces indices est comparable pour l'objectif global H, l'espérance de progrès génétique pour chaque composante de H est plus sensible au type d'indice utilisé.

Si la sélection sur indice se confirme être une méthode robuste, la prédiction de son efficacité, par contre, est très hasardeuse lorsque les estimations des paramètres ne sont pas fiables. De telles situations sont susceptibles de se présenter lorsque la matrice S_{gg} de covariances génétiques entre caractères est « non définie positive » (HILL & THOMPSON, 1978). Retenant cette propriété comme règle de validité des paramètres estimés, HAYES & HILL (1980) ont proposé une transformation canonique des variables permettant de déterminer facilement les propriétés statistiques des coefficients de l'indice. Cette démarche doit permettre de lever une part d'indétermination sur les réelles possibilités de la sélection ; elle passe nécessairement par une meilleure estimation des paramètres, notamment pour les caractères faiblement héritables.

V. Application à la sélection sur la prolificité chez le porc

La taille de la portée à la naissance, chez le porc, est un caractère dont l'héritabilité et la répétabilité ont été étudiées par de nombreux auteurs et sont faibles, respectivement de l'ordre de 0,10 et 0,15 (BOLET & LEGAULT, 1982). L'écart-type phénotypique est de l'ordre de 2,7 pour la première portée et 2,9 pour les suivantes. On suppose classiquement que les portées successives d'une truie sont l'expression d'un même caractère, c'est-à-dire que la corrélation génétique est égale à 1. C'est sur cette base qu'est proposé aux sélectionneurs français un indice de prolificité qui est, à un coefficient de proportionnalité près, la moyenne des performances de la truie, corrigées pour les effets de milieu (LEGAULT *et al.*, 1981). C'est également sur cette base que OLLIVIER & BOLET (1981) ont calculé la réponse théorique à une sélection individuelle sur la somme des 2 premières portées : avec un taux de sélection de 1/10 dans le choix des mâles, et 1/3 dans le choix des femelles, la réponse théorique est de 0,25 porcelet par portée, soit, pour l'objectif H défini précédemment, 0,83 porcelet. Or, un certain nombre d'auteurs ont montré que la corrélation génétique entre portées est inférieure à 1 ; les valeurs obtenues varient énormément : 0 pour OLLIVIER & BOLET (1981), de 0,1 à 0,9 pour BOLET & FELGINES (1981), 0,5 à 1 pour JOHANSSON & KENNEDY (1982) ou VIDOVIC (1982). Le tableau 4 présente, pour 4 valeurs de x , défini précédemment, correspondant à cette gamme de valeurs possibles, l'indice de sélection combiné optimal I_c .

A la suite des résultats du § II, nous savons maintenant que, en fait, l'indice actuellement proposé aux éleveurs, qui est un indice de type I_m , est pratiquement aussi efficace, à hypothèses génétiques identiques et peut donc continuer à être utilisé, en attendant d'avoir une estimation fiable des corrélations génétiques entre portées. SCHULTECOERNE (1983) arrive à une conclusion sensiblement identique pour la sélection sur la production laitière chez les bovins. Cette démarche a toutefois l'intérêt d'affecter une pondération économique à chaque mesure, et donc de relativiser la part du progrès génétique global revenant à chaque mesure, en fonction de son importance économique.

Par contre, la prédiction de l'efficacité de cette sélection dépend de ces corrélations génétiques : le tableau 4 montre que, entre l'hypothèse d'indépendance ($x = 0$) et celle d'identité génétique ($x = 1$), la précision de l'indice est multipliée par 1,6 à 1,9 et son efficacité par 3,4 à 4, suivant le nombre de portées dans l'indice. Ceci peut expliquer,

TABLEAU 4

Coefficients de l'indice et espérance de progrès génétique pour chaque portée, en fonction de x et du nombre de portées figurant dans l'indice.

Coefficients of the index and expected genetic progress by parity, according to x and the number of litters in the index.

Nombre de portées (Number of litters) x	Coefficient de chaque portée dans l'indice (Coefficient of each litter in the index)												
	P1		P2		P3		P4		P5		R _{HI}	E (ΔH) pour i = 1	
	Coef. G ₁	Coef. G ₂	Coef. G ₃	Coef. G ₄	Coef. G ₅	Coef. G ₆	Coef. G ₇	Coef. G ₈	Coef. G ₉				
0	1	1	0,27		0,00		0,00		0,00		0,00	0,20	0,27
	2	1	0,20	0,8	0,18		0,00		0,00		0,00	0,25	0,34
	3	1	0,18	0,8	0,16	0,6	0,12		0,00		0,00	0,28	0,37
	4	1	0,17	0,8	0,15	0,6	0,11	0,4	0,07		0,00	0,29	0,39
	5	1	0,17	0,8	0,15	0,6	0,11	0,3	0,07	0,2	0,04	0,29	0,39
0,4	1	1	0,27		0,12		0,09		0,06		-0,03	0,22	0,45
	2	1	0,25	1,1	0,28		0,16		0,12		0,08	0,32	0,65
	3	1	0,23	1,1	0,29	1	0,28		0,17		0,12	0,37	0,75
	4	1	0,22	1,1	0,30	1	0,29	0,8	0,25		0,16	0,40	0,81
	5	1	0,22	1,1	0,30	1	0,30	0,8	0,27	0,5	0,22	0,41	0,83
0,8	1	1	0,27		0,23		0,20		0,17		0,15	0,27	0,71
	2	1	0,32	1,1	0,35		0,31		0,27		0,38	0,38	0,98
	3	1	0,34	1,1	0,40	1,1	0,38		0,34		0,30	0,44	1,14
	4	1	0,35	1,1	0,42	1	0,42	1	0,40		0,35	0,48	1,24
	5	1	0,35	1	0,44	1	0,44	0,9	0,43	0,8	0,40	0,50	1,29
1	1	1	0,27		0,29		0,29		0,29		0,29	0,32	0,91
	2	1	0,35	0,9	0,38		0,38		0,38		0,38	0,41	1,18
	3	1	0,40	0,9	0,43	0,9	0,43		0,43		0,43	0,47	1,35
	4	1	0,44	0,9	0,47	0,9	0,47	0,9	0,47		0,47	0,51	1,47
	5	1	0,47	0,9	0,50	0,9	0,50	0,9	0,50	0,9	0,50	0,55	1,56

Coef. : Coefficients de l'indice.

(Coefficients of the index.)

G : Espérance de progrès génétique par portée, pour i = 1.
(Expected genetic progress for each litter, with i = 1.)

R_{HI} : Coefficient de corrélation entre l'objectif et l'indice de sélection.
(Correlation coefficient between the breeding goal and the selection index.)

E (ΔH) : Espérance de progrès génétique sur l'objectif de sélection.
(Expected genetic progress for the breeding goal.)

en partie seulement, la contradiction apparente entre les estimations théoriques classiques et les résultats de la sélection sur la prolificité chez le porc. Ainsi, dans l'expérience citée ci-dessus (OLLIVIER & BOLET, 1981), l'espérance de progrès génétique passe de 0,25 porcelet en 1^{re} portée pour x = 1 à 0,14 pour x = 0 (soit de 0,83 à 0,24 pour H).

Les hypothèses classiques (x = 1) surestiment les possibilités de la sélection individuelle sur la prolificité, et ce d'autant plus que le nombre de portées prises en compte

dans l'indice est élevé. Ceci pose particulièrement un problème lorsqu'on compare des truies sur des indices de prolificité prenant en compte un nombre variable de portées : c'est le cas par exemple du programme de détection de truies « hyperprolifiques » dans les élevages français (LEGAULT *et al.*, 1981). Ceci renforce donc l'intérêt d'une sélection précoce sur la prolificité, qui permet par ailleurs de bénéficier d'une réduction de l'intervalle de générations ; cette conclusion doit cependant être modulée du fait que pour la première portée, intervenant chez la truie de plus en plus précocement, les conditions physiologiques semblent différentes des portées suivantes. L'intérêt de la prise en compte de la prolificité des truies dans les stratégies de sélection a déjà été discuté, et démontré pour certaines situations (LEGAULT, 1970 ; ELSÉN & SELLIER, 1978). Mais, étant donnée la variabilité des données existantes, il est indispensable d'améliorer avant tout la précision des estimations de corrélations génétiques.

VI. Conclusion

Cette étude constitue une première approche d'un problème qui, à notre connaissance, n'a jamais été étudié sous cet angle. Nous avons dû nous imposer un certain nombre de contraintes concernant les valeurs de certains paramètres fixés *a priori*. Par contre, les valeurs d'héritabilité retenues et le champ de variation des corrélations génétiques donnent à cette analyse un caractère tout à fait général, applicable à des caractères très divers, tels que la production laitière ou la prolificité. Il apparaît très nettement qu'en considérant les mesures successives comme des répétitions d'un caractère, on peut être amené à se tromper sur les possibilités d'amélioration génétique par sélection individuelle. Si les paramètres génétiques sont mal estimés, il n'est peut-être pas utile d'élaborer un indice « combiné », puisque la moyenne des mesures, ou, dans certains cas, l'indice « de base », se révèle constituer un indice simple presque aussi efficace ; par contre, il faut savoir, que, en supposant que la corrélation entre mesures est égale à 1, on a toutes les chances de surestimer fortement la précision et l'efficacité de la sélection mise en œuvre. Ceci est particulièrement vrai pour les caractères à faible héritabilité.

Reçu le 3 mai 1983.

Accepté le 8 mars 1984.

Références bibliographiques

- BARKER J.S.F., ROBERTSON A., 1966. Genetic and phenotypic parameters for the first three lactations in Friesian cows. *Anim. Prod.*, **8**, 221-240.
- BOLET G., FELGINES C., 1981. Héritabilité de la prolificité, corrélations phénotypiques et génétiques entre les quatre premières portées chez des truies de race *Large White*. *32nd Annual Meeting of the European Association for Animal Production, 31 août-3 septembre 1981, Zagreb. Ann. Génét. Sél. Anim.*, **14**, 121-122 (abstr.).
- BOLET G., LEGAULT C., 1982. New aspects of genetic improvement of prolificacy in pigs. *2nd World Congress on Genetics applied to Livestock Production, Madrid, 4-8 octobre 1982*, **5**, 548-567, Editorial Garsi, Madrid.

- ELSEN J.M., SELLIER P., 1978. Etude conjointe de l'intérêt de la sélection sur la prolificité et de l'utilisation d'une lignée mâle spécialisée chez le porc. *Ann. Génét. Sél. Anim.*, **10**, 403-442.
- FALCONER D.S., 1974. *Introduction à la génétique quantitative*. Masson, Paris, 248 p.
- FREEMAN A.E., 1960. Genetic relationships among the first three lactations of *Hosltain* cows. *J. Dairy Sci.*, **43**, 876-877 (abstr.).
- HAYES S.F., HILL W.G., 1980. A reparameterization of a genetic selection index to locate its sampling properties. *Biometrics*, **36**, 237-248.
- HAZEL L.N., 1943. The genetic basis for constructing selection indexes. *Genetics*, **28**, 476-490.
- HILL W.G., THOMPSON R., 1978. Probabilities of non-positive definite between-group of genetic covariance matrices. *Biometrics*, **34**, 429-439.
- JOHANSSON K., KENNEDY B.W., 1982. Estimation of genetic parameters of fertility traits in pigs. *2nd World Congress on Genetics applied to Livestock Production, Madrid, 4-8 octobre 1982*, **7**, 503-508, Editorial Garsi, Madrid.
- LEGAULT C., 1970. Recherche d'un taux optimum de sélection des jeunes truies sur la prolificité de leur mère. *2^e Journées de la Recherche Porcine en France, Paris, 19-20 février 1970*, 241-249, I.T.P., Paris.
- LEGAULT C., GRUAND J., BOLET G., 1981. Résultats de l'utilisation en race pure et en croisement d'une lignée dite « hyperprolifique ». *13^e Journées de la Recherche Porcine en France, Paris, 4-5 février 1981*, 255-260, I.T.P., Paris.
- MAIJALA K., HANNA M., 1974. Reliable phenotypic and genetic parameters on dairy cattle. *1st World Congress on Genetics applied to Livestock Production, Madrid, 7-11 octobre 1974*, **1**, 541-563, Editorial Garsi, Madrid.
- OLLIVIER L., BOLET G., 1981. La sélection sur la prolificité chez le porc : résultats d'une expérience de sélection sur dix générations. *13^e Journées de la Recherche Porcine en France, Paris, 4-5 février 1981*, 261-267, I.T.P., Paris.
- PEASE A.H.R., COOK G.L., GREIG M., CUTHBERTSON A., 1967. Combined testing. Report DA 188 of the pig industry development authority. Hitchin, Herts, England.
- RENDEL J.M., ROBERTSON A., ASKER A.A., KISHIN S.S., RAGAB M.T., 1957. The inheritance of milk production characteristics. *J. Agric. Sci.*, **48**, 426-431.
- ROUVIER R., 1977. Mise au point sur le modèle classique d'estimation de la valeur génétique. *Ann. Génét. Sél. Anim.*, **9**, 17-26.
- SALES J., HILL W.G., 1976. Effect of sampling errors on efficiency of selection indices. *Anim. Prod.*, **23**, 1-14.
- SCHULTE-COERNE H., 1983. Comparison of selection indices using repeatability models and multiple trait models. *34th Annual Meeting of the European Association for Animal Production, Madrid, 3-6 octobre 1983*, **1**, 83, Neografis, Madrid.
- VIDOVIC V., 1982. Estimation of genetic parameters for reproductive performance in purebred and crossbred sows. *2nd World Congress on Genetics applied to Livestock Production, Madrid, 4-8 octobre 1982*, **7**, 288-292, Editorial Garsi, Madrid.
- WILLIAMS J.S., 1962. Some statistical properties of a genetic selection index. *Biometrika*, **19**, 325-337.