Predicting observed and latent responses to BLUP selection involving logistic variates

Jean-Jacques Colleau, Elisabeth Le Bihan-Duval

Genet. Sel. Evol. 32 (2000) 265–287

Original article

Predicting observed and latent responses to BLUP selection involving logistic variates

Jean-Jacques COLLEAU[a,*], Elizabeth LE BIHAN-DUVAL[b]

[a] Station de génétique quantitative et appliquée, Institut national de la recherche agronomique 78352 Jouy-en-Josas Cedex, France
[b] Station de recherches avicoles, Institut national de la recherche agronomique 37380 Nouzilly, France

(Received 19 May 1999; accepted 31 January 2000)

Abstract – Multivariate BLUPs can be derived when data are a mixture of continuous traits and observed discrete traits controlled by logistic latent traits. Algorithms were developed for predicting discrete responses to BLUP selection, and latent responses when the selection process included additional culling on scores. These algorithms were Taylor expansions using well-known expressions such as the probabilities and the two first moments of the truncated multinormal distribution, after appropriate re-parametrizations They were compared to very accurate quadrature integrations. The test examples were suggested by a situation found in chickens where selection can involve body weight and leg deformity described by two logistic latent variates. Quadratic Taylor expansions generally provided a good accuracy. Therefore, they could be recommended when quadrature methods are too demanding, *e.g.*, for complex breeding schemes.

logistic / BLUP / selection / response

Résumé – Prédiction des réponses observées et sous-jacentes à la sélection sur le BLUP. On peut calculer des BLUPs à partir de données mélangeant des caractères continus et des caractères discrets contrôlés par des variables logistiques. On développe des algorithmes de prédiction des réponses discrètes à la sélection sur le BLUP. De même, on cherche à prédire les réponses sur les variables latentes quand le processus de sélection inclut une sélection directe sur les variables discrètes en plus de la sélection sur le BLUP. Ces algorithmes sont des développements de Taylor qui utilisent des expressions bien connues, telles que la probabilité et les deux premiers moments d'une distribution multinormale tronquée, après modification appropriée des paramètres Ils sont comparés à des quadratures numériques très précises. Les exemples de test proviennent d'une situation rencontrée en sélection avicole où la sélection peut impliquer le poids corporel et les déformations des pattes, décrites par deux variables latentes logistiques. Les développements de Taylor du second degré sont

* Correspondence and reprints
E-mail. ugencjj@dga2 jouy.inra.fr

généralement précis. Par conséquent, ils pourraient être utilisés quand les méthodes de résolution par quadratures sont trop exigeantes en temps de calcul.

logistique / BLUP / sélection / réponse

1. INTRODUCTION

Real breeding schemes in livestock often rely on many traits. Some of them are clearly categorical: dichotomous such as fertility or mortality, polychotomous ordered such as "linear" type traits, calving difficulty, prolificity, conformation scores or unordered such as abnormality assessments. Much statistical work has been devoted to studying relevant models for these traits, quantifying the corresponding dispersion parameters and finally defining optimal procedures for estimating breeding values. The usual method that deals with such traits assumes the existence of latent variates expressed through a link function in a generalized linear mixed model [9,19]. In the threshold model, this function is related to the probability that the latent variate is higher than the threshold(s) concerned. In the multivariate logistic model, the link function is related to the probability that a given latent variate will be maximal among the whole set of latent variates [2,16,17]. In the context of the generalized linear mixed model, procedures for estimating breeding values were developed using either the estimates with maximum posterior probability density(MAP) [12, 13] or quasi-likelihood approaches [14]. The equation system used for estimation presents striking similarities with its counterpart for conventional BLUPs, except that mixed model equation (MME) matrices depend on the true position parameters. Solutions should be obtained after iteratively modifying these matrices.

Finding the optimal design of a breeding scheme requires accurate predictions of responses on the observed scale to given selection procedures. Prediction over several generations should include recurrent predictions, from one generation to the subsequent one, about the expectations and the variance-covariance matrix of the genetic latent variates. The last prediction is needed to take into account the linkage disequilibria induced by selection *i.e.*, the so-called Bulmer effect [5].

Authors often focus on predicting observed responses of discrete traits to BLUP selection. Exact solutions are given for instance for the threshold model and for the Poisson distributed model [10,11], assuming normality of the estimated breeding values (EBVs), although it might be questioned [6,7].

Logistic latent variates can be encountered for unordered polychotomous traits such as leg deformities, for binary traits or sequential discrete traits [9,19]. The objective of this paper is now to develop analytical approaches exploiting assumed multinormality of EBVs, in this logistic setting: (i) to predict discrete responses to selection on indices involving BLUPs; (ii) to predict responses on the first two moments of latent variates when the selection process is complicated by some direct culling on the corresponding observed scores, due to natural or artificial reasons.

It should be reminded that the forthcoming algorithms will address the issue of predicting genetic gains from future selection, with known initial dispersion

parameters. Hence, they will be quite distinct from posterior estimation of genetic trends, with unknown dispersion parameters, in an observed population with a past selection history, where the use of Markov chain Monte Carlo methods might be relevant.

2. PREDICTING SCORE PROBABILITIES AFTER BLUP SELECTION

Let us consider a situation where we want to predict score probabilities after BLUP selection. We want to predict the score probabilities of the selected candidates because, for instance, scores affect the length of breeding period. Alternatively, we might have considered the score probabilities for the future progeny of candidates based on BLUPs obtained from the sampling progeny.

2.1. Notations

A candidate for selection is described by a set of relevant variates: his breeding values and his EBVs for the continuous variates (latent variates and observed continuous traits). The joint distribution corresponds to the compound distribution of n_g Gaussian variates and of n_l latent logistic variates. The latter govern the expression of one categorical trait concerning the candidate, according to n_l+1 possible scores with probabilities π_i $(i = 1, \ldots, n_l+1)$. In the Gaussian-logistic model, n_l Gaussian variates correspond to the expectation of their logistic counterparts [16,17]. Therefore, the other $n_g - n_l$ Gaussian variates are the breeding values of continuous traits and the EBVs.

Let $\boldsymbol{x}$ be the vector of standardized Gaussian variates of dimension $n_g \times 1$ *i.e.*,

$$E(\boldsymbol{x}) = \mathbf{0} \qquad \mathrm{Var}(\boldsymbol{x}) = \boldsymbol{R}$$

where $\boldsymbol{R}$ is the correlation matrix. Then, in this logistic setting, we can write [17,18]:

$$\pi_i|\boldsymbol{x} = \frac{w_i}{1 + \sum_{j=1}^{n_l} w_j} \qquad i = 1, \ldots, n_l$$

$$\pi_{n_l+1}|\boldsymbol{x} = \frac{1}{1 + \sum_{j=1}^{n_l} w_j}$$

where the w_i terms are of the form $\exp\left[k + \boldsymbol{h}'\boldsymbol{x}\right]$ after considering a scalar k and a vector $\boldsymbol{h}$. Equivalently, for any $i = 1, \ldots, n_l + 1$, these probabilities can be re-written as follows:

$$\pi_i|\boldsymbol{x} = \frac{1}{\sum_{j=1}^{n_l+1} v_{ij}} \tag{1}$$

where $v_{ij} = \exp\left[k_{ij} + \boldsymbol{h}_{ij}'\boldsymbol{x}\right]$. In this expression, k_{ij} is a scalar and $\boldsymbol{h}_{ij}$ is a vector of dimension $n_g \times 1$. For any i, $k_{ii} = 0$, $\boldsymbol{h}_{ii} = \mathbf{0}$.

The selection procedure is summarized by the $(n_g \times 1)$ vector $\boldsymbol{t}$ of selection thresholds. If certain variates are not directly selected for, then the corresponding thresholds are set to $-\infty$.

2.2. Prediction algorithm

The selection domain S is defined by $x_i \geq t_i$ for $i = 1, \ldots, n_g$ and summarized by the notation $\boldsymbol{x} \geq \boldsymbol{t}$. We have to calculate score probabilities for the selected population *i.e.*, $E_S(\pi_j)$ for $j = 1, \ldots, 1+n_l$. They cannot be integrated out analytically even when the joint distribution of continuous variates is multinormal. Therefore, numerical methods should be investigated.

2.2.1. Numerical quadrature

The Gauss-Hermite quadrature approximates the integral $\int g(x)f(x)\mathrm{d}x$ by the finite sum

$$\sum_{s=1}^{n_s} \alpha_s f(x_s)$$

where the x_s terms are the nodes and the α_s terms are the corresponding weights, depending on the support function g involved [4,20]. The approximation is exact if $f(x)$ can be described by a polynomial of degree smaller than $2n_s$. If $g(x)$ is the Gaussian probability density function and if integration is not bounded, then the x_s terms are the roots of the Hermite polynomial of order n_s [4]. Numerical values of x_s and α_s can be found in [1]. This approach can be extended very easily to a multivariate function if the multiple integral can be decomposed into the product of simple integrals, after considering new independent variates (*e.g.*, after a Cholesky transformation of the variance-covariance matrix of the variates to be integrated out).

Here, the integrals considered describe the effect of selection among an infinite set of independent candidates. Then, this effect can be approximated by considering that the population submitted to selection is a mixture of a finite number of candidates with known breeding values and frequencies. In our case, the full Gauss-Hermite quadrature consists of calculating the probability that each configuration of unselected variates (each dummy candidate) be indirectly selected through the variates under culling. Then, the probability of each configuration in the selected population is calculated, which leads directly to the expectation of the π_i terms in this population.

2.2.2. Taylor expansions

As shown by expression (1), any score probability π_i can be written as the inverse of a function denoted $f_1(\boldsymbol{x})$ or f_1 in short, for the sake of simplicity.

If we are interested in calculating $E_S(\pi_i)$, then the function f_1 is equal to $\sum_{j=1}^{n_l+1} v_{ij}$. Its expectation $\tilde{f}_1$ in the selection domain

$$\tilde{f}_1 = E_S(f_1) = \sum_{j=1}^{n_l+1} E_S(v_{ij}) = \sum_{j=1}^{n_l+1} E(v_{ij}|\boldsymbol{x} \geq \boldsymbol{t}).$$

The Taylor expansion of $\pi_i = \frac{1}{f_1}$ up to the order n in the vicinity of $\tilde{f}_1$ is equal to:

$$\frac{1}{\tilde{f}_1}\left(1 + \sum_{i=1}^{n} \left(1 - \frac{f_1}{\tilde{f}_1}\right)^i\right).$$

The expectation of this expansion in the selection domain involves calculating expressions such as $E_S\left(f_1^i\right)$ for $i = 2, \ldots, n$. Finally, the expectation of the quadratic Taylor expansion is equal to $\frac{1}{\tilde{f}_1^3} E_S\left(f_1^2\right)$ and the expectation of the cubic expansion is equal to:

$$-\frac{2}{\tilde{f}_1} + \frac{4}{\tilde{f}_1^3} E_S\left(f_1^2\right) - \frac{1}{\tilde{f}_1^4} E_S\left(f_1^3\right).$$

The functions f_1^i are still similar to function f_1 itself because they can be written as linear combinations of expressions such as $\exp\left[k + \boldsymbol{h}'\boldsymbol{x}\right]$, where k is a constant and $\boldsymbol{h}$ is a vector. Finally, the problem amounts to calculating $E_S(\exp\left[\boldsymbol{h}'\boldsymbol{x}\right])$. This expression is simply the moment generating function of the truncated multinormal distribution [3,21].

$$E_S\left(\exp\left[\boldsymbol{h}'\boldsymbol{x}\right]\right) = \frac{\int_{t_1}^{\infty} \cdots \int_{t_{n_g}}^{\infty} \phi(\boldsymbol{x}) \exp\left[\boldsymbol{h}'\boldsymbol{x}\right] \mathrm{d}x_1 \cdots \mathrm{d}x_{n_g}}{\int_{t_1}^{\infty} \cdots \int_{t_{n_g}}^{\infty} \phi(\boldsymbol{x}) \, \mathrm{d}x_1 \cdots \mathrm{d}x_{n_g}}$$

where $\phi(\boldsymbol{x})$ is the probability density of vector $\boldsymbol{x}$ *i.e.*, the product of the integration constant $(2\pi)^{-\frac{n_g}{2}} \det \boldsymbol{R}^{-\frac{1}{2}}$ by $\exp\left[-\frac{1}{2}\boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{x}\right]$. As already mentioned by Tallis [21], if we pose $\boldsymbol{x}^* = \boldsymbol{x} - \boldsymbol{R}\boldsymbol{h}$, $\boldsymbol{t}^* = \boldsymbol{t} - \boldsymbol{R}\boldsymbol{h}$ with $\boldsymbol{x}^* \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$, then the nominator of the fraction is equal to:

$$\exp\left[\frac{1}{2}\boldsymbol{h}'\boldsymbol{R}\boldsymbol{h}\right] \int_{t_1^*}^{\infty} \cdot \cdot \int_{t_{n_g}^*}^{\infty} \phi(\boldsymbol{x}^*) \, \mathrm{d}x_1^* \cdot \cdot \, \mathrm{d}x_{n_g}^*.$$

Let us denote $\mathrm{Prob}(\boldsymbol{x} \geq \boldsymbol{t}; \boldsymbol{R})$ by $P(\boldsymbol{t}; \boldsymbol{R})$. Then,

$$E_S\left(\exp\left[\boldsymbol{h}'\boldsymbol{x}\right]\right) = \exp\left[\frac{1}{2}\boldsymbol{h}'\boldsymbol{R}\boldsymbol{h}\right] \frac{P(\boldsymbol{t} - \boldsymbol{R}\boldsymbol{h}, \boldsymbol{R})}{P(\boldsymbol{t}; \boldsymbol{R})}.$$

For calculating probabilities P, we can use the usual approximations such as the very accurate Dutt approximation [8].

Unfortunately, the accuracy of the expectations of Taylor expansions for functions such as $\frac{1}{f_1}$ is not certain to increase automatically along with the expansion order. The reason is that the Taylor expansion of function f_1 around $\tilde{f}_1$ is divergent when $f_1 \geq 2\tilde{f}_1$: in this case, $\left|1 - \frac{f_1}{\tilde{f}_1}\right|^i$ does not tend towards zero with an increasing i. This can be examplified by a very simple situation where function f_1 is x and where x is a discrete variate with three values x_1, x_2, x_3, with probabilities equal to $\frac{1}{3}$. Then

$$E\left(\frac{1}{f_1}\right) = \frac{1}{3}\sum_{i=1}^{3}\frac{1}{x_i}\cdot$$

When $x_1 = 2$, $x_2 = 3$, $x_3 = 4$, the correct value is obtained for a Taylor expansion of order 8. However, if $x_3 = 15$, then the Taylor expansion does not converge at all. The largest values of the ratio $\frac{x}{\tilde{x}}$ are 1.33 and 2.25 respectively. The coefficients of variation are 0.288 and 0.595 respectively.

3. PREDICTIONS ABOUT THE GAUSSIAN VARIATES AFTER SELECTION BASED ON BLUPS AND SCORES

In some cases, candidates can be naturally or artificially selected on scores in addition to BLUPs. For instance, selected cows should still be alive for further breeding, regardless of their EBVs. Poultry breeding schemes can discard candidates exhibiting leg deformities, at least for not impairing reproduction. In these cases, the first two moments of variates x can no longer be calculated exactly for the selected population.

3.1. First moments of the Gaussian variates

For instance, let i be the score required for selected candidates and let j be the Gaussian variate we are interested in. Let us denote S_i the selection domain. Then, using simplified notations:

$$E_{S_i}(x_j) = \frac{\int\limits_{S_i} x_j \pi_i}{\int\limits_{S_i} \pi_i} = \frac{E_S(x_j \pi_i)}{E_S(\pi_i)}$$

after dividing both members of the fraction by $P(\boldsymbol{t}; \boldsymbol{R})$ and calling S the selection domain without considering scores. Calculation by quadrature is straightforward, given the previously described guidelines.

The denominator of the second fraction was already calculated using a Taylor expansion. The nominator corresponds to the expectation of fraction $\frac{f_2}{f_1}$ where

function $f_2 = x_j$ and function $f_1 = \sum_j v_{ij}$, as defined previously. Let us denote:

$$\tilde{f}_1 = E_S(f_1) \qquad \tilde{f}_2 = E_S(f_2).$$

Then, the expectation of the quadratic expansion is:

$$\frac{\tilde{f}_2}{\tilde{f}_1} + \frac{1}{\tilde{f}_1^3}\left(\tilde{f}_2 E_S(f_1^2) - \tilde{f}_1 E_S(f_1 f_2)\right)$$

and the expectation of the cubic expansion is:

$$\frac{\tilde{f}_2}{\tilde{f}_1} + \frac{3}{\tilde{f}_1^3}\left(\tilde{f}_2 E_S(f_1^2) - \tilde{f}_1 E_S(f_1 f_2)\right) + \frac{1}{\tilde{f}_1^4}\left(\tilde{f}_1 E_S(f_1^2 f_2) - \tilde{f}_2 E_S(f_1^3)\right).$$

In these expressions, $\tilde{f}_2 = E_S(x_j)$ has an already known analytic solution [21] and $E_S(f_1 f_2)$ is a linear combination of terms such as $E_S\left(x_j \exp(\boldsymbol{h}'\boldsymbol{x})\right)$ where the $\boldsymbol{h}$ terms are known vectors. After changing variates as previously indicated, we get:

$$E_S\left(x_j \exp[\boldsymbol{h}'\boldsymbol{x}]\right) = \left(E_S \exp[\boldsymbol{h}'\boldsymbol{x}]\right)\left(h_j^* + E\left(x_j^* | \boldsymbol{x}^* \geq \boldsymbol{t} - \boldsymbol{h}^*\right)\right)$$

where $\boldsymbol{x}^* \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$ and $\boldsymbol{h}^* = \boldsymbol{R}\boldsymbol{h}$. The first term was previously calculated and the second term involves calculations similar to those of $\tilde{f}_2$ *i.e.*, the expectation of variates in a truncated multinormal distribution (See Appendix 1).

3.2. Second moments of the Gaussian variates

We need to calculate expressions such as:

$$E_{S_i}(x_j x_m) = \frac{E_S(\pi_i x_j x_m)}{E_S(\pi_i)}.$$

The nominator of this fraction is still $E_S\left(\frac{f_2}{f_1}\right)$ where $f_2 = x_j x_m$. It can be calculated using a Taylor expansion. $E_S(f_1 f_2)$ and $E_S(f_1^2 f_2)$ are linear combinations of expressions such as $E_S\left(x_j x_m \exp[\boldsymbol{h}'\boldsymbol{x}]\right)$, which is equal to:

$$\left(E_S \exp[\boldsymbol{h}'\boldsymbol{x}]\right)\ E_{\boldsymbol{x}^* \geq \boldsymbol{t} - \boldsymbol{h}^*}\left[(h_j^* + x_j^*)(h_m^* + x_m^*)\right]$$

using the afore-mentioned notations. The first term of the latter equation was previously calculated and the second one involves calculating first and second moments of the truncated multinormal distribution, the expressions of which are well-known [21] and reminded in Appendix 1.

4. TEST EXAMPLES

4.1. The situation corresponding to BLUPs

This situation was inspired, after some simplification, by the investigation of Le Bihan *et al.* [18] who analysed the joint distribution of latent variates for leg scores (score 1 = valgus abnormality, score 2 = varus abnormality, score 3 = normal) and of body weights in chickens. Independent sires were considered as candidates for selection. Each of them was evaluated based on the performances of n progeny recorded for a discrete trait with three categories and a continuous trait. The discrete variate could be interpreted as the expression of two latent logistic variates. The $(n \times 1)$ vectors of the expectations of the latent variates for the progeny of a given sire was $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$ and the vector of the observed performances for the continuous variate was $\boldsymbol{y}_3$. Let u_1, u_2, u_3 be the corresponding sire's breeding values for these variates. The (3×1) vector of the sire's breeding values was $\boldsymbol{u} \sim \text{MVN}\ (\boldsymbol{0},\ \boldsymbol{G})$. Then, we wrote:

$$\boldsymbol{\mu}_1 = \lambda_1 \left(\mathbf{1}_n \theta_1 + \mathbf{1}_n \frac{u_1}{2} + \beta_1 \boldsymbol{e}_3 \right)$$

$$\boldsymbol{\mu}_2 = \lambda_2 \left(\mathbf{1}_n \theta_2 + \mathbf{1}_n \frac{u_2}{2} + \beta_2 \boldsymbol{e}_3 \right)$$

$$\boldsymbol{y}_3 = \mathbf{1}_n \frac{u_3}{2} + \boldsymbol{e}_3.$$

Overall means θ_1, θ_2 were assumed to be known. We took $\theta_3 = 0$ without a loss of generality. β_1 and β_2 were regression coefficients assumed to be known and possibly different from 0: therefore, the continuous trait and the latent traits could be correlated at both levels *i.e.*, between and within-sire. The residuals, $\boldsymbol{e}_3$, affected individual risks and could be interpreted as random nuisance factors when estimating the sire's breeding values. Finally, λ_1 and λ_2 were scale factors defined as follows

$$\lambda_k = \left(\frac{\pi^2}{\pi^2 - 3\beta_k^2 \sigma_{e_3}^2} \right)^{\frac{1}{2}}.$$

This scale adjustment was carried out so that the residual variance of the model was still $\frac{\pi^2}{3}$ [16,18] *i.e.*, the variance of the standard logistic distribution.

Let p be the individual index $(p = 1, \ldots, n)$. Then, the probabilities of score 1 and 2 of the discrete variate were respectively

$$\pi_{1p} = \frac{\exp(\mu_{1p})}{1 + \exp(\mu_{1p}) + \exp(\mu_{2p})}$$

$$\pi_{2p} = \frac{\exp(\mu_{2p})}{1 + \exp(\mu_{1p}) + \exp(\mu_{2p})}.$$

Let us define the following variates

$$
\begin{aligned}
w_{ii p} &= \lambda_i^2 \pi_{ip}(1-\pi_{ip}) \qquad i = 1,2 \\
w_{12p} &= w_{21p} = -\lambda_1 \lambda_2 \pi_{1p} \pi_{2p} \\
w_{i3p} &= w_{3ip} = -\beta_1 w_{i1p} - \beta_2 w_{i2p} \\
w_{33p} &= \sum_{i=1}^{2} \sum_{i'=1}^{2} \beta_i \beta_{i'} w_{ii'p} \\
v_{ip} &= \lambda_i \left(n_{ip} - \pi_{ip}\right) \\
e_{3p} &= y_{3p} - \frac{1}{2} u_3
\end{aligned}
$$

where $n_{ip} = 1$ if individual p expresses score i, $n_{ip} = 0$ otherwise.

Le Bihan *et al.* [18] showed that the vector $\hat{\boldsymbol{u}}$ of MAP estimates was obtained iteratively by solving the system

$$
\boldsymbol{M}^{[t]} \left(\hat{\boldsymbol{u}}^{[t+1]} - \hat{\boldsymbol{u}}^{[t]} \right) = \boldsymbol{T}^{[t]} - \boldsymbol{G}^{-1} \hat{\boldsymbol{u}}^{[t]}
$$

where $\hat{\boldsymbol{u}}^{[t+1]}$ was the solution vector obtained at iteration $t+1$ starting from matrix $\boldsymbol{M}^{[t]}$ and vector $\boldsymbol{T}^{[t]}$ built from π_{ip} and e_{3p} estimated at iteration t, *i.e.*, $\hat{\pi}_{ip}^{[t]}$ and $\hat{e}_{3p}^{[t]}$ respectively. In this system, we had

$$
\boldsymbol{M}^{[t]} = 0.25 \begin{pmatrix} \sum_p \hat{w}_{11p}^{[t]} & \sum_p \hat{w}_{12p}^{[t]} & \sum_p \hat{w}_{13p}^{[t]} \\ & \sum_p \hat{w}_{22p}^{[t]} & \sum_p \hat{w}_{23p}^{[t]} \\ Sym & & \sum_p \hat{w}_{33p}^{[t]} + \dfrac{n}{\sigma_{e_3}^2} \end{pmatrix} + \boldsymbol{G}^{-1}
$$

and

$$
\boldsymbol{T}^{[t]} = \begin{pmatrix} 0.5 \sum_p \hat{v}_{1p}^{[t]} \\ 0.5 \sum_p \hat{v}_{2p}^{[t]} \\ 0.5 \sum_p \dfrac{\hat{e}_{3p}^{[t]}}{\sigma_{e_3}^2} - 0.5\beta_1 \sum_p \hat{v}_{1p}^{[t]} - 0.5\beta_2 \sum_p \hat{v}_{2p}^{[t]} \end{pmatrix}.
$$

Prediction algorithms directly considering the final solutions after convergence would be quite hard to develop. Then, the estimated breeding values used for prediction were those of the first iteration, starting from:

$$
\hat{\boldsymbol{u}}_1^{[0]} = \hat{\boldsymbol{u}}_2^{[0]} = \hat{\boldsymbol{u}}_3^{[0]} = \boldsymbol{0}, \quad \boldsymbol{e_3}^{[0]} = \boldsymbol{y_3}
$$

and $\hat{\boldsymbol{\pi}}_i^{[0]}$ calculated accordingly, for $t = 0$. Then, assuming multinormality, the expression of both $\mathrm{Var}(\hat{\boldsymbol{u}})$ or $\mathrm{Cov}(\hat{\boldsymbol{u}}, \boldsymbol{u})$ was equal to $\boldsymbol{G} - \left(\boldsymbol{M}_{[0]} + \boldsymbol{G}^{-1}\right)^{-1}$ [15]. Matrix $\boldsymbol{M}^{[0]}$ was obtained from null values for $\hat{\boldsymbol{u}}_1$, $\hat{\boldsymbol{u}}_2$, $\hat{\boldsymbol{u}}_3$ and $\hat{\boldsymbol{e}}_{3p}$.

4.2. The parameter values

Trait 1 *i.e.*, the latent variable for expression of score 1 of the discrete variate, was assumed to have a heritability equal to 0.2. The overall probability of score 1 was assumed to be 0.2. Trait 2 *i.e.*, the latent variable for expression of score 2, was assumed much more difficult to select for because heritability was only 0.05. On the contrary, trait 3 was assumed much more heritable: h^2 was equal to 0.4. Genetic variances were 0.692 1, 0.166 6, 1 for traits 1,2 and 3 respectively. Traits 1 and 2 were assumed to be genetically independent. Both traits were assumed to exhibit the same genetic correlation coefficient with trait 3, *i.e.*, +0.2. The number n of progeny per candidate was equal to 20. For each trait (1 or 2), the corresponding $\beta's$ were calculated so that the variance controlled by the covariate e_3 was twice as high as the variance controlled by the sire effect. Therefore, for $i = 1$ or 2,

$$\beta_i^2 \sigma_{e_3}^2 = 2\frac{\sigma_{u_i}^2}{4} \cdot$$

Then, the pair of β values was 0.392 and 0.192.

4.3. The predictions to be carried out

4.3.1. Predicting phenotypic scores

In the first illustrating exercise, we assumed that the phenotypic scores of sires were not available and we wanted to predict score phenotypes of sires selected based on their BLUPs. The logistic latent values corresponding to the phenotype of sires had expectations $\psi_1(\theta_1 + u_1)$ and $\psi_2(\theta_2 + u_2)$ respectively, where θ_1, θ_2 were constant mean effects and ψ_1, ψ_2 were scale effects so that the residual variance of the model was still $\frac{\pi^2}{3}$ [16,18]. Then, for latent variate 1 for instance,

$$h_1^2 = \frac{\psi_1^2 \sigma_{u_1}^2}{\psi_1^2 \sigma_{u_1}^2 + \dfrac{\pi^2}{3}} \cdot$$

Given the values of h_1^2 and h_2^2 used in [18], the values for ψ_1 and ψ_2 were 1.09 and 1.02 respectively. Let $x_1, \ldots, x_5$ be the standardized variates corresponding to $u_1, u_2, u_3, -\hat{u}_1, \hat{u}_3$ respectively. Table I shows the corresponding correlation matrix ($\boldsymbol{R}$). Table II shows the analytic expressions of the different score probabilities involving the scalars k and vectors $\boldsymbol{h}$ mentioned previously. A two-threshold selection was carried out on $-\hat{u}_1$ and $\hat{u}_3$ with two subsituations $t_4 = t_5 = 0$ and $t_4 = t_5 = 1.5$. Of course, $t_1 = t_2 = t_3 = -\infty$.

4.3.2. Predicting first and second moments of standardized breeding values

Here, score phenotypes of sires were known. For the sake of simplicity, these phenotypes were not incorporated into the EBVs so that Tables I and II could be re-used. The same selection procedure as in 4.3.1. was considered, but with the restriction that selected sires should exhibit normal score (*i.e.*, score 3).

Table I. Definition of the variates involved $x_1, \ldots, x_5$ and corresponding correlation matrix.

Variate	1	2	3	4	5
1	1	0	0.2	−0.601	0.131
2		1	0.2	0.006	0.165
3			1	−0.182	0.832
4				1	−0.218
5					1

$x_1 = u_1/\sigma_{u_1}$ = standardized latent breeding value for score 1.
$x_2 = u_2/\sigma_{u_2}$ = standardized latent breeding value for score 2.
$x_3 = u_3/\sigma_{u_3}$ = standardized breeding value for the observed continuous trait.
$x_4 = -\hat{u}_1/\sigma_{\hat{u}_1}$ = − standardized EBV for u_1.
$x_5 = \hat{u}_3/\sigma_{\hat{u}_3}$ = standardized EBV for u_3.

Table II. Full expression of score probabilities.

$$\pi_1 = \text{Prob(score 1)} = \frac{1}{1 + \exp(1.441 - 0.907x_1) + \exp(-1.321 - 0.907x_1 + 0.416x_2)}$$

$$\pi_2 = \text{Prob(score 2)} = \frac{1}{1 + \exp(1.321 + 0.907x_1 - 0.416x_2) + \exp(2.762 - 0.416x_2)}$$

$$\pi_3 = \text{Prob(score 3)} = \frac{1}{1 + \exp(-1.441 + 0.907x_1) + \exp(-2.762 + 0.416x_2)}$$

Score 1, 2 and 3 = valgus deformity, varus deformity and no deformity respectively.

5. TESTING GOODNESS OF THE PREDICTIONS

5.1. Accepting multinormality

Taylor expansions (quadratic and cubic) were compared to the values approximated by an accurate quadrature, because Hermite polynomials of degree 20 were used. The functions were evaluated for each combination of u_1, u_2 and the expectations were obtained after weighting the corresponding values. For each level of quadrature, 20 nodes [1] were sampled so that $(20)^2 = 400$ pairs of values u_1, u_2 were considered in the quadrature calculations.

5.2. Questioning multinormality

In a preliminary investigation, the complications due to score predictions or constraints were excluded: selection was carried out only on BLUPs. This allowed us to focus on the possible consequences of the joint distribution not

being strictly multinormal. We examined the potential of alternative prediction methods assuming that the joint distribution was a mixture of multinormal distributions with specific parameters.

We examined the genetic responses for each trait (*i.e.*, on u_1, u_2, u_3) for three selection pressures (50%, 10%, 1% respectively) and for three selection objectives (selection on $\hat{u}_1$, $\hat{u}_2$, $\hat{u}_3$ respectively).

The observed genetic responses were obtained based on sets of 40 000 independent candidates evaluated with the iterative MAP methodology using the relevant parameters.

The predicted genetic responses were obtained by three methods: the conventional method (C) assuming multinormality, method Q (quadrature approach) and method Q^* (mixed quadrature-Taylor expansion approach), where normality was assumed only locally. Methods Q and Q^* are described in Appendix 2 and 3 respectively. Let us call S_{ijk} the observed responses of trait k (k = 1 to 3) for selection objective j ($j = 1$ to 3) and selection pressure i ($i = 1$ to 3), expressed in corresponding genetic standard deviation. Corresponding predicted values were C_{ijk}, Q_{ijk}, Q^*_{ijk} for the conventional approach, the quadrature approach and the partial quadrature approach respectively. The overall prediction potential of the prediction method P (either C, Q, or Q^*) was assessed through the following quadratic error norms:

$$E_P^{[i]} = \frac{\sum_j \sum_k (P_{ijk} - S_{ijk})^2}{\sum_j \sum_k S_{ijk}^2}$$

established for each selection pressure i.

6. RESULTS AND DISCUSSION

6.1. Predicting categorical probabilities of selected sires

The results obtained are shown in Table III. The agreement between quadrature and quadratic Taylor expansions was fairly good for categories 2 and 3 and less good for category 1. The sum of the predictions for the three categories was not equal to 1. By contrast with the quadratic expansions, the cubic Taylor expansion gave quite misleading results for the prediction of probabilities for score 1. As mentioned previously, the probability of divergence for Taylor expansions was expected to increase along with the coefficient of variation of the function corresponding to the inverse of the score probability. Table IV shows that this coefficient of variation was definitely high for score 1 ($f_1 = \frac{1}{\pi_1}$) in both subsituations. For the other scores (2 and 3), extending the expansion to higher degrees led to predictions affected more and more by the high values of function f_1. For instance, in the first subsituation, the prediction of π_2 was 0.019 and 4.0×10^8, for $n = 5$ and 10 respectively. Corresponding values for π_3 were 0.775 and 2.2×10^6. In this circumstance, the use of only quadratic expansions seems to be the most robust procedure. Quadratic expansion could be recommended, except for the score probability exhibiting the highest coefficient of variation, which should be calculated indirectly, so that score probabilities sum to 1. The

Table III. Testing score probabilities of selected sires. Values between brackets for score 1 were obtained from values for scores 1 and 2. Bold values correspond to the proposed method.

Thresholds for x_4 and x_5	Method	Score probabilities		
		1	2	3
$t_4 = t_5 = 0$	Quadrature	0.153	0.057	0.789
	Taylor 2	0.172(**0.149**)	**0.058**	**0.793**
	Taylor 3	−0.094	0 051	0.785
$t_4 = t_5 = 1.5$	Quadrature	0.090	0.069	0.843
	Taylor 2	0.095(**0.088**)	**0.068**	**0.844**
	Taylor 3	−0.015	0 062	0.842

Table IV. Coefficient of variation of the inverse of score probabilities in the sire population selected for continuous variates.

Thresholds for x_4 and x_5	Score		
	1	2	3
$t_4 = t_5 = 0$	0.869	0.442	0.153
$t_4 = t_5 = 1.5$	0.829	0.414	0.086

results of the recommended procedure are shown in bold in Table III. In this case, accuracy was fairly good for any score in comparison with the quadrature.

6.2. Predicting moments of the constrained selected sire population

Table V shows the quadratic expansions of the first moment of variates u_1, u_2, u_3 in the population of sires selected above the thresholds and constrained to exhibit score 3 (*i.e.*, no leg deformity). The accuracy was reasonably good. Table VI shows that the same was true for the second moments. Using Taylor expansions in this kind of calculation already required a good accuracy of the Taylor expansion concerning the probability of the constraining category. This was clearly the case for score 3. Had score 1 been chosen for constraining the phenotype of the sires, then Prob (score 1) should have been calculated indirectly after considering the function 1 -Prob (score 2) -Prob (score 3).

6.3. Sensitivity to normality assumptions

The relative prediction error norms observed for the three prediction methods are shown in Table VII. The relative prediction error norms observed for the three prediction methods increased when the selection pressure became more and more severe. Methods Q and Q^* showed the same accuracy as method C except for the strongest selection pressure (1%) where the conventional

Table V. Testing first moments of the selected sire population constrained to exhibit no leg deformity.

Thresholds for x_4 and x_5	Method	E(x_1)	E(x_2)	E(x_3)
$t_4 = t_5 = 0$	Quadrature	−0.532	0.135	0.588
	Taylor 2	−0.545	0.136	0.586
	Taylor 3	−0.518	0.136	0.590
$t_4 = t_5 = 1.5$	Quadrature	−1.176	0.381	1.546
	Taylor 2	−1.181	0.381	1.546
	Taylor 3	−1 174	0.382	1.547

Table VI. Testing second moments of the selected sire population constrained to exhibit no leg deformity.

Thresholds for x_4 and x_5	Method	$E(x_1^2)$	$E(x_2^2)$	$E(x_3^2)$	$E(x_1x_2)$	$E(x_1x_3)$	$E(x_2x_3)$
$t_4 = t_5 = 0$	Quadrature	0.992	0.990	0.870	−0.073	−0.217	0.186
	Taylor 2	0.970	0.990	0 867	−0.079	−0.229	0.186
	Taylor 3	1.006	0.991	0.873	−0.070	−0.205	0.187
$t_4 = t_5 = 1.5$	Quadrature	2.036	1.110	2.775	−0 446	−1.730	0.667
	Taylor 2	2.037	1.108	2.773	−0.451	−1.739	0.666
	Taylor 3	2.035	1.110	2.776	−0.445	−1.726	0.668

Table VII. Preliminary study: relative error norms (10^{-4}) observed for the three prediction methods.

Prediction method	50% Selection	10% Selection	1% Selection
Conventional	48	56	258
Quadrature	52	61	193
Quadrature-Taylor	47	63	208

approach was found inferior. Table VIII shows the detailed predictions. If selection was carried out on continuous traits (selection on $\hat{u}_3$), then direct and indirect predicted responses could be done reliably by conventional methodology assuming multinormality of the vector of latent values and corresponding EBVs. Conversely, if emphasis was laid on the discrete traits (selection on $\hat{u}_1$ or $\hat{u}_2$), then conventional predictions were still reliable provided selection pressure was reasonable. In our preliminary simulation, substantial errors occurred only if a very severe (about a few % selected) selection pressure was favouring a rare score (score 2). The use of quadrature methods assuming mixture of distributions would be useful only for this case. Then, in the large majority of situations, conventional methods assuming multinormality can be used, or to be exact, are hard to be significantly improved.

Table VIII. Preliminary study: detailed genetic gains, in genetic standard units.

	Selection on $\hat{u}_1$			Selection on $\hat{u}_2$			Selection on $\hat{u}_3$		
	u_1	u_2	u_3	u_1	u_2	u_3	u_1	u_2	u_3
	0.487 (a)	0.016	0.199	−0.010	0.225	0.454	0.110	0.150	0.677
Simulation	1.134 (b)	−0.015	0.338	−0.064	0.466	0.827	0.215	0.306	1.456
(S)	1.753 (c)	0.033	0.519	−0.206	0.816	1.030	0.378	0.365	2.165
	0.480	−0.004	0.145	−0.010	0.202	0.433	0.105	0.132	0.664
Conventional	1.055	−0.010	0.319	−0.024	0.443	0.954	0.230	0.290	1.460
(C)	1.601	−0.015	0.484	−0.036	0.673	1.448	0.349	0.440	2.216
Full	0.490	−0.008	0.144	−0 016	0.201	0.436	0.094	0.127	0.668
quadrature	1.179	−0.036	0.250	−0.091	0.550	0.922	0.217	0.283	1.470
(Q)	1.834	−0.074	0.296	−0.206	0.966	1.338	0.371	0.459	2.194
Partial	0.486	−0.007	0.145	−0.014	0.200	0.438	0.105	0.132	0.664
quadrature	1.176	−0.035	0.253	−0.083	0.551	0.932	0.229	0.290	1.461
(Q^*)	1.834	−0.072	0.300	−0.187	0.971	1.362	0.346	0.439	2.215

(a) selection pressure: 50%. (b) selection pressure: 10%. (c) selection pressure: 1%.

7. CONCLUSION

The problem of predicting responses on logistic scores or on latent variates when culling on scores is practised can be addressed by using quadrature methods or Taylor expansions. In the latter case, standard tools for calculating probabilities of the truncated multinormal distribution and their corresponding two first moments can still be used after appropriate re-parametrizations. This would allow one to easily incorporate other constraints not studied here, so that the overall fraction of selected individuals will be equal to a specified value. The use of standard approaches such as a Newton-Raphson algorithm would then be straightforward.

The test examples analyzed a situation concerning a set of variates observed in chickens. In these examples, both numerical methods were easy to implement *i.e.*, were not too computationally demanding. This allowed us to observe that the accuracy of Taylor expansions might be considered as fairly reasonable provided that correcting procedures be carried out in some cases.

The formal algorithms would allow one to address a very wide set of situations which could range from a single-step selection up to a complete sequence of selection steps involving categorical scores. In this case, the EBVs and the scores (*e.g.*, survival) considered might depend on the ages of the breeding individuals during lifetime. Probabilities π_i would then represent the product of elementary probabilities at each selection step. In these complex cases, the overall number of selection thresholds and of variates to be integrated out might be high enough.

The first limiting factor for calculations whatever the approach chosen (quadrature or Taylor expansion) is the number of thresholds. When it exceeds 5 or 6, numerical integrals of the multinormal distribution are either inaccurate or very long to compute [8]. Furthermore, quadrature methods are very much demanding when the number of variates to integrate out is high because it affects computation times exponentially. Finally, Taylor expansions can be useful whenever quadrature methods would lead to a substantial computational burden. In practice, this might be the case for complex selection schemes with sequential discrete traits.

REFERENCES

[1] Abramowitz M., Stegun I.A., Handbook of mathematical functions, 9th edn., Dover Publications, New York, 1970.

[2] Albert J.H , Chib S , Bayesian analysis of binary and polychotomous response data, J. Am. Stat. Assoc. 88 (1993) 669–679.

[3] Anderson T.W., An introduction to multivariate statistical analysis, 2nd edn., John Wiley and Sons, New-York, 1984.

[4] Atkinson K.E , An introduction to numerical analysis, John Wiley and Sons, New-York, 1978.

[5] Bulmer M.G., The effect of selection on genetic variability, Am. Nat. 105 (1971) 201–221.

[6] Danell O , Ronningen K., All-or-none traits in index selection, Z. Tierzuchtg. Zuchtungsbiol. 93 (1981) 226–236.

[7] Dempster E.R., Lerner I.M., Robertson A., Heritability of threshold traits, Genetics 35 (1950) 212–236.
[8] Ducrocq V., Colleau J.J., Interest in quantitative genetics of Dutt's and Deak's methods for numerical computation of multivariate normal probability integrals, Génét. Sél Évol. 18 (1986) 447–474.
[9] Fahrmeir L., Tutz G., Multivariate statistical modelling based on generalized linear models, Springer-Verlag, New-York, 1994.
[10] Foulley J.L., Prediction of selection response for threshold dichotomous traits, Genetics 32 (1992) 1187–1194.
[11] Foulley J.L., Prediction of selection response for Poisson distributed traits, Genet. Sel. Evol. 25 (1993) 297–303.
[12] Foulley J.L., Gianola D., Thompson R., Prediction of genetic merit from data on binary and quantitative variates with an application to calving difficulty, birth weight and pelvic opening, Génét. Sél. Évol. 15 (1983) 401–424.
[13] Gianola D., Foulley J.L., Sire evaluation of ordered categorical data with a threshold model, Génét. Sél. Évol.,15 (1983) 201–224.
[14] Gilmour A., Anderson R.D., Rae A , The analysis of binomial data by a generalized linear mixed model, Biometrika 72 (1985) 593–599.
[15] Henderson C.R., Sire evaluation and genetic trends, in: Proceedings of the animal breeding and genetics symposium in honor of Dr J.L Lush, American Society of Animal Science, Champaign, IL, 1973, pp. 10–41.
[16] Johnson N.L., Kotz S., Distributions in statistics: continuous univariate distributions -2, John Wiley and Sons, New-York, 1970.
[17] Le Bihan-Duval E., Beaumont C., Colleau J.J., Genetic parameters of twisted legs syndrome in broiler chickens, Genet. Sel. Evol 28 (1996) 177–195.
[18] Le Bihan-Duval E , Beaumont C., Colleau J.J., Estimation of the genetic correlations between twisted legs and growth or conformation traits in broiler chickens, J. Anim. Breed. Genet. 114 (1997) 239–259.
[19] McCullagh P., Nelder J.A., Generalized linear models, 2nd edn., Chapman and Hall, London, 1992.
[20] Stoer J , Bulirsch R., Introduction to numerical analysis, Springer-Verlag, New-York, 1980.
[21] Tallis G.M., The moment generating function of the truncated multinormal distribution, J. R. Statist. Soc. B., 323 (1961) 223–229

APPENDIX 1

First and second moments of the truncated multinormal distribution

Let $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$. The selection domain S corresponds to $\boldsymbol{x} \geq \boldsymbol{t}$ with the probability $P(\boldsymbol{t}; \boldsymbol{R}) = \pi_S$. The moment generating function is:

$$\Psi(\boldsymbol{h}) = \exp\left[\frac{1}{2}\boldsymbol{h}'\boldsymbol{R}\boldsymbol{h}\right] \frac{P(\boldsymbol{t} - \boldsymbol{R}\boldsymbol{h}; \boldsymbol{R})}{P(\boldsymbol{t}; \boldsymbol{R})}.$$

Then,

$$E_S(x_i) = \left(\frac{\partial \Psi}{\partial h_i}\right)_{[\boldsymbol{h}=\boldsymbol{0}]} = \pi_S^{-1} \sum_j r_{ij} \phi(t_j) P(\boldsymbol{t}_{.j}; \boldsymbol{R}_{.j}).$$

In this expression, $\boldsymbol{t}_{.j}$ and $\boldsymbol{R}_{.j}$ are vectors of thresholds and matrix of correlations conditioned on threshold t_j and variate j respectively. $\phi()$ stands for the probability density of the standard Gaussian distribution. Furthermore,

$$E_S(x_i x_j) = \left(\frac{\partial^2 \Psi}{\partial h_i \partial h_j}\right)_{[\boldsymbol{h}=\boldsymbol{0}]}$$

$$= r_{ij} + \pi_S^{-1} \sum_k r_{ki} \phi(t_k)$$

$$\times \left[r_{kj} t_k P(\boldsymbol{t}_{.k}; \boldsymbol{R}_{.k}) + \sum_{q \neq k} r_{qj\,k} \sqrt{1 - r_{kj}^2}\, \phi(t_{q\,k}) P(\boldsymbol{t}_{.kq}; \boldsymbol{R}_{.kq}) \right].$$

APPENDIX 2

Preliminary study: full quadrature approach

Continuous variates to be integrated out were u_1, u_2, u_3 and the n e_{3p}s. To avoid combinatorial explosion between the corresponding roots, a good accuracy (*e.g.*, 6 roots) could be retained when integrating out the u terms whereas a low accuracy (2 roots) was accepted when integrating out the e_{3p} terms. Because the us were correlated, roots of the corresponding Cholesky transforms were calculated and then re-transformed into the original scale: this gave 216 ($= 6^3$) triplets of us with weights corresponding to the product of weights for roots of Cholesky transforms. Then, at least theoretically, each of these situations should have been subdivided into 2^n subsituations according to the number n^* of positive roots for e_3 (number of negative roots $= n - n^*$). Each subsituation should have been weighted by the expression $\left(\frac{1}{2}\right)^n \frac{n!}{n^*!n-n^*!}$ because $\frac{1}{2}$ is the weight of both roots of the quadratic Hermite polynomial (+1 and −1 respectively in standard units). In fact, an additional approximation was introduced by allowing n^* to vary only between two bounds n_L^* and n_H^* so that the total weight involved was almost 1 *e.g.*, 0.999 8. After eliminating these outlying values, the weights were re-transformed so that they would sum up to 1 again. Finally, the overall number of distinct situations to be investigated was $216(1 + n_H^* - n_L^*)$.

Let a basic situation be referenced by a quadruplet i, j, k, l of indices with i, j, k varying from 1 to 6 and l varying from 1 to $1 + n_H^* - n_L^*$. The corresponding matrix and right-hand side of the equation system used for estimating breeding values were $\boldsymbol{M}_{ijkl}^{[0]}$ and $\boldsymbol{T}_{ijkl}^{[0]}$ respectively. In the MME matrix, the w terms did not depend on u_{1i} and u_{2j}, the roots involved for u_1 and u_2. They indirectly depended on u_{3k} because $e_{3p}^{[0]} = y_{3p}$. Then, in this situation, n^* values of y_{3p} were equal to $\frac{u_{3k}}{2} + \sigma_{e_3}$ and $n - n^*$ values of y_{3p} were equal to $\frac{u_{3k}}{2} - \sigma_{e_3}$. We now had to calculate $E(\hat{\boldsymbol{u}})$ and $\mathrm{Var}(\hat{\boldsymbol{u}})$ in the situation $ijkl$.

We had:

$$E(\hat{\boldsymbol{u}}) = \boldsymbol{M}_{ijkl}^{[0]-1} E\left(\boldsymbol{T}_{ijkl}^{[0]}\right)$$

with

$$E\left(\boldsymbol{T}_{ijkl}^{[0]}\right) = \begin{pmatrix} 0.5\lambda_1 \sum_p \left(\pi_{1p} - \hat{\pi}_{1p}^{[0]}\right) \\ 0.5\lambda_2 \sum_p \left(\pi_{2p} - \hat{\pi}_{2p}^{[0]}\right) \\ \begin{bmatrix} 0.5\dfrac{(0.5nu_{3k}) + n^*\sigma_{e_3} - (n-n^*)\sigma_{e_3}}{\sigma_{e_3}^2} \\ -0.5\lambda_1\beta_1 \sum_p \left(\pi_{1p} - \hat{\pi}_{1p}^{[0]}\right) \\ -0.5\lambda_2\beta_2 \sum_p \left(\pi_{2p} - \hat{\pi}_{2p}^{[0]}\right) \end{bmatrix} \end{pmatrix}.$$

In this vector, the terms π_{1p} and π_{2p} depended on every kind of the roots corresponding to the situation. Terms $\pi_{1p}^{[0]}$ depended on root u_{3k}. Then, we had:

$$\mathrm{Var}(\hat{\boldsymbol{u}}) = \boldsymbol{M}_{ijkl}^{[0]-1} \quad \mathrm{Var}\left(\boldsymbol{T}_{ijkl}^{[0]}\right) \quad \boldsymbol{M}_{ijkl}^{[0]-1}.$$

The central variance matrix simply corresponded to the w terms of $\boldsymbol{M}_{ijkl}$ calculated based on the true probabilities π_{1p}, π_{2p} influenced by the different roots of the u terms and of the e_3 terms. Next, we were interested in the response to selection on the overall index $I = \boldsymbol{a}'\hat{\boldsymbol{u}}$ where $\boldsymbol{a}$ was a vector of economic weights. We selected candidates such that $I \geq I_0$ where I_0 was a predetermined threshold. After assuming that the local distribution of $\hat{\boldsymbol{u}}$ was still multinormal, although admitting that parameters might vary across situations, we could easily calculate the expected contribution of each "candidate" to the selected population. Then, we could calculate characteristics of this population such as the average breeding value and the genetic variance.

APPENDIX 3

Preliminary study: mixed quadrature-Taylor expansion approach

In this method, the numerical quadrature for the u terms was maintained but the effects of the different values of e_3 were integrated using linear and quadratic Taylor expansions of the following functions:

$$E(\hat{\boldsymbol{u}}|\boldsymbol{u}) = E_{\boldsymbol{e_3}}\left(E\left(\hat{\boldsymbol{u}}|\boldsymbol{u}, \boldsymbol{e_3}\right)\right)$$
$$\mathrm{Var}(\hat{\boldsymbol{u}}|\boldsymbol{u}) = \mathrm{Var}_{\boldsymbol{e_3}}\left(E\left(\hat{\boldsymbol{u}}|\boldsymbol{u}, \boldsymbol{e_3}\right)\right) + E_{\boldsymbol{e_3}}\left(\mathrm{Var}\left(\hat{\boldsymbol{u}}|\boldsymbol{u}, \boldsymbol{e_3}\right)\right).$$

These expansions were calculated for each triplet (u_1, u_2, u_3) involved in the quadrature, in the vicinity of $\boldsymbol{e_3} = \boldsymbol{0}$ after considering n independent variates e_{3p}. Because each of these variates brought the same information, we only needed to calculate the derivatives with respect to only one of them *e.g.*, e_{3n}.

(i) Calculating $E_{e_3}\left(E\left(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{e_3}\right)\right)$

This expectation was equal to:

$$E(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{0})+\frac{1}{2}n\sigma_{e_3}^2\left(\frac{\partial^2}{\partial_{e_{3n}}^2}E\left(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{e_3}\right)\right)_{[\boldsymbol{e_3}=\boldsymbol{0}]}.$$

For the sake of simplicity, let matrix $\boldsymbol{M}$ and vector $\boldsymbol{T}$ stand for $\boldsymbol{M}^{[0]}$ and $\boldsymbol{T}^{[0]}$ conditioned on the current u terms and on $\boldsymbol{e_3}=\boldsymbol{0}$. Then, $E(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{0})=\boldsymbol{M}^{-1}E(\boldsymbol{T})$. In the matrix $\boldsymbol{M}$, the w terms were influenced only by u_3 because terms $\hat{e}_{3p}^{[0]}=y_{3p}=\frac{1}{2}u_3$ when $e_{3p}=0$. Next,

$$E(\boldsymbol{T})=\begin{pmatrix}0.5\sum_p\lambda_1\left(\pi_{1p}-\hat{\pi}_{1p}^{[0]}\right)\\0.5\sum_p\lambda_2\left(\pi_{2p}-\hat{\pi}_{2p}^{[0]}\right)\\\begin{bmatrix}0.25\dfrac{nu_3}{\sigma_{e_3}^2}\\-0.5\sum_p\beta_1\lambda_1\left(\pi_{1p}-\hat{\pi}_{1p}^{[0]}\right)\\-0.5\sum_p\beta_2\lambda_2\left(\pi_{2p}-\hat{\pi}_{2p}^{[0]}\right)\end{bmatrix}\end{pmatrix}.$$

First and second derivatives could be obtained from first and second derivatives of $\boldsymbol{M}$ and $E(\boldsymbol{T})$: see Appendix 4.

(ii) Calculating $\mathrm{Var}_{e_3}\left(E(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{e_3})\right)$

The first order expansion of $E(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{e_3})$ around $\boldsymbol{e_3}=\boldsymbol{0}$ was equal to:

$$E\left(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{0}\right)+\left(\boldsymbol{1}_n'\boldsymbol{e_3}\right)\left(\frac{\partial}{\partial e_{3n}}E(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{e_3})\right)_{[\boldsymbol{e_3}=\boldsymbol{0}]}.$$

Then, $\mathrm{Var}_{e_3}E\left(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{e_3}\right)\approx n\sigma_{e_3}^2\boldsymbol{h}\boldsymbol{h}'$ where $\boldsymbol{h}=\left(\frac{\partial}{\partial e_{3n}}E\left(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{e_3}\right)\right)_{[\boldsymbol{e_3}=\boldsymbol{0}]}$. The expression of $\boldsymbol{h}$ is given in Appendix 4.

(iii) Calculating $E_{e_3}\left(\mathrm{Var}\left(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{e_3}\right)\right)$

We had: $\mathrm{Var}\left(\hat{\boldsymbol{u}}|\boldsymbol{u},\boldsymbol{e_3}\right)=\left(\boldsymbol{M}|\boldsymbol{u},\boldsymbol{e_3}\right)^{-1}\mathrm{Var}\left(\boldsymbol{T}|\boldsymbol{u},\boldsymbol{e_3}\right)\left(\boldsymbol{M}|\boldsymbol{u},\boldsymbol{e_3}\right)^{-1}$.

This variance was approximated by a quadratic expansion around $\boldsymbol{e_3}=\boldsymbol{0}$ *i.e.*:

$$\boldsymbol{M}^{-1}\mathrm{Var}(\boldsymbol{T})\boldsymbol{M}^{-1}+\frac{1}{2}n\sigma_{e_3}^2\left(\frac{\partial^2\boldsymbol{\Psi}}{\partial^2 e_{3n}}\right)_{[e_{3n}=0]}$$

with $\boldsymbol{\Psi} = \boldsymbol{M}^{-1} \operatorname{Var}(\boldsymbol{T}) \boldsymbol{M}^{-1}$ In this expression,

$$\operatorname{Var}(\boldsymbol{T}) = 0.25 \begin{pmatrix} \sum_p w_{11p} & \sum_p w_{12p} & \sum_p w_{13p} \\ & \sum_p w_{22p} & \sum_p w_{23p} \\ Sym & & \sum_p w_{33p} \end{pmatrix}$$

where the w_{ijp} terms were calculated from the true probabilities. The derivatives of function $\boldsymbol{\Psi}$ are shown in Appendix 5.

APPENDIX 4

Calculating the derivatives of $\boldsymbol{f} = E(\hat{\boldsymbol{u}}|\boldsymbol{u}, \boldsymbol{e_3}) = \boldsymbol{M}^{-1}E(\boldsymbol{T})$

Let e_{3n} be the last term of vector $\boldsymbol{e_3}$. Then,

$$\frac{\partial \boldsymbol{f}}{\partial e_{3n}} = \boldsymbol{M}^{-1} \left(\frac{\partial E(\boldsymbol{T})}{\partial e_{3n}} - \frac{\partial \boldsymbol{M}}{\partial e_{3n}} \boldsymbol{f} \right)$$

$$\frac{\partial^2 \boldsymbol{f}}{\partial^2 e_{3n}} = \boldsymbol{M}^{-1} \left(\frac{\partial^2 E(\boldsymbol{T})}{\partial^2 e_{3n}} - \frac{\partial^2 \boldsymbol{M}}{\partial^2 e_{3n}} \boldsymbol{f} - 2 \frac{\partial \boldsymbol{M}}{\partial e_{3n}} \frac{\partial \boldsymbol{f}}{\partial e_{3n}} \right)$$

after repeatedly using

$$\frac{\partial \boldsymbol{M}^{-1}}{\partial e_{3n}} = -\boldsymbol{M}^{-1} \frac{\partial \boldsymbol{M}}{\partial e_{3n}} \boldsymbol{M}^{-1}$$

and

$$\boldsymbol{M}^{-1}E(\boldsymbol{T}) = \boldsymbol{f}.$$

(a) First derivatives of $\boldsymbol{M}$

Term (i, i) for $i = 1, 2$ is equal to:

$$\frac{\partial \hat{w}_{iin}^{[0]}}{\partial e_{3n}} = \lambda_i \left(1 - 2\hat{\pi}_{in}^{[0]}\right) \left(\beta_1 \hat{w}_{i1n}^{[0]} + \beta_2 \hat{w}_{i2n}^{[0]}\right).$$

Term(1,2) is equal to:

$$\frac{\partial \hat{w}_{12n}^{[0]}}{\partial e_{3n}} = -\lambda_1 \hat{\pi}_{1n}^{[0]} \left(\beta_1 \hat{w}_{12n}^{[0]} + \beta_2 \hat{w}_{22n}^{[0]}\right) - \lambda_2 \hat{\pi}_{2n}^{[0]} \left(\beta_1 \hat{w}_{11n}^{[0]} + \beta_2 \hat{w}_{12n}^{[0]}\right).$$

Term(i,3) for $i = 1$, 2 and term (3,3) can easily be obtained from the previous terms (see definition of these terms given previously). In these expressions, the $\hat{\pi}_{in}$'s and $\hat{w}_{ijn}$'s correspond to individual n with performance $y_{3n} = \frac{1}{2}u_3$.

(b) Second derivatives of M

The corresponding terms can be obtained from the preceding ones and by noting that:

$$\frac{\partial \hat{\pi}_{in}^{[0]}}{\partial e_{3n}} = \frac{1}{\lambda_i}\left(\beta_1 \hat{w}_{i1n}^{[0]} + \beta_2 \hat{w}_{i2n}^{[0]}\right).$$

For example,

$$\frac{\partial^2 \hat{w}_{11n}^{[0]}}{\partial^2 e_{3n}} = \lambda_1\left(1 - 2\hat{\pi}_{1n}^{[0]}\right)\left(\beta_1 \frac{\partial \hat{w}_{11}^{[0]}}{\partial e_{3n}} + \beta_2 \frac{\partial \hat{w}_1 2^{[0]}}{\partial e_{3n}}\right) - 2\lambda_1 \frac{\partial \hat{\pi}_1^{[0]}}{\partial e_{3n}}\left(\beta_1 \hat{w}_{1i}^{[0]} + \beta_2 \hat{w}_{i2}^{[0]}\right)$$

(c) First derivatives of E (T)

$$\text{Term } 1 = 0.5\left[\beta_1\left(w_{11n} - \hat{w}_{11n}^{[0]}\right) + \beta_2\left(w_{12n} - \hat{w}_{12n}^{[0]}\right)\right].$$

$$\text{Term } 2 = 0.5\left[\beta_1\left(w_{12n} - \hat{w}_{12n}^{[0]}\right) + \beta_2\left(w_{22n} - \hat{w}_{22n}^{[0]}\right)\right].$$

$$\text{Term } 3 = -\beta_1(\text{Term } 1) - \beta_2(\text{Term } 2) + \frac{1}{\sigma_{e_3}^2}\cdot$$

In these expressions, the w_{ijn}'s are calculated from the true probabilities stemming from the local combinations u_1, u_2, u_3 and $e_{3n} = 0$. The $\hat{w}_{ijn}^{[0]}$'s consider probabilities with $\hat{u}_1 = \hat{u}_2 = 0$ and $\hat{e}_{3n} = \frac{1}{2}u_3$ d) second derivatives of E (T).

From (c) and (a), the calculations are straightforward. For instance, term 1 is equal to

$$0.5\left[\beta_1\left(\frac{\partial w_{11n}}{\partial e_{3n}} - \frac{\partial \hat{w}_{11n}^{[0]}}{\partial e_{3n}}\right) + \beta_2\left(\frac{\partial w_{12n}}{\partial e_{3n}} - \frac{\partial \hat{w}_{12n}^{[0]}}{\partial e_{3n}}\right)\right]$$

where the $\dfrac{\partial \hat{w}_{ijn}^{[0]}}{\partial e_{3n}}$ were previously defined and where the $\dfrac{\partial w_{ijn}}{\partial e_{3n}}$ are calculated based on the true probability.

APPENDIX 5
Calculating the derivatives of $\boldsymbol{\Psi} = \mathrm{Var}\,(\hat{\boldsymbol{u}}|\boldsymbol{u}, \boldsymbol{e_3})$

Then, $\boldsymbol{\Psi} = \boldsymbol{M}^{-1}\boldsymbol{V}\boldsymbol{M}^{-1}$ and $\frac{\partial^2 \boldsymbol{\Psi}}{\partial^2 e_{3n}} = \boldsymbol{M}^{-1}\frac{\partial^2 \boldsymbol{V}}{\partial^2 e_{3n}}\boldsymbol{M}^{-1} + \boldsymbol{\Psi}_1 + \boldsymbol{\Psi}_1'$ where

$$\begin{aligned}\boldsymbol{\Psi}_1 &= \frac{\partial^2 \boldsymbol{M}^{-1}}{\partial^2 e_{3n}}\boldsymbol{V}\boldsymbol{M}^{-1} + \frac{\partial \boldsymbol{M}^{-1}}{\partial e_{3n}}\boldsymbol{V}\frac{\partial \boldsymbol{M}^{-1}}{\partial e_{3n}} + 2\frac{\partial \boldsymbol{M}^{-1}}{\partial e_{3n}}\frac{\partial \boldsymbol{V}}{\partial e_{3n}}\boldsymbol{M}^{-1} \\ &= \left(2\boldsymbol{M}^{-1}\frac{\partial \boldsymbol{M}}{\partial e_{3n}}\boldsymbol{M}^{-1}\frac{\partial \boldsymbol{M}}{\partial e_{3n}} - \boldsymbol{M}^{-1}\frac{\partial^2 \boldsymbol{M}}{\partial^2 e_{3n}}\right)\boldsymbol{\Psi} \\ &\quad + \boldsymbol{M}^{-1}\frac{\partial \boldsymbol{M}}{\partial e_{3n}}\boldsymbol{\Psi}\frac{\partial \boldsymbol{M}}{\partial e_{3n}}\boldsymbol{M}^{-1} - 2\boldsymbol{M}^{-1}\frac{\partial \boldsymbol{M}}{\partial e_{3n}}\frac{\partial \boldsymbol{V}}{\partial e_{3n}}\boldsymbol{M}^{-1}\end{aligned}$$

The derivatives of $\boldsymbol{M}$ were given previously and the derivatives of $\boldsymbol{V}$ can be obtained from the definition of $\boldsymbol{V} = \mathrm{Var}(\boldsymbol{T})$ in paragraph 2.3.3 and derivatives of terms w_{ijn} given in Appendix 4(a) and 4(b).