

## Relation entre BLUP (Best Linear Unbiased Prediction) et estimateurs bayésiens <sup>(1)</sup>

L. DEMPFLÉ

Lehrstuhl für Tierzucht der TU-München  
D-8050 Freising-Weißenstephan

---

### Résumé

Les relations entre le BLUP (Best Linear Unbiased Prediction) et la méthode bayésienne appliquée au modèle linéaire sont examinées dans cet article. On indique dans quel cas les deux méthodes sont identiques et on donne une approche simple et intuitive du BLUP.

---

### Introduction

Dans l'amélioration génétique des bovins laitiers on rencontre souvent une situation du type suivant : on dispose des performances en première lactation des vaches d'un département, enregistrées sur plusieurs années. Les vaches sont issues d'un grand nombre de taureaux, et ont vêlé dans des étables et à des années différentes. Les taureaux appartiennent à deux souches distinctes : *U.S. Holstein Friesian* et *Française Frisonne*. Un modèle linéaire pour la performance d'une vache en première lactation est alors :

$$y_{ijkl} = h_k + g_i + u_{ij} + e_{ijkl} \quad (1)$$

où

$y_{ijkl}$  est la performance de la vache  $ijkl$

$h_k$  est l'effet de l'étable-année-saison  $k$

$g_i$  est l'effet de la souche  $i$

$u_{ij}$  est l'effet du taureau  $j$  dans la souche  $i$

$e_{ijkl}$  est l'effet propre à la vache  $ijkl$

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

En utilisant une notation matricielle ce modèle s'écrit :

$$\mathbf{y} = \mathbf{X}_1 \mathbf{h} + \mathbf{X}_2 \mathbf{g} + \mathbf{Z} \mathbf{u} + \mathbf{e} \quad (1a)$$

$$= \mathbf{X} \mathbf{b} + \mathbf{Z} \mathbf{u} + \mathbf{e} \quad ; \quad \mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2] \quad ; \quad \mathbf{b}' = [\mathbf{h}' \mathbf{g}'] \quad (1b)$$

$$= \mathbf{W} \mathbf{t} + \mathbf{e} \quad ; \quad \mathbf{W} = [\mathbf{X} \mathbf{Z}] \quad ; \quad \mathbf{t}' = [\mathbf{b}' \mathbf{u}'] \quad (1c)$$

La matrice de variance-covariance de  $\mathbf{e}$  est dénotée par

$$\text{Var}(\mathbf{e}) = \mathbf{R} \quad (2)$$

On suppose, que les taureaux sont un échantillon aléatoire des populations constituées par les différentes souches. Avec cette hypothèse le vecteur  $\mathbf{u}$  est un vecteur aléatoire d'espérance zéro et de matrice de variance-covariance  $\mathbf{G}$ . Nous supposons également que de bons estimateurs de  $\mathbf{R}$  et  $\mathbf{G}$  existent.

On cherche à estimer les valeurs génétiques additives des taureaux. Ces valeurs sont définies par :

$$A_{ij} = 2w_{ij} = 2(g_i + u_{ij}) \quad (3)$$

La méthode des moindres carrés peut être utilisée pour l'estimation. Mais cette méthode n'utilise que l'information liée aux performances des vaches. Le fait que les taureaux constituent un échantillon aléatoire de populations bien définies est en lui-même un élément d'information supplémentaire. En effet si nous n'avions pas de performances sur les apparentés d'un taureau nous pourrions toujours utiliser comme estimation de la valeur génétique de ce taureau la moyenne de la population dont il a été tiré. Les méthodes suivantes utilisent cette information.

## Blup

Le BLUP (Best Linear Unbiased Prediction) a été défini par HENDERSON (1963, 1973) comme un estimateur qui est :

$$\text{linéaire : } \hat{w}_{ij} = \lambda_{ij} \mathbf{y} \quad (4)$$

$$\text{« non biaisé » : } E(\hat{w}_{ij}) = E(w_{ij}) \quad (5)$$

$$\text{le meilleur : } E(\hat{w}_{ij} - w_{ij})^2 \text{ est minimum} \quad (6)$$

Dans notre exemple nous considérons  $\mathbf{b}$  comme fixe. Le modèle est alors défini par :

$$\mathbf{y} = \mathbf{X} \mathbf{b} + \mathbf{Z} \mathbf{u} + \mathbf{e} \quad (7)$$

$$E(\mathbf{y}) = \mathbf{X} \mathbf{b} \quad (8)$$

$$\text{Var}(\mathbf{y}) = \mathbf{Z} \mathbf{G} \mathbf{Z}' + \mathbf{R} = \mathbf{V} \quad (9)$$

$$\text{Cov}(\mathbf{u}, \mathbf{y}) = \mathbf{G} \mathbf{Z}' \quad (10)$$

et le BLUP s'écrit (HENDERSON, 1973)

$$\hat{\mathbf{w}} = \mathbf{P} \hat{\mathbf{b}} + \text{Cov}(\mathbf{u}, \mathbf{y}) \text{Var}(\mathbf{y})^{-1} [\mathbf{y} - \mathbf{X} \hat{\mathbf{b}}] \quad (11)$$

$$= \mathbf{P} \hat{\mathbf{b}} + \mathbf{G} \mathbf{Z}' [\mathbf{Z} \mathbf{G} \mathbf{Z}' + \mathbf{R}]^{-1} [\mathbf{y} - \mathbf{X} \hat{\mathbf{b}}] = \mathbf{P} \hat{\mathbf{b}} + \hat{\mathbf{u}} \quad (12)$$

où

$$\hat{\mathbf{b}} = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1} \mathbf{y} \quad (13)$$

$\mathbf{P}$  est une matrice d'indicateurs et dans notre exemple l'estimation du taureau  $ij$  est

$$\hat{w}_{ij} = \hat{g}_i + \hat{u}_{ij} \tag{14}$$

On peut démontrer (HENDERSON, 1963) que  $\hat{\mathbf{b}}$  et  $\hat{\mathbf{u}}$  peuvent aussi être calculés plus facilement dans la plupart des cas par l'équation suivante :

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \tag{15}$$

*Modèle linéaire bayésien*

Certains aspects des relations entre le BLUP et la méthode bayésienne ont été discutés par RONNINGEN (1971) et SALOMON (1971). Nous utilisons ici la formulation de la méthode bayésienne donnée par LINDLEY et SMITH (1972), qui ont expliqué en détail l'emploi de cette méthode dans le cas d'un modèle linéaire. Dans un contexte bayésien il n'y a pas de différence entre fixe et aléatoire puisque chaque paramètre est considéré comme aléatoire. LINDLEY et SMITH utilisent un modèle hiérarchique pour ces paramètres. Dans l'hypothèse de **lois normales** on a

pour  $\mathbf{t}$  donné

$$\mathbf{y} \sim \mathbf{N}(\mathbf{W}\mathbf{t}, \mathbf{R}) \tag{16}$$

et pour  $\mathbf{r}$  donné, où  $\mathbf{r}$  est l'espérance de  $\mathbf{t}$

$$\mathbf{t} \sim \mathbf{N}(\mathbf{r}, \mathbf{M}) \tag{17}$$

Dans ce cas  $\mathbf{r}$  et  $\mathbf{M}$  sont appelés hyperparamètres et dans notre exemple

$$\mathbf{r} = \mathbf{E}(\mathbf{t}) = \mathbf{E} \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{0} \end{bmatrix} \tag{18}$$

et

$$\mathbf{M} = \text{Var} \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix} \tag{19}$$

La forme de  $\mathbf{M}$  indique que  $\mathbf{b}$  et  $\mathbf{u}$  sont indépendants.

Dans le langage traditionnel des statistiques bayésiennes, la fonction de vraisemblance pour  $\mathbf{y}$  donné s'écrit :

$$L(\mathbf{t}, \mathbf{y}) = k_1 e^{-\frac{1}{2} \{ (\mathbf{y} - \mathbf{W}\mathbf{t})' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{W}\mathbf{t}) \}} \tag{20}$$

La loi *a priori* de  $\mathbf{t}$  est :

$$f(\mathbf{t}) = k_2 e^{-\frac{1}{2} \{ (\mathbf{t} - \mathbf{r})' \mathbf{M}^{-1} (\mathbf{t} - \mathbf{r}) \}} \tag{21}$$

La loi *a posteriori* de  $\mathbf{t}$ , d'après le théorème de Bayes, est alors, à un coefficient de proportionnalité près :

$$f(\mathbf{t}, \mathbf{y}) \propto L(\mathbf{t}, \mathbf{y}) \cdot f(\mathbf{t}) \tag{22}$$

$$\propto e^{-\frac{1}{2} \{ (\mathbf{y} - \mathbf{W}\mathbf{t})' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{W}\mathbf{t}) + (\mathbf{t} - \mathbf{r})' \mathbf{M}^{-1} (\mathbf{t} - \mathbf{r}) \}}$$

Cette loi *a posteriori* est normale, donc symétrique et unimodale. Dans un contexte bayésien, notre inférence est basée seulement sur (22). En général le choix rationnel d'un estimateur de  $\mathbf{t}$  dépend de la fonction de perte,  $l(\hat{\mathbf{t}} - \mathbf{t})$ . Si la fonction de perte est  $l(\hat{\mathbf{t}} - \mathbf{t}) = \sum_i (\hat{t}_i - t_i)^2$  il est préférable de choisir la moyenne, si  $l(\hat{\mathbf{t}} - \mathbf{t}) = \sum_i |\hat{t}_i - t_i|$  il est préférable de choisir la médiane (FERGUSON, 1967, p. 46). Dans notre exemple la moyenne, la médiane et le mode de la loi *a posteriori* sont identiques, si bien, que plusieurs fonctions des pertes mènent au même estimateur. Pour calculer cet estimateur, le plus facile est de chercher le mode. Pour cela on dérive  $f(\mathbf{t}|\mathbf{y})$  par rapport à  $\mathbf{t}$  et on annule les dérivées.

On obtient les équations suivantes :

$$[\mathbf{W}'\mathbf{R}^{-1}\mathbf{W} + \mathbf{M}^{-1}] \hat{\mathbf{t}} = [\mathbf{W}'\mathbf{R}^{-1}\mathbf{y} + \mathbf{M}^{-1}\mathbf{r}] \quad (23)$$

Soit dans notre exemple :

$$\left\{ \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} \end{bmatrix} + \begin{bmatrix} \mathbf{S} & \mathbf{O} \\ \mathbf{O} & \mathbf{G} \end{bmatrix}^{-1} \right\} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} + \begin{bmatrix} \mathbf{S}^{-1}\mathbf{r}_1 \\ \mathbf{G}^{-1}\mathbf{O} \end{bmatrix} \quad (24)$$

Jusque là,  $\mathbf{S}$  et  $\mathbf{r}_1$  n'étaient pas spécifiés. Si nous ne disposons d'aucune information *a priori* nous pouvons exprimer cela en considérant que  $\mathbf{S}$  tend vers l'infini. Ce qui signifie que les composantes de  $\mathbf{b}$  sont distribuées sur l'ensemble des réels avec la même « probabilité ».  $\mathbf{b}$  n'a pas alors de loi propre, c'est-à-dire que la somme des probabilités des événements élémentaires n'est pas égale à 1, et l'estimateur obtenu est un estimateur bayésien généralisé (FERGUSON, 1967). Puisque  $\mathbf{S}$  tend vers l'infini, son inverse  $\mathbf{S}^{-1}$  tend vers  $\mathbf{O}$ . La valeur de  $\mathbf{r}_1$  n'a alors plus d'importance et nous obtenons la même solution qu'avec le BLUP.

#### Autre dérivation du BLUP

Une autre approche plus simple et plus intuitive peut être utilisée pour obtenir le BLUP.

Nous sommes en présence de deux estimateurs indépendants de  $\mathbf{t}$  :

1) l'estimateur des moindres carrés généralisés :

$$\hat{\mathbf{t}}_1 = (\mathbf{W}'\mathbf{R}^{-1}\mathbf{W})^{-1}\mathbf{W}'\mathbf{R}^{-1}\mathbf{y} \quad (25)$$

2) Un estimateur *a priori* :

$$\hat{\mathbf{t}}_2 = \mathbf{E}(\mathbf{t}) = \mathbf{r} \quad (26)$$

Les matrices de variance-covariance de ces deux estimateurs sont

$$\text{Var}(\hat{\mathbf{t}}_1) = (\mathbf{W}'\mathbf{R}^{-1}\mathbf{W})^{-1} + \mathbf{M} \quad (27)$$

$$\text{Var}(\hat{\mathbf{t}}_2) = \mathbf{O} \quad (28)$$

et la variance des erreurs, qui est une expression de la précision, est donnée par :

$$\text{Var}(\hat{\mathbf{t}}_1 - \mathbf{t}) = (\mathbf{W}'\mathbf{R}^{-1}\mathbf{W})^{-1} \quad (29)$$

$$\text{Var}(\hat{\mathbf{t}}_2 - \mathbf{t}) = \mathbf{M} \quad (30)$$

Quand on dispose de deux estimateurs indépendants, de précision inégale on en utilise généralement une moyenne pondérée, où la pondération est proportionnelle à l'inverse de la variance.

On obtient ainsi :

$$\hat{\mathbf{t}}_p = [\text{Var}(\hat{\mathbf{t}}_1 - \mathbf{t})^{-1} + \text{Var}(\hat{\mathbf{t}}_2 - \mathbf{t})^{-1}]^{-1} [\text{Var}(\hat{\mathbf{t}}_1 - \mathbf{t})^{-1}\hat{\mathbf{t}}_1 + \text{Var}(\hat{\mathbf{t}}_2 - \mathbf{t})^{-1}\hat{\mathbf{t}}_2] \quad (31)$$

$$= [\mathbf{W}'\mathbf{R}^{-1}\mathbf{W} + \mathbf{M}^{-1}]^{-1} [\mathbf{W}'\mathbf{R}^{-1}\mathbf{y} + \mathbf{M}^{-1}\mathbf{r}]$$

La comparaison entre (23) et (31) montre que l'estimateur bayésien et par conséquent le BLUP peuvent s'exprimer comme une moyenne pondérée de deux estimateurs indépendants : un estimateur de moindres carrés généralisés et un estimateur *a priori*. Cette façon d'obtenir le BLUP est similaire à celle utilisée par ROBERTSON (1955) pour la recherche d'équations de prédiction dans le cas d'un modèle purement aléatoire.

### Discussion

On doit noter qu'en général le BLUP et la méthode bayésienne (généralisée) donnent des estimateurs différents. Mais si les lois de  $\mathbf{u}$  et  $\mathbf{e}$  sont supposées normales et indépendantes et si l'information *a priori* sur  $\mathbf{b}$  est vague, le BLUP et la méthode bayésienne donnent le même estimateur. Différentes fonctions de perte peuvent alors être utilisées sans que les résultats en soient modifiés.

Le BLUP et la moyenne pondérée exprimée par (31) sont pour une situation donnée toujours identiques. Ils ne dépendent que des variances des lois.

Cette autre façon d'obtenir et d'interpréter le BLUP est particulièrement utile dans l'enseignement et pour certains aspects de la recherche, car elle permet mieux de cerner les hypothèses formulées pour le BLUP.

*Reçu pour publication en février 1977.*

### Remerciements

Je remercie L. OLLIVIER et J. CHESNAIS pour leurs commentaires sur ce manuscrit.

### Summary

#### *Relationship between BLUP (Best Linear Unbiased Prediction) and the Bayesian method*

The relationship between BLUP and the Bayes method for linear models is investigated. It is shown under what circumstances both methods lead to the same estimates. To derive BLUP an other approach has been used which is very simple and very intuitive and which gives some additional insight about the method.

### Références bibliographiques

- FERGUSON T. S., 1967. Mathematical statistics. A decision theoretic approach. *Academic Press*, New York, 50.
- HENDERSON C. R., 1963. Selection index and expected genetic advance. In *Statistical Genetics and Plant Breeding*. *NAS-NRC Publications*, **982**, 141-163.

- HENDERSON C. R., 1972. Sire evaluation and genetic trends. *Proc. anim. Breed. Genet. Symp.* in honor of J. L. Lush, Balcksburg, Virginia, 10-41.
- LINDLEY D. V., SMITH A. F. M., 1972. Bayes estimates for the linear model (with discussion). *J. Roy. Stat. Soc. B* **34**, 1-42.
- ROBERTSON A., 1955. Prediction equation in Quantitative Genetics. *Biometrics*, **11**, 95-98.
- RONNINGEN K., 1971. Some Properties of the Selection Index derived by « Henderson's Mixed Model Method ». *Z. Tierz. Züchtbiol.*, **88**, 186-193.
- SALOMON D. L., 1971. *A bayesian Interpretation and Extension of the genetic selection index*. Paper BU-362, Biometrics Unit, Cornell University, Ithaca, N. Y.
-