

## Relations a priori entre BLUP (\*) méthode de comparaison aux contemporaines et méthodes des différences cumulées en vue de l'évaluation des pères

J. L. FOULLEY et J. M. ELSEN

Station de Génétique quantitative et appliquée  
Centre national de Recherche zootechniques, I.N.R.A.,  
78350 Jouy en Josas

---

### Résumé

La méthode de comparaison aux contemporaines, la méthode des différences cumulées de BAR ANAN et SACKS ainsi que les versions modifiées qu'en ont données DEMPFLÉ et THOMPSON respectivement, ont été comparées d'un point de vue théorique par rapport au BLUP en vue de l'évaluation des pères avec plusieurs performances par fille.

Cette comparaison a été effectuée dans l'hypothèse d'une répartition des performances selon 2 facteurs; le père et le troupeau. Des expressions *a priori* de l'accroissement relatif de l'erreur de prédiction par rapport au BLUP, du biais occasionné par chacune de ces méthodes par rapport aux valeurs génétiques vraies d'un échantillon fixé de pères ont été établies. La convergence des itérations qui peuvent être pratiquées à un instant donné avec les méthodes de BAR ANAN, DEMPFLÉ et THOMPSON est également discutée. Quand la solution de départ est celle de la méthode de comparaison aux contemporaines, seule la méthode de THOMPSON converge vers le BLUP.

Un exemple numérique concernant l'évaluation de béliers *Rava* illustre cette comparaison.

---

### Introduction

Si parmi les méthodes d'évaluation de la valeur des reproducteurs, le BLUP d'HENDERSON (1973) est statistiquement le plus satisfaisant, d'autres méthodes, demandant moins de calculs, peuvent lui être préférées. La plus simple est sans doute la méthode de comparaison aux contemporaines de ROBERTSON et RENDEL (1954) qui supposent que les pères auxquels sont comparés chacun des jeunes reproducteurs à indexer aient tous des valeurs génétiques voisines. Pour lever cette hypothèse, qui, du fait de la sélection, apparaît de moins en moins fondée, BAR ANAN et SACKS (1974) proposent la méthode des « différences cumulées »

(\*) Abréviation consacrée de « Best linear unbiased predictor ».

dans laquelle cette comparaison aux contemporaines est corrigée par une estimation de ces valeurs génétiques. THOMPSON (1976) et DEMPFLÉ (1976), jugeant cette correction peu satisfaisante, en préconisent des modifications. Enfin, plusieurs autres méthodes ont été proposées, que nous ne discuterons pas ici (CUNNINGHAM, 1965; HENDERSON *et coll.*, 1954).

THOMPSON (1976) et DEMPFLÉ (1976) ont récemment comparé le BLUP, la comparaison aux contemporaines et la méthode des différences cumulées. Nous nous proposons ici de poursuivre cette comparaison, en particulier en dégageant des formules pour les valeurs caractéristiques des différentes estimations (espérances, variances, corrélations) et en analysant la convergence des itérations proposées par THOMPSON (1976) et DEMPFLÉ (1976) pour les méthodes approchées.

## I. — Description des méthodes

Nous nous placerons dans les conditions du modèle suivant :

$$Y_{ijkl} = \mu + T_i + H_j + F_{ijk} + E_{ijkl} \quad (1)$$

où  $Y_{ijkl}$  est la  $l^e$  performance de la  $k^e$  fille du  $i^e$  père réalisée dans le  $j^e$  troupeau

$T_i$  est l'effet du père  $i$  ( $i = 1, 2, \dots, p$ )

$H_j$  est l'effet du troupeau  $j$  ( $j = 1, 2, \dots, q$ )

$F_{ijk}$  est l'effet propre de la  $k^e$  fille du père  $i$  intervenant dans le  $j^e$  troupeau ( $k = 1, 2, \dots, r_{ij}$ )

$E_{ijkl}$  est l'effet résiduel propre à la  $l^e$  performance de cette fille ( $l = 1, 2, \dots, m_{ijk}$  pour la fille  $ijk$ ).

Nous supposons par ailleurs que :

les  $T_i$  sont indépendantes entre elles et de loi  $N(0, \sigma^2_T)$

les  $F_{ijk}$  sont indépendantes entre elles et de loi  $N(0, \sigma^2_F)$

les  $E_{ijkl}$  sont indépendantes entre elles et de loi  $N(0, \sigma^2_E)$

les  $T_i$ ,  $F_{ijk}$  et  $E_{ijkl}$  sont indépendantes entre elles.

Nous traiterons enfin que les effets « troupeau » comme fixés, conformément aux recommandations d'HENDERSON (1973) afin d'éliminer le biais dû à une répartition non aléatoire des pères selon ce facteur.

### A. — Modèle BLUP

Le modèle (1) s'écrit aussi :

$$y = X.t + Z.h + W.f + e \quad (2)$$

où  $y$  est le vecteur des observations

$t$  le vecteur des effets « père »

$h$  le vecteur des effets « troupeau »

$f$  le vecteur des effets propres aux filles

$e$  le vecteur des erreurs.

$X$ ,  $Z$  et  $W$  sont les matrices d'incidence correspondant respectivement aux effets  $t$ ,  $h$  et  $f$ .

Avec cette écriture :

$$\text{Var} (y) = V = X.{}^tX\sigma^2_T + W.{}^tW\sigma^2_F + I\sigma^2_E \tag{3}$$

Dans ces conditions les équations du modèle « mixte » qui permettent d'obtenir le BLUP sont les suivantes :

$$\begin{bmatrix} {}^tX.X + I.\frac{\sigma^2_E}{\sigma^2_T} & {}^tX.Z & {}^tX.W \\ {}^tZ.X & {}^tZ.Z & {}^tZ.W \\ {}^tW.X & {}^tW.Z & {}^tW.W + I.\frac{\sigma^2_E}{\sigma^2_F} \end{bmatrix} \begin{bmatrix} \widehat{t} \\ \widehat{h} \\ \widehat{f} \end{bmatrix} = \begin{bmatrix} {}^tX.y \\ {}^tZ.y \\ {}^tW.y \end{bmatrix} \tag{4}$$

L'absorption des équations relatives à *f* dans celles de *h* et *t* puis celles de *h* dans *t* conduit au système linéaire réduit de rang *p* (nombre de pères à indexer) s'écrivant ainsi :

$$(U + \lambda I).\widehat{t} = C$$

où  $\lambda = \frac{\sigma^2_E}{\sigma^2_T}$  et où U est une matrice (p x p) symétrique telle que :

$$U_{ii} = n_{i00} - \sum_j \frac{n^2_{ijo}}{n_{oj0}} \tag{6}$$

$$U_{i'j'} = - \sum_j \frac{n_{ijo} n_{i'j'o}}{n_{oj0}} \tag{7}$$

avec

$$n_{ijk} = \frac{m_{ijk} \alpha}{m_{ijk} + \alpha} \tag{8}$$

$$\alpha = \frac{\sigma^2_E}{\sigma^2_F}$$

Les symboles o et . indiquant une sommation et une moyenne sur l'indice correspondant à leurs places.

Quant au second membre, il a pour terme général

$$C_i = \sum_{jk} n_{ijk} Y_{ijk} . - \sum_j \frac{n_{ijo}}{n_{oj0}} \left( \sum_{ik} Y_{ijk} . n_{ijk} \right) \tag{9}$$

En effet, l'absorption des effets *f* propres aux filles conduit à

$$\sum_{jk} Y_{ijk0} - \sum_{jk} Y_{ijk0} \frac{m_{ijk}}{m_{ijk} + \alpha}$$

soit,

$$\sum_{jk} Y_{ijk0} \left( 1 - \frac{m_{ijk}}{m_{ijk} + \alpha} \right) = \sum_{jk} n_{ijk} Y_{ijk} . \tag{10}$$

Cette première absorption faite, celle relative aux effets *h* conduit à retrancher de (10)

$$\sum_j \frac{n_{ijo}}{n_{oj0}} \left( \sum_{ik} Y_{ijk} . n_{ijk} \right)$$

Enfin, on notera aussi la matrice  $U + \lambda I$ ,  $A + K$  où  $A$  est la matrice diagonale des  $U_{ii} + \lambda$  et  $K$  est formée des éléments non diagonaux  $U_{ii'}$  de  $U$ . La matrice  $A'$  sera la matrice diagonale des  $U_{ii}$  ( $A - A' = \lambda I$ ).

*B. — Méthode de comparaison aux contemporaines*

1) Dans cette méthode, le prédicteur  $I_i$  de la moitié de la valeur génétique additive  $G_i$  est donné par

$$I_i = CD_i \times D_i \quad (11)$$

où  $CD_i = \frac{1}{2} \frac{\text{Cov}(G_i, D_i)}{\text{Var}(D_i)}$  est la moitié du coefficient de régression de  $G_i$  en  $D_i$ , le prédicteur  $D_i$  étant l'écart moyen des performances des filles du père  $i$  à leurs contemporaines, écart calculé par la formule

$$D_i = \frac{\sum_j W_{ij} D_{ij}}{\sum_j W_{ij}} \quad (12)$$

où l'écart  $D_{ij}$  pour le troupeau  $j$  est donné par :

$$D_{ij} = \tilde{Y}_{ij} - c\tilde{Y}_{ij} \quad (13)$$

$\tilde{Y}_{ij} = \frac{\sum_k v_{ijk} Y_{ijk}}{\sum_k v_{ijk}}$  étant la moyenne pondérée des performances des filles du père  $i$  dans le troupeau  $j$  et  $c\tilde{Y}_{ij} = \frac{\sum_{i' \neq i} \sum_{k'} v_{i'jk'} Y_{i'jk'}}$  la moyenne de leurs contemporaines.

Ces moyennes sont pondérées en raison inverse de leur variance ( $\sigma^2 F + \sigma^2 E / m_{ijk}$ ). On prendra ici  $v_{ijk} = \frac{\sigma^2 F + \sigma^2 E}{\sigma^2 F + \sigma^2 E / m_{ijk}}$  qui peut s'écrire en fonction de la seule répétabilité intratroupeau et père  $\rho = \frac{\sigma^2 F}{\sigma^2 F + \sigma^2 E}$ , puisque :

$$v_{ijk} = \frac{m_{ijk}}{1 + \rho(m_{ijk} - 1)} \quad (14)$$

Les coefficients  $W_{ij}$  de (12) sont les moyennes harmoniques

$$\frac{1}{W_{ij}} = \frac{1}{v_{ijo}} + \frac{1}{\sum_{i' \neq i} \sum_{k'} v_{i'jk'}} \quad (15)$$

ce qui revient à pondérer chacun des écarts  $D_{ij}$  par l'inverse de sa variance.

Enfin, le coefficient de régression  $CD_i$  s'écrit :

$$CD_i = \frac{W_{io}}{W_{io} \left[ 1 + \frac{\sum_{i' \neq i} \left[ \sum_j \frac{v_{i'j_0} \cdot v_{ij_0}}{v_{oj_0}} \right]^2}{W_{io}^2} \right]} = 1 + \frac{4}{h^2} (1 - e^2) \quad (16)$$

$h^2$  étant l'héritabilité et  $e^2$  la part de la variance due aux troupeaux.

En pratique, on néglige le terme

$$\sum_{i' \neq i} \left( \sum_j \frac{v_{i'j_0} v_{ij_0}}{v_{oj_0}} \right)^2 / W_{io}^2 (*) \text{ et le prédicteur } Ii \text{ devient}$$

$$Ii = \frac{\sum_j W_{ij} D_{ij}}{W_{io} + \lambda'}$$

$$\text{avec } \lambda' = \frac{4}{h^2} (1 - e^2) - 1$$

## 2) Relations avec le BLUP.

En notant que (14) s'écrit aussi  $v_{ijk} = \frac{(1 + \alpha)m_{ijk}}{m_{ijk} + \alpha}$  on voit que d'après (8),

$$v_{ijk} = \frac{1 + \alpha}{\alpha} n_{ijk} \quad (17)$$

$$\text{Dans ces conditions, } W_{io} = v_{ioo} - \sum_j \frac{v_{ij_0}^2}{v_{oj_0}} = \frac{1 + \alpha}{\alpha} U_{ii} \quad (18)$$

$$\text{Par ailleurs } \sum_j W_{ij} D_{ij} = \sum_j \frac{v_{ij_0}(v_{oj_0} - v_{ij_0})}{v_{oj_0}} (\tilde{Y}_{ij} - c \tilde{Y}_{ij})$$

$$\text{et on trouve que } \sum_j W_{ij} D_{ij} = \sum_{jk} v_{ijk} Y_{ijk} - \sum_j \frac{v_{ij_0}}{v_{oj_0}} \left( \sum_{ik} v_{ijk} Y_{ijk} \right) \quad (19)$$

$$\text{Compte tenu de (9) et (19) on voit que } \sum_j W_{ij} D_{ij} = \frac{1 + \alpha}{\alpha} C_i \quad (20)$$

$$\text{Enfin, } \lambda' = \frac{4}{h^2} (1 - e^2) - 1 = \frac{1 - e^2 - h^2/4}{h^2/4} = \frac{\sigma^2 E + \sigma^2 F}{\sigma^2 T} = \lambda \frac{1 + \alpha}{\alpha} \quad (21)$$

Avec les relations (18), (20) et (21) on peut donc écrire

$$Ii = \frac{C_i}{U_{ii} + \lambda} \quad (22)$$

(\*) Dans le cas équilibré où chacun des  $p$  pères a dans chacun des  $q$  élevages  $r$  filles dont on connaît  $m$  performance :

$$\sum_{i' \neq i} \left( \sum_j \frac{v_{i'j_0} v_{ij_0}}{v_{oj_0}} \right)^2 = (p-1) \left( \frac{qrv}{p} \right)^2 \quad \text{avec } v = \frac{m}{1 + r(m-1)}$$

$$W_{io} = qrv \frac{p-1}{p} \quad \text{et} \quad \frac{\sum_{i' \neq i} \left( \sum_j \frac{v_{i'j_0} v_{ij_0}}{v_{oj_0}} \right)^2}{W_{io}^2} = \frac{1}{p-1}$$

On retrouve là l'interprétation de la méthode de comparaison aux contemporaines mentionnée notamment par DEMPFLÉ (1976) et THOMPSON (1976) dans l'étude de systèmes d'indexation ne faisant intervenir qu'une performance par descendant : les solutions obtenues sont celles que donnerait la méthode d'HENDERSON (1973) si on supprimait les éléments non diagonaux de la matrice conduisant au BLUP.

En termes matriciels (22) s'écrit  $\widehat{At}_c = C$ , soit

$$\widehat{t}_c = A^{-1} C \quad (23)$$

### C. — Méthodes des différences cumulées.

Ces méthodes consistent à corriger l'estimation de l'effet père obtenue par la méthode de comparaison aux contemporaines ( $\widehat{t}_{ci}$ ) par un facteur  $B_i = \sum_j W_{ij} \left( \sum_{i' \neq i} \frac{n_{i'j_0}}{n_{oj_0} - n_{ijo}} \widehat{t}_{1i'} \right)$  où  $\widehat{t}_{1i'}$  est, soit l'estimation  $\widehat{t}_{div}^*$  de la valeur du père  $i'$  calculée par la méthode des différences cumulées à la période précédente (BAR ANAN et SACKS, 1974; THOMPSON, 1976), soit l'estimation non régressée  $D_{i'}$  obtenue dans la méthode de comparaison aux contemporaines (DEMPFLÉ, 1976).

Ces méthodes rappellent ainsi un BLUP dans lequel l'effet du groupe serait le niveau génétique moyen des contemporaines de chaque père à indexer (POUTOUS, communication personnelle).

Compte tenu de (15) et (17), on trouve que :

$$B_i = \sum_j \frac{v_{ijo}(v_{oj_0} - v_{ijo})}{v_{oj_0}} \left( \sum_{i' \neq i} \frac{n_{i'j_0}}{n_{oj_0} - n_{ijo}} \widehat{t}_{1i'} \right)$$

$$B_i = \sum_j \frac{n_{ijo}}{n_{oj_0}} \sum_{i' \neq i} n_{i'j_0} \widehat{t}_{1i'}$$

En remarquant que  $\sum_j \frac{n_{ijo}}{n_{oj_0}} n_{i'j_0} = -U_{ii'}$ , on peut alors donner une interprétation matricielle des trois méthodes proposées

1) Méthode de Bar Anan et Sacks (1974).

$$\widehat{t}_{a_1} = A^{-1} \cdot C - A'^{-1} \cdot K \cdot \widehat{t}_{a_1}^* \quad (24)$$

où  $\widehat{t}_{a_1}$  et  $\widehat{t}_{a_1}^*$  sont les estimations des valeurs des pères pour la période actuelle et la période précédente.

2) Méthode de Thompson (1976).

$$\widehat{t}_{a_2} = A^{-1} \cdot C - A^{-1} \cdot K \cdot \widehat{t}_{a_2}^* \quad (25)$$

$\widehat{t}_{a_2}$  et  $\widehat{t}_{a_2}^*$  ayant la même signification que  $\widehat{t}_{a_1}$  et  $\widehat{t}_{a_1}^*$ .

3) *Méthode de Dempfle* (1976).

DEMPFLE (1976) corrige la déviation relative au père  $i$  à partir de celles relatives aux pères des contemporaines par :

$$\hat{\theta}_2 = A'^{-1} (C - KA'^{-1}C),$$

ce qui revient à résoudre le système  $A'\hat{\theta}_2 = C - K\hat{\theta}_1$  où  $\hat{\theta}_1 = A'^{-1}C$ .

$\hat{\theta}_2$  est ensuite régressé (multiplication de chaque  $\hat{\theta}_{2i}$  par  $\frac{W_{i0}}{W_{i0} + \lambda'}$ ) pour donner

$$\hat{t}_{a3} = A^{-1}A' \cdot \hat{\theta}_2.$$

soit 
$$\hat{t}_{a3} = A^{-1}C - A^{-1}KA'^{-1}C.$$

Ces méthodes ne sont pas directement comparables entre elles, ni à la méthode de comparaison aux contemporaines, puisque les unes font intervenir une information antérieure (BAR ANAN et SACKS, 1974 et THOMPSON, 1976) et les autres non. Cependant, on peut les comparer facilement pour la première année d'indexation. Dans cette situation, on aura :

$$\hat{t}_{a1} = A^{-1} \cdot C - A'^{-1} \cdot K \cdot A^{-1}C \quad (27) \text{ (BAR ANAN et SACKS)}$$

$$\hat{t}_{a2} = A^{-1} \cdot C - A^{-1} \cdot K \cdot A^{-1} \cdot C \quad (28) \text{ (THOMPSON)}$$

$$\hat{t}_{a3} = A^{-1}C - A^{-1} \cdot K \cdot A'^{-1} \cdot C \quad (29) \text{ (DEMPFLE)}$$

Rappelons ici que les autres méthodes s'écrivent :

$$\hat{t}_b = (A + K)^{-1} \cdot C \quad (30) \text{ (BLUP)}$$

$$\hat{t}_c = A^{-1} \cdot C \quad (23) \text{ (comparaison aux contemporaines)}$$

## II. — Estimation des valeurs caractéristiques

Dans le cas particulier cité ci-dessus (première année d'indexation), on va pouvoir calculer les valeurs caractéristiques des estimations des valeurs des pères obtenues par l'application de chacune des méthodes approchées. Pour ce faire, il suffit de donner les expressions *a priori* des différences  $\hat{t}_b - \hat{t}_a$ , où  $\hat{t}_a$  est l'estimation approchée des valeurs des pères ( $\hat{t}_{a1}$ ,  $\hat{t}_{a2}$ ,  $\hat{t}_{a3}$  ou  $\hat{t}_c$ ), en fonction de  $\hat{t}_b$  (BLUP) :

### A. — Expressions de $\hat{t}_b - \hat{t}_a$

#### 1) *Méthode de comparaison aux contemporaines.*

$$\hat{t}_c = A^{-1} \cdot C = A^{-1} \cdot (A + K) \cdot (A + K)^{-1}C = A^{-1} \cdot (A + K) \cdot \hat{t}_b$$

et on trouve

$$\hat{t}_b - \hat{t}_c = -A^{-1} \cdot K \cdot \hat{t}_b \quad (31)$$

2) *Méthode de Bar Anan et Sacks* (1974).

$$\widehat{t}_{a_1} = A^{-1}.C - A'^{-1}.K.A^{-1}C = (I - A'^{-1}.K)A^{-1}(A + K)\widehat{t}_b$$

Soit  $\widehat{t}_{a_1} = \widehat{t}_b + [(I - A'^{-1}.K).A^{-1}.K - A'^{-1}.K]\widehat{t}_b$

Alors  $\widehat{t}_b - \widehat{t}_{a_1} = (A'^{-1}.K - I + A'^{-1}.A).A^{-1}.K.\widehat{t}_b$

$$\widehat{t}_b - \widehat{t}_{a_1} = A'^{-1}(K - A' + A).A^{-1}.K.\widehat{t}_b$$

$$\widehat{t}_b - \widehat{t}_{a_1} = A'^{-1}(K + I.\lambda).A^{-1}.K.\widehat{t}_b \quad (32)$$

3) *Méthode de Thompson* (1976).

$$\widehat{t}_{a_2} = A^{-1}.C - A^{-1}KA^{-1}C = (I - A^{-1}.K).A^{-1}.(A + K).\widehat{t}_b$$

Soit  $\widehat{t}_b - \widehat{t}_{a_2} = A^{-1}.K.A^{-1}.K.\widehat{t}_b \quad (33)$

4) *Méthode de Dempfle* (1976).

On trouve de la même façon :

$$\widehat{t}_b - \widehat{t}_{a_3} = A^{-1}.K.A'^{-1}.(K + I.\lambda)\widehat{t}_b \quad (34)$$

On peut donc dans tous les cas écrire les différences  $\widehat{t}_b - \widehat{t}_a$  sous la forme  $\widehat{t}_b - \widehat{t}_a = H.\widehat{t}_b$ , la matrice H variant avec la méthode d'indexation.

### B. — Expressions a priori des valeurs caractéristiques.

1) *Espérances.*

HENDERSON (1973) a montré que les espérances des estimations du BLUP, si elles sont centrées sur l'ensemble de la population des pères, sont biaisées pour un ensemble de pères donné et sont telles que, dans nos notations :

$$E(\widehat{t}_b | t) = (I - (A + K)^{-1}.\lambda).t \quad (35)$$

On en déduit immédiatement les expressions suivantes :

$$E(\widehat{t}_b - \widehat{t}_a | t) = H.(I - (A + K)^{-1}.\lambda).t \quad (36)$$

et  $E(\widehat{t}_a | t) = (I - H).(I - (A + K)^{-1}.\lambda).t \quad (37)$

On remarquera que dans le cas de la méthode de comparaison aux contemporaines (37) se simplifie en :

$$E(\widehat{t}_c | t) = A^{-1}.U.t \quad (38)$$

2) *Accroissements de l'erreur quadratique moyenne de prédiction par rapport au BLUP dus à l'application des méthodes approchées.*

Ces accroissements relatifs peuvent être mesurés, pour chaque père  $i$ , par les rapports  $E[(\hat{t}_{bi} - \hat{t}_{ai})^2] / E[(\hat{t}_{bi} - t_i)^2]$  (\*).

On sait que  $E[(\hat{t}_{bi} - t_i)^2] = [(A + K)^{-1}]_{ii} \sigma^2 E$  (HENDERSON, 1973). Par ailleurs,  $E[(\hat{t}_{bi} - \hat{t}_{ai})^2] = E[(H \cdot \hat{t}_b)^2]_{ii} = [(H \cdot E(\hat{t}_b^2) \cdot {}^t H)]_{ii}$ , où  ${}^t H$  est la transposée de  $H$ .

Comme les estimateurs  $\hat{t}_b$  sont centrés  $V(\hat{t}_b) = E(\hat{t}_b^2)$ . D'autre part

$$V(\hat{t}_b) = \left[ \frac{I}{\lambda} \cdot I - (A + K)^{-1} \right] \cdot \sigma^2 E = Q \cdot \sigma^2 E \quad (39)$$

d'après HENDERSON (1973).

Dans ces conditions :

$$\frac{E[(\hat{t}_{bi} - \hat{t}_{ai})^2]}{E[(\hat{t}_{bi} - t_i)^2]} = \frac{[H \cdot Q \cdot {}^t H]_{ii}}{[(A + K)^{-1}]_{ii}} \quad (40)$$

3) *Corrélations entre les estimations obtenues par le BLUP et par les méthodes approchées.*

Ces corrélations qui sont, pour le  $i^e$  père,  $\rho_i = \frac{\text{Cov}(\hat{t}_{bi}, \hat{t}_{ai})}{\sqrt{v(\hat{t}_{bi}) \cdot v(\hat{t}_{ai})}}$  sont calculables

à partir des matrices de variances et covariances :

$V(\hat{t}_b) = Q \cdot \sigma^2 E$  d'après (39).

$\text{Cov}(\hat{t}_b, \hat{t}_a) = E(\hat{t}_a \cdot \hat{t}_b)$ . puisque  $E(\hat{t}_b) = 0$  et  $E(\hat{t}_a \cdot \hat{t}_b) = (I - H) \cdot V(\hat{t}_b)$

$V(\hat{t}_a) = (I - H) \cdot V(\hat{t}_b) \cdot (I - {}^t H)$

On a donc

$$\rho_i = \frac{[(I - H) \cdot Q]_{ii}}{\sqrt{Q_{ii} \cdot [(I - H) \cdot Q \cdot (I - {}^t H)]_{ii}}} \quad (41)$$

$$\text{où } Q = \frac{I}{\lambda} \cdot I - (A + K)^{-1}$$

### III. — Étude de la convergence des itérations préconisées par Thompson (1976) et Dempfle (1976)

Ces auteurs suggèrent de ne pas s'arrêter à la méthode des différences cumulées mais d'affiner les estimations en faisant des itérations. THOMPSON (1976) propose de calculer un  $l^e$  vecteur des effets pères à partir de la relation  $\hat{t}_{a_2} =$

(\*) En effet  $E[(\hat{t}_{ai} - t_i)^2] = E[(\hat{t}_{ai} - \hat{t}_{bi})^2] + E[(\hat{t}_{bi} - t_i)^2]$  puisque, d'après HENDERSON (1973)  $V(\hat{t}_b) = \text{Cov}(\hat{t}_b, t)$  donc  $E[(\hat{t}_{ai} - \hat{t}_{bi})(\hat{t}_{bi} - t_i)] = [H E(\hat{t}_b^2(t_b - t))]_{ii} = [H \cdot (E(\hat{t}_b^2) - E(\hat{t}_b \cdot t))]_{ii} = 0$ .

$A^{-1}.C - A^{-1}.K.{}_{l-1}\hat{t}_{a2}$ , la première estimation  ${}_{1}\hat{t}_{a2}$  étant, comme on l'a dit plus haut,  $\hat{t}_{a2}^*$  (effets pères obtenus au cours de la période précédente) ou, pour la première année d'indexation,  $\hat{t}_c = A^{-1}.C$  (estimation de la méthode de comparaison aux contemporaines).

Pour sa part, DEMPFLÉ (1976) fait des itérations sur les estimations non régressées des valeurs des pères : le  $l^e$  vecteur  ${}_{l}\hat{\theta}d$  est donné par  ${}_{l}\hat{\theta}d = A'^{-1}.C - A'^{-1}.K.{}_{l-1}\hat{\theta}d$  où  ${}_{1}\hat{\theta}d = A'^{-1}.C$ . Quand le nombre  $l$  d'itérations est assez grand, l'estimation  ${}_{l}a_3$  est alors calculée par  ${}_{l}a_3 = A^{-1}.A'.{}_{l}\hat{\theta}d_3$ .

En généralisant cette démarche à la méthode de BAR ANAN et SACKS (1974), les résultats de ces itérations s'écrivent :

$${}_{l}\hat{a}_1 = A^{-1}.C - A'^{-1}.K.(A^{-1}.C - A'^{-1}.K.(A^{-1}.C \dots (A^{-1}.C - A'^{-1}.K.{}_{1}\hat{t}_{a1}^*) \dots))$$

soit

$${}_{l}\hat{a}_1 = [I - A'^{-1}.K + (A'^{-1}.K)^2 + \dots + (-1)^{l-1}(A'^{-1}.K)^{l-1}] \cdot A^{-1}.C + (-1)^l.(A'^{-1}.K)^l.{}_{1}\hat{t}_{a1}^* \quad (42)$$

De même

$${}_{l}\hat{a}_2 = [I - A^{-1}.K + (A^{-1}.K)^2 + \dots + (-1)^{l-1}(A^{-1}.K)^{l-1}] \cdot A^{-1}.C + (-1)^l.(A^{-1}.K)^l.{}_{1}\hat{t}_{a2}^* \quad (43)$$

et

$${}_{l}\hat{a}_3 = A^{-1}.A'.[I - A'^{-1}.K + (A'^{-1}.K)^2 + \dots + (-1)^{l-1}.(A'^{-1}.K)^{l-1}] \cdot A'^{-1}.C + (-1)^l.(A'^{-1}.K)^l.A'^{-1}.C$$

On peut alors se demander si ces itérations convergent et vers quelles limites.

#### A. — Convergence

On voit que les suites d'estimations  ${}_{l}\hat{a}$  convergent si la matrice  $G^l = (-A'^{-1}.K)^l$  ou  $(-A^{-1}.K)^l$  tend vers 0 quand  $l$  tend vers l'infini. Dans ce cas, en effet, la série  $S = I + G + G^2 + \dots$  converge et a pour limite  $(I - G)^{-1}$ .

Nous notons  ${}_{i}g_{i'}$ , le terme  $(i, i')$  de la matrice  $G^l$ .

#### I. — Méthode de Thompson

La matrice  $G$  a pour éléments :

$${}_{i}g_{i'} = \frac{\sum_j \frac{n_{ijo}n_{i'jo}}{n_{ojo}}}{n_{i00} - \sum_j \frac{n_{ij0}^2}{n_{ojo}} + \lambda} \quad \text{pour } i \neq i'$$

$${}_{i}g_{ii} = 0$$

Comme  $n_{i00} - \sum_j \frac{n_{ij0}^2}{n_{ojo}} = \sum_j n_{ij0} \left( \frac{n_{ojo} - n_{ij0}}{n_{ojo}} \right) \geq 0$ ,  ${}_{i}g_{i'} \geq 0$  pour tout  $i$  et  $i'$ .

La matrice  $G^l$  tend vers 0 si chacun de ses termes  $ig_{ii}$ , tend vers 0. Or  ${}_{i+1}g_{ii'} = \sum_i ig_{ii} \cdot ig_{ii'}$ .

Si  $im_{i'}$  est la plus grande valeur des  $ig_{ii'}$  :  $im_{i'} = \text{Max}_i(ig_{ii'})$

$${}_{i+1}g_{ii'} \leq im_{i'} \sum_i ig_{ii}$$

La somme

$$\sum_i ig_{ii} = \frac{\sum_{i \neq j} \sum_j \frac{n_{ijo} \cdot n_{ijo}}{n_{ojo}}}{n_{i00} - \sum_j \frac{n_{ijo}^2}{n_{ojo}} + \lambda}$$

Or 
$$\sum_{i \neq j} \sum_j \frac{n_{ijo} \cdot n_{ijo}}{n_{ojo}} = \sum_j \frac{n_{ijo}}{n_{ojo}} (n_{ojo} - n_{ijo})$$

Donc  $\sum_i ig_{ii} < 1$  dès que  $\lambda$  est plus grand que 0. Alors :

$${}_{i+1}g_{ii'} < im_i$$

En particulier  ${}_{i+1}m_i < im_i$ . La suite des maxima, qui est minorée par 0, est décroissante et converge donc vers cette valeur. Les suites  $ig_{ii'}$ , convergent donc a fortiori vers 0, bien que d'une itération à la suivante, on puisse avoir pour certains couples  $(i, i')$   ${}_{i+1}g_{ii'} \geq ig_{ii'}$ .

Les itérations préconisées par THOMPSON (1976) convergent donc.

2. — Méthodes de Bar Anan et Sacks et de Dempfle.

La matrice  $G$  vaut dans ce cas  $-A^{-1} \cdot K$  et, avec un raisonnement analogue à ci-dessus, on voit que  $\sum_i ig_{ii'} = 1$  et que les suites  $ig_{ii'}$ , peuvent ne pas converger.

On peut d'ailleurs noter qu'ici la plus grande valeur propre de  $G$  est 1 (\*) et que  $G^l$  ne tend donc pas vers 0.

On retrouve ce résultat en remarquant que la matrice  $I - G = I + A^{-1} \cdot K$  soit  $A^{-1} \cdot (A' + K)$  n'est pas inversible : la matrice  $A' + K$  des moindres carrés est de rang  $p - 1$  ( $p$ , nombre de pères) puisqu'on a la relation  $\sum_{i'} U_{ii'} = 0$  pour tout

$i$ . DEMPFFLE (1977) a, cependant, montré que les solutions de son système d'itération convergent vers une limite qui satisfait au système des équations de moindres carrés :

$$(A' + K)\hat{\theta} = C$$

Il s'agit en fait de la solution particulière correspondant à la contrainte  $\sum_i U_{ii}\hat{\theta}_i = 0$ . De façon générale, il suffit de se donner une contrainte sur les valeurs des pères  $\hat{t}$ . Si on prend, par exemple,  $\hat{t}_1 = 0$ , ce qui est obtenu en supprimant la

(\*) Si  ${}_iU$  est un vecteur quelconque, de termes  ${}_oU_i$ , la plus grande valeur propre de  $G$  peut être obtenue par le rapport  $\lim_{l \rightarrow \infty} \frac{{}_{i+1}U_i}{{}_iU_i}$  où  ${}_{i+1}U = G \cdot {}_iU$  (PHAM, 1962). Si on prend pour  ${}_oU$  le vecteur  $(1, 1, \dots, 1)$  on voit que  ${}_iU = {}_oU$  pour tout  $l$ .

première ligne et la première colonne de  $A'$  et de  $K$ , la nouvelle matrice  $A' + K$  est inversible et on peut montrer que les itérations convergent.

Dans ce cas, en effet,

$$\sum_{i=1}^p \mathbf{1}g_{i'} = \sum_{i=2}^p \mathbf{1}g_{i'} = \mathbf{I} - \mathbf{1}g_{1'} < \mathbf{I}$$

B. — *Limites*

Quand  $l$  tend vers l'infini, en se plaçant dans une situation où  $(A' + K)^{-1}$  existe, on voit que les équations (42), (43) et (44) deviennent

$$\widehat{t_{a_1}}_{\infty} = (\mathbf{I} + A'^{-1}.K)^{-1}.A^{-1}.C = (A' + K)^{-1}.A'.A^{-1}.C \quad (45)$$

$$\widehat{t_{a_2}}_{\infty} = (A + K)^{-1}.C \quad (46)$$

$$\widehat{t_{a_3}}_{\infty} = A^{-1}.A'.(A' + K)^{-1}.C \quad (47)$$

La méthode de THOMPSON ( $\widehat{t_{a_2}}_{\infty}$ ) donne des solutions qui convergent donc vers le BLUP, alors que les itérations proposées par DEMPFLÉ ( $\widehat{t_{a_3}}_{\infty}$ ) reviennent à régesser les estimées des moindres carrés, procédé qui avait été préconisé par CUNNINGHAM (1965). Nous n'avons pas trouvé d'interprétation simple de la valeur  $\widehat{t_{a_1}}_{\infty}$  de BAR ANAN et SACKS.

On remarquera que si les effectifs sont assez grands, ou si l'héritabilité est élevée (donc  $\lambda$  petit), les matrices  $A$  et  $A'$  sont peu différentes et toutes ces méthodes doivent donner à peu près les mêmes valeurs.

Puisque  $E(C/t) = (A + K).E(\widehat{t_b}/t)$ , on a les espérances conditionnées des estimations :

$$E(\widehat{t_{a_1}}_{\infty}/t) = (A' + K)^{-1}.A'.A^{-1}.(A + K).[\mathbf{I} - (A + K)^{-1}.\lambda]t$$

Soit  $E(\widehat{t_{a_1}}_{\infty}/t) = (A' + K)^{-1}.A'.A^{-1}.(A + K).t \quad (48)$

$$E(\widehat{t_{a_2}}_{\infty}/t) = E(\widehat{t_b}/t) = [\mathbf{I} - (A + K)^{-1}.\lambda]t$$

Soit  $E(\widehat{t_{a_2}}_{\infty}/t) = (A + K)^{-1}.(A' + K).t \quad (49)$

Enfin  $E(\widehat{t_{a_3}}_{\infty}/t) = A^{-1}.A'.(A' + K)^{-1}.(A + K).[\mathbf{I} - (A + K)^{-1}.\lambda].t$

Soit  $E(\widehat{t_{a_3}}_{\infty}/t) = A^{-1}.A'.t \quad (50)$

ce qui revient à régesser les effets vrais des pères.

On peut calculer de la même façon  $E[(\widehat{t_{bi}} - \widehat{t_{ai}})^2]$  et  $\rho(\widehat{t_{bi}}, \widehat{t_{ai}})$ . (formules 40 et 41) avec ici

$$H = \mathbf{I} - (A' + K)^{-1}.A'.A^{-1}.(A + K) \quad (\text{BAR ANAN et SACKS})$$

$$H = \mathbf{I} - A^{-1}.A'.(A' + K).(A + K) \quad (\text{DEMPFLÉ})$$

## IV. — Exemple

Les valeurs des espérances conditionnées, des accroissements de variances dus à l'application des méthodes approchées dépendent de la structure de la population dans laquelle vont être indexés les pères. Il n'est donc pas possible de comparer ces méthodes de façon absolue, indépendamment de cette structure. A titre d'illustration, nous allons présenter ici les résultats obtenus sur un cas concret d'indexation de béliers de race *Rava*.

TABLEAU I

*Nombre de filles par élevage et par père*

N° père	N° élevage						Total
	1	2	3	4	5	6	
1	16	10	16				42
2	26	8					34
3	24	19			35		78
4	3			10	11		24
5			12	20			32
6					20	15	35
7	40						40
8	12						12
9	8						8
10		7					7
11		3					3
12		1					1
13		13					13
14		15					15
15			1				1
16			2				2
17			2				2
18			4				4
19			3				3
20			1				1
21			6				6
22				77			77
23				21			21
24				4			4
25				3			3
26					10		10
27					14		14
28					28		28
29						4	4
30						17	17
31						3	3
Inconnus	85	63		70			228
Total	214	139	47	205	128	39	772

A. — *Structure étudiée*

On compare 31 mâles utilisés dans 6 élevages différents. Parmi ces 31 béliers, 6 sont communs à plusieurs troupeaux, chacun des 25 restant n'ayant sailli de brebis que dans un élevage. En plus de ces 31 mâles, on trouve dans les troupeaux des filles de pères inconnus. L'information apportée par les performances de ces femelles a, malgré tout, été utilisée (\*).

Les nombres de filles par père et par élevage sont donnés au tableau 1. Les nombres de performances par fille varient de 1 à 5. Dans le troupeau 5, en cours de constitution, aucune femelle n'a plus d'une performance.

On ne donnera les résultats que pour quelques uns des béliers (mâles 3, 6, 7, 9, 16 et 31). Le choix de ces béliers a été fait sur la structure de leur descendance de façon à couvrir les situations les plus représentatives (tabl. 2).

Enfin on a supposé que la répétabilité entre performances d'une fille intra-troupeau et père était de 0,2 et on a étudié 2 valeurs de l'héritabilité (0,1 et 0,5).

TABLEAU 2

*Structures de la descendance des pères retenus pour la comparaison des méthodes*

Numéro du père	Nombre de filles de ce père	Nombre de troupeaux où il est utilisé	Nombre total de femelles auxquelles ses filles sont comparées	Nombre total de béliers utilisés dans les mêmes troupeaux que lui
3	++ (78)	++ (3)	++ (481)	++ (18)
6	+ (35)	+ (2)	+ (167)	+ (8)
7	+ (40)	— (1)	+ (214)	+ (6)
9	— (8)	— (1)	+ (214)	+ (6)
16	— — (2)	— (1)	— (47)	+ (8)
31	— — — (3)	— (1)	— (39)	— (4)

B. — *Résultats*

L'accroissement relatif de l'erreur quadratique moyenne de prédiction par rapport au BLUP est d'autant plus important que le coefficient d'héritabilité

(\*) On a supposé à cet effet que les valeurs des pères inconnus  $\hat{t}$  étaient nulles : il suffit alors, dans les termes de la matrice U de prendre  $n_{ojo}$  par sommation des  $n_{ijk}$  sur toutes les femelles du troupeau  $j$  et, dans le second membre  $C_i$ , de calculer  $\frac{1}{n_{ojo}} \sum_{ik} Y_{ijk} \cdot n_{ijk}$  pour toutes les femelles  $ik$  du troupeau  $j$ .

TABLEAU 3

Accroissements de l'erreur quadratique moyenne due à l'application des méthodes approchées

<i>Héritabilité 0,1</i>					
N° Père	Erreur moyenne $E[(\hat{t}_b - t)^2]$	Accroissement relatif de l'erreur moyenne : $E[(\hat{t}_a - \hat{t}_b)^2] / E[(\hat{t}_b - t)^2]$ (en %)			
		Comparaison aux contemporaines	Bar Anan et Sacks	Thompson	Dempfle
3	0,0074	2,5	0,9	0,4	13,9
6	0,0112	3,1	3,3	0,5	8,3
7	0,0093	4,2	1,1	0,3	2,2
9	0,0200	0,2	1,9	0,02	0,02
16	0,0229	0,02	22,2	0,00	0,05
31	0,0219	0,2	4,9	0,01	0,2
Moyenne des 31	0,0171	1,4	7,2	0,1	1,6

  

<i>Héritabilité 0,5</i>					
N° Père	Erreur moyenne $E[(\hat{t}_b - t)^2]$	Accroissement relatif de l'erreur moyenne : $E[(\hat{t}_a - \hat{t}_b)^2] / E[(\hat{t}_b - t)^2]$ (en %)			
		Comparaison aux contemporaines	Bar Anan et Sacks	Thompson	Dempfle
3	0,0096	30,8	20,0	18,3	53,9
6	0,0212	38,7	25,5	21,3	33,3
7	0,0124	51,0	8,1	8,4	9,5
9	0,0520	6,7	0,3	1,1	0,8
16	0,0832	6,3	13,6	0,4	0,8
31	0,0707	6,7	3,1	1,4	1,6
Moyenne des 31	0,0460	22,1	7,9	4,4	6,8

est élevé quelle que soit la méthode considérée (tabl. 3). La variation est particulièrement sensible pour la méthode de comparaison aux contemporaines (+ 22,1 contre + 1,4 %) et à un moindre degré pour celle de DEMPFLÉ (6,8 contre 1,6 %). Par contre, sur l'ensemble de l'échantillon étudié, on n'enregistre globalement dans la méthode de BAR ANAN et SACKS qu'un faible accroissement de cette variance (7,9 contre 7,2 %). Pour  $h^2 = 0,5$  l'accroissement de l'erreur quadratique moyenne s'avère élevé dans toutes les méthodes pour les pères largement représentés tels les pères n° 3, 6 et 7 (38,7; 33,3; 25,5 et 21,3 p. 100 par exemple pour le père n° 6 avec les méthodes de comparaison aux contemporaines, de DEMPFLÉ, de BAR ANAN et SACKS et de THOMPSON respectivement).

De façon plus générale, la supériorité du BLUP apparaît ainsi d'autant plus marquée que les pères sont mieux connus; cela est particulièrement sensible quand on le compare à la méthode de DEMPFLÉ qui se classe en dernière position de ce point de vue, pour les mâles 3,6 et 7 par exemple, aussi bien quand  $h^2 = 0,1$  que 0,5. La méthode de BAR ANAN et SACKS échappe quelque peu à cette règle puisqu'elle prédit également avec une erreur appréciable par rapport au BLUP les pères très peu représentés tels 16 et 31.

Dans tous les cas, et comme l'indiquent sous une autre forme les valeurs des corrélations (tabl. 4) la méthode de THOMPSON s'avère la meilleure par rapport au BLUP et celle de BAR ANAN et SACKS la plus imprécise. Les méthodes de comparaison aux contemporaines et de DEMPFLÉ sont assez voisines, la première pouvant être préférée à la seconde (et inversement dans le cas contraire) quand on contrôle beaucoup de descendants par père et que l'héritabilité est faible.

TABLEAU 4

Corrélation  $\rho(\hat{t}_{bi}, \hat{t}_{ai})$  entre les estimations des valeurs des pères obtenus par le BLUP et par chacune des méthodes approchées

<i>Hératibilité 0,1</i>				
N° Père	Corrélations entre les estimations BLUP et méthodes approchées			
	Comparaison aux contemporaines	Bar Anan et Sacks	Thompson	Dempfle
3	0,9952	0,9998	0,9998	0,9855
6	0,9921	0,9964	0,9994	0,9874
7	0,9888	0,9986	0,9997	0,9960
9	0,9953	0,9628	0,9997	0,9996
16	0,9893	0,4854	0,9999	0,9976
31	0,9947	0,8333	0,9997	0,9945
Moyenne des 31	0,9903	0,8572	0,9997	0,9948
<i>Hératibilité 0,5</i>				
N° Père	Corrélations entre les estimations BLUP et méthodes approchées			
	Comparaison aux contemporaines	Bar Anan et Sacks	Thompson	Dempfle
3	0,9878	0,9989	0,9985	0,9972
6	0,9697	0,9903	0,9894	0,9876
6	0,9736	0,9988	0,9979	0,9986
7	0,9774	0,9991	0,9962	0,9971
16	0,9450	0,8799	0,9963	0,9920
31	0,9631	0,9795	0,9911	0,9896
Moyenne des 31	0,9630	0,9664	0,9951	0,9939

## V. — Discussion-Conclusion

Pour déterminer les expressions algébriques des valeurs caractéristiques des estimations des effets pères obtenus par les différentes méthodes d'indexation comparées, nous nous sommes placés dans la situation très restrictive où aucune information antérieure n'est disponible : nous avons raisonné sur la première année d'indexation. En toute rigueur seules les méthodes de comparaison aux contemporaines et de DEMPFLÉ sont directement comparables au BLUP puisque les deux autres méthodes de différences cumulées (BAR ANAN et SACKS et THOMPSON) utilisent les estimations des pères obtenus l'année  $t-1$  pour corriger celles de l'année  $t$ . En pratique les méthodes de BAR ANAN et SACKS et de THOMPSON seront donc probablement meilleures qu'elles n'apparaissent au vue de nos résultats.

Les propriétés de convergence nous donnent cependant quelques informations sur la valeur relative de ces méthodes. En effet, si on se place dans la situation purement théorique où les mêmes mâles sont réindexés chaque année, sans apport d'information nouvelle, les solutions de BAR ANAN et SACKS et de THOMPSON après 2 années d'indexation seront celles de la 2<sup>e</sup> itération telles que nous les avons explicitées dans notre 3<sup>e</sup> partie. Or, seule la méthode de THOMPSON converge vers le BLUP. On peut donc en déduire que, dans le cas général où les matrices A et K et le second membre C varient d'une période à la suivante, la méthode de BAR ANAN et SACKS (et d'ailleurs aussi celle de DEMPFLÉ si on suppose que les estimations des effets pères obtenues dans le passé sont utilisées pour corriger celles de la période présente) sera toujours moins efficace que celle de THOMPSON.

Ces itérations à un instant donné nous paraissent d'ailleurs peu justifiées. En effet, si l'objectif est de parvenir au BLUP, on peut penser qu'il existe d'autres méthodes certainement plus efficaces pour résoudre au moindre coût un système d'équations linéaires, telle que celle de SCHAEFFER (1975). Sinon, quand faut-il s'arrêter?

En définitive, le choix entre les différentes méthodes est celui d'un compromis entre coût de calcul et qualité des résultats.

Si, comme dans le modèle étudié ici, la structure de répartition des données selon les facteurs de variation est simple et que le nombre de variantes de ces facteurs reste dans des limites acceptables (de l'ordre de la centaine), il ne fait pas de doute que la meilleure solution pratique est la résolution quasi directe des équations BLUP, après absorption des facteurs autres que les pères. Toutefois, en présence de certains facteurs tels que l'année ou la saison, adjoints au facteur troupeau, les calculs nécessités par les absorptions deviennent beaucoup plus complexes. Les résultats obtenus en considérant les facteurs « année » et « troupeau » comme additifs sont donnés en annexe. En présence d'une interaction (effet « troupeau  $\times$  année »), la matrice d'incidence de ces effets n'est plus diagonale après absorption des effets propres aux filles d'où le recours obligatoire au calcul numérique pour les absorptions ultérieures (SCHAEFFER, 1975).

Dans ces situations et lorsqu'en outre d'autres paramètres (indices des femelles, évaluation de différences de niveaux génétiques entre régions, etc...) sont à fournir aux unités de sélection, on peut s'interroger sur l'opportunité d'une méthode de calcul conventionnelle du BLUP. C'est alors que les autres méthodes reprennent de l'intérêt. La méthode de comparaison aux contemporaines est la plus simple; ses limites de validité sont maintenant bien connues. En particulier, s'il faut tenir compte du niveau génétique des pères des contemporaines et de l'évolution

du niveau génétique de la population dans le temps, le recours à une méthode des différences cumulées telle que celle de THOMPSON peut parfaitement se justifier vu sa bonne précision vis-à-vis du BLUP et ses propriétés de convergence qui laissent espérer que d'année en année les solutions données par cette méthode se rapprochent de plus en plus du BLUP.

*Reçu pour publication en juillet 1977.*

## Remerciements

Les auteurs tiennent à remercier M. POUTOUS (INRA-Jouy) pour la constante attention et les conseils qu'il leur a prodigués au cours de la réalisation de cette étude, le Dr DEMPFFLE (Freising), MM. COLLEAU et OLLIVIER (INRA-Jouy) pour leurs utiles critiques et suggestions à la lecture du manuscrit ainsi que M. BERNY (ITOVIC-Paris) pour avoir mis à leur disposition les données de l'exemple numérique.

## Summary

### *A priori relationships between BLUP, contemporary-comparison method and cumulative difference methods of sire evaluation*

The following sire evaluation methods: contemporary-comparison, BAR-ANAN and SACK's cumulative difference methods, two modifications of the latter (proposed by DEMPFFLE and THOMPSON) and BLUP were compared from a theoretical point of view.

Performance data were classified according to sire and herd and repeated records per daughter were considered. The increase in mean square prediction error relatively to BLUP and the bias of predicted from true sire values for a given sample of sires were determined.

The convergence of solutions obtained by using iteratively, at a given moment, methods of BAR-ANAN, DEMPFFLE and THOMPSON, was also discussed. When the initial solution is that of the contemporary-comparison method, only THOMPSON's method converges to BLUP.

A numerical application is given for rams of the *Rava* breed.

## Références bibliographiques

- BAR ANAN R., SACKS J. M., 1974. Sire evaluation and estimation of genetic gain in Israeli dairy herds. *Anim. Prod.*, **18**, 59-66.
- CUNNINGHAM E. P., 1965. The evaluation of sires from progeny test data. *Anim. Prod.* **7**, 221-231.
- DEMPFFLE L., 1976. A note on the properties of the cumulative difference method for sire evaluation. *Anim. Prod.*, **23**, 121-124.
- DEMPFFLE L., 1977. On the properties of two iterative methods for sire evaluation. *Z. Tierz. Zücht. Biol.* (in press).
- HENDERSON C. R., CARTER H. W., GOFREY J. T., 1954. Use of contemporary herd average in appraising progeny tests of dairy bull. *J. Anim. Sci.*, **13**, 959 (Abstr.).
- HENDERSON C. R., 1973. Sire evaluation and genetic trends. In *Proc. Anim. Breed. Genet. Symp.*, Blacksburg, Virginia p. 10-41. American Society of Animal Science, Champaign, Illinois.
- PHAM D., 1962. *Techniques du calcul matriciel*. 279 p. Collection universitaire de mathématiques, DUNOD, Paris.

- ROBERTSON A., RENDEL J. M., 1954. The performance of heifers got by artificial insemination. *J. Agric. Sci., Camb.*, **44**, 184-192.
- SCHAFFER L. R., 1975. *Dairy sire evaluation for milk and fat production*. Dept. of Animal and Poultry Science, University of Guelph, Ontario, Canada: 46 p.
- THOMPSON R., 1976. Relationship between the cumulative difference and best linear unbiased predictor methods of evaluating bulls. *Anim. Prod.*, **23**, 15-24.

### Annexe

En rajoutant le facteur « année » ( $A_h$ ) comme additif et fixé, le modèle décrit en (1) devient, avec les mêmes notations que précédemment :

$$Y_{ihjkl} = \mu + T_i + A_h + H_j + F_{ijh} + E_{ihjkl} \quad (51)$$

L'absorption successive des effets « fille » ( $F_{ijk}$ ) puis « troupeau » ( $H_j$ ) conduit au système suivant :

$$\begin{pmatrix} U_{11} + \lambda I & ; & U_{12} \\ U_{21} & ; & U_{22} \end{pmatrix} \begin{pmatrix} t \\ \hat{a} \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \quad (52)$$

1)  $U_{11}$  est une matrice  $(p,p)$  symétrique relative aux  $p$  effets « père » ( $t$ ) telle que :

$$(U_{11})_{ii} = n_{i000} - \sum_j \frac{n_{iojo}^2}{n_{oojo}} \quad (53)$$

$$(U_{11})_{ii} = - \sum_j \frac{n_{iojo}n_{i'ojo}}{n_{oojo}} \quad (54)$$

avec

$$n_{ihjk} = \frac{\alpha \cdot m_{ihjk}}{m_{iojk} + \alpha} \quad (55)$$

$$\alpha = \frac{\sigma^2 E}{\sigma^2 F}$$

$m_{ihjk}$  étant le nombre de performances de la  $k^{\text{ième}}$  fille du père  $i$  dans l'élevage  $j$ , l'année  $h$ .

2)  $U_{12}$  est une matrice  $(p,q)$  relative aux  $p$  effets « père » et  $q$  effets « année » telle que :

$$U_{12} = {}^t U_{21}$$

$$(U_{12})_{ih} = n_{ih00} - \sum_j \frac{n_{iojo}n_{ohjo}}{n_{oojo}} \quad (56)$$

3)  $U_{22}$  est une matrice  $(q, q)$  symétrique relative aux  $q$  effets « année » ( $\hat{a}$ ) telle que :

$$(U_{22})_{hh} = m_{ohoo} - \sum_{ijk} \frac{m_{ihjk}^2}{m_{iojk} + \alpha} - \sum_j \frac{n_{ohjo}^2}{n_{oojo}} \quad (57)$$

$$(U_{22})_{hh'} = - \sum_{ijk} \frac{m_{ihjk}m_{ih'jk}}{m_{iojk} + \alpha} - \sum_j \frac{n_{ohjo}n_{oh'jo}}{n_{oojo}} \quad (58)$$

4)  $C_1$  est un vecteur  $(p, I)$  du second membre relatif aux  $p$  effets « père » tel que :

$$(C_1)_i = \sum_{jk} n_{iojk} Y_{iojk} - \sum_j \frac{n_{iojo}}{n_{oojo}} \left( \sum_{ik} n_{iojk} Y_{iojk} \right) \quad (59)$$

5)  $C_2$  est un vecteur  $(q, I)$  du second membre relatif aux  $q$  effets « année » tel que :

$$(C_2)_h = Y_{ohoo} - \sum_{ijk} \frac{m_{ihjk}}{m_{iojk} + \alpha} Y_{iojko} - \sum_j \frac{n_{ohjo}}{n_{oojo}} \left( \sum_{ik} n_{iojk} Y_{iojk} \right) \quad (60)$$

Par ailleurs, du fait de la relation  $\sum_{h'} (U_{22})_{hh'} = 0$  tout pour  $h$ , il est nécessaire d'adjoindre au système (51) une relation supplémentaire sur les  $\hat{a}_h$ .