THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Predicting rates of inbreeding for livestock improvement schemes

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher final version (usually the publisher pdf)

Published In:
Journal of Animal Science

OPEN ACCESS

Predicting rates of inbreeding for livestock improvement schemes.
P Bijma, J A Van Arendonk and J A Woolliams

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://www.journalofanimalscience.org/content/79/4/840

American Society of Animal Science

www.asas.org

Predicting rates of inbreeding for livestock improvement schemes[1]

P. Bijma*,[2], J. A. M. Van Arendonk* and J. A. Woolliams†

*Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen University, Wageningen, The Netherlands and †Roslin Institute (Edinburgh), Roslin Midlothian, EH25 9PS, United Kingdom

ABSTRACT: This article presents a deterministic method to predict rates of inbreeding (ΔF) for typical livestock improvement schemes. The method is based on a recently developed general theory to predict rates of inbreeding, which uses the concept of long-term genetic contributions. A typical livestock breeding population was modeled, with overlapping generations, BLUP selection, and progeny testing of male selection candidates. Two types of selection were practiced: animals were either selected by truncation on estimated breeding values (EBV) across age classes, or the number of parents selected from each age class was set to a fixed value and truncation selection was practiced within age classes. Bulmer's equilibrium genetic parameters were obtained by iterating on a pseudo-BLUP selection index and ΔF was predicted for the equilibrium situation. Predictions were substantially more accurate than predictions from other available methods, which ignore the effect of selection on ΔF. Predictions were accurate for schemes with up to 20 sires. Predicted ΔF was somewhat too low for schemes with more than 20 sires, which was due to the use of simple linear models to predict genetic contributions. The present method provides a computationally feasible (i.e., deterministic) tool to consider both the rate of inbreeding and the rate of genetic gain when optimizing livestock improvement schemes.

Introduction

When optimizing breeding schemes, breeding companies have to consider both short- and long-term effects of selection decisions. In the short term, breeding companies require a sufficiently high rate of genetic gain (ΔG) to either maintain or strengthen their competitive position. In the long term, maintenance of genetic variance and avoidance of inbreeding depression are important and require a restriction of the rate of inbreeding. Particularly when applying truncation selection on BLUP-EBV, rates of inbreeding may become very high, which reduces long-term response (e.g., Verrier et al., 1993) and clearly merits attention.

Optimization of breeding schemes over a wide range of alternatives requires methods to assess rates of gain and inbreeding that take little computing time. Deterministic prediction methods, therefore, are preferred, rather than stochastic simulation. Methods to deterministically predict rates of genetic gain are based on selection index theory and are well-established (e.g., Wray and Hill, 1989). Recently, Woolliams et al. (1999) and Woolliams and Bijma (2000) developed a general theory to predict rates of inbreeding in populations undergoing selection. However, explicit prediction equations that can be applied directly to typical livestock improvement schemes have not yet been developed.

The objective of this research, therefore, was to derive equations to predict rates of inbreeding for typical livestock improvement schemes. For this purpose, we modeled a population with overlapping generations, truncation selection on BLUP-EBV, and progeny testing of male candidates for selection. Subsequently, equations to predict the rate of inbreeding for these populations were derived on the basis of the general theory presented by Woolliams et al. (1999) and Woolliams and Bijma (2000).

[2]Correspondence: Department of Animal Sciences, P.O. Box 338, Marijkeweg 40, 6700 AH Wageningen (phone: +31 317 482731/482335; fax: +31 317 483929; E-mail: piter.bijma@alg.vf.wau.nl).

840

Table 1. Notation used

Equation[a]	Symbol	Definition
P	N_m, N_f, d, n_{prg}	Number of sires, number of dams, mating ratio $d = N_f/N_m$, number of progeny per sire
P	n_o, T	Number of selection candidates born per dam, total number of candidates per sex
P	c_{max}, n_k	Maximum age of parents, number of parents selected from category k
6	$\mathbf{N}$	$2c_{max} \times 2c_{max}$ diagonal matrix of the number of parents selected from each category k
P, G	k, s	Indicator for sex-age class categories, indicator for lifetime categories
B, S$_2$	p_k, τ_k	Selected proportion and standardized truncation point for category k
A	l_k, κ_k	Selection intensity, variance reduction coefficient for category k
P, PB	$P, A, \hat{A}$	Phenotype, breeding value, estimated breeding value (EBV)
PB	$\mathbf{b}_1(\mathbf{b}_2), \mathbf{x}_{1,i}(\mathbf{x}_{2,i})$	6×1 (7×1) vector of index weights, 6×1 (7×1) vector of index information sources
A	$\sigma_A^2, \sigma_{\hat{A}}^2$	Additive genetic variance, variance of the estimated breeding value
A, R	ρ, h^2	Accuracy of selection, heritability
A	$\overline{\rho}_{FS,kl} (\overline{\rho}_{HS,kl})$	Sample correlation between EBV of full-sibs (half-sibs) in category k and l
I, G	$\Delta F, r_{i,k}$	Rate of inbreeding, long-term genetic contribution of individual i in category k
2, 12	$s_{i,k}, \sigma_{s_k}^2$	Selective advantage of individual i in category k, variance of $s_{i,k}$
4	$u_{i,k}, \alpha_k, \beta_k$	Expectation of $r_{i,k}$ conditional on $s_{i,k}$, linear model for $u_{i,k} = \alpha_x + \beta_k s_{i,k}$
6	$\mathbf{G}, g_{kl}$	$2c_{max} \times 2c_{max}$ gene flow matrix, element of $\mathbf{G}$
6	$\mathbf{D}, d_{kl}$	$2c_{max} \times 2c_{max}$ matrix of deviations of breeding values from mean, element of $\mathbf{D}$
6	Π, π_{kl}	Matrix of regression coefficients of $s_{offspring}$ on s_{parent}, element of Π
6	Λ	Matrix of regression coefficients of the number of selected offspring on s_{parent}
6, 18	$\lambda_{kl}, n_{i,k}(l)$	Element of Λ, number of selected offspring in category l of parent i in category k
17	$\mathbf{V}_{n,k}, \Delta\mathbf{V}_{n,k}$	Matrix of variance of family size, deviation of $\mathbf{V}_{n,k}$ from Poisson variance
1	δ_k	Correction term required when $\Delta\mathbf{V}_{n(x)} \neq 0$

[a]Equation number or section where symbol is first defined, P = "Population Structure," G = "General," B = "Appendix B," S$_2$ = "Step 2," A = "Appendix A," PB = "Pseudo-BLUP Selection Index," R = "Results," and I = "Introduction."

Materials and Methods

In this section we will first describe the population structure for which rates of inbreeding will be predicted. Because deterministic prediction of rates of inbreeding (ΔF) requires a deterministic analogy to BLUP, a pseudo-BLUP selection index will be used to provide the necessary parameters (e.g., equilibrium genetic parameters; Bulmer, 1971). Next, the procedure for predicting rates of inbreeding will be outlined in three steps, which will be implemented using equilibrium genetic parameters. Finally, we will describe a stochastic simulation procedure that will be used to evaluate the accuracy of the deterministic prediction of ΔF. Table 1 shows the notation used.

The general theory to predict ΔF is derived fully in previous papers (e.g., Bijma and Woolliams, 2000; Woolliams and Bijma, 2000). This paper, therefore, focuses on the implementation of the theory for typical livestock improvement schemes.

Population Structure

Selection was for a trait determined by the additive infinitesimal model (Fisher, 1918; Bulmer, 1971). Phenotypic values (P) were the sum of additive genetic values (A, breeding values) and environmental values (E), $P = A + E$. A closed nucleus population with overlap-ping generations was modeled and selection was on BLUP-EBV for a single trait. With two sexes and a maximum age of c_{max}, there are $2c_{max}$ categories of animals, one for each sex and age of parent. Categories will be indexed by k or l, so $k,l = 1 .. c_{max}$ are males and $k,l = c_{max}+1 .. 2c_{max}$ are females.

Phenotypes of selection candidates were recorded prior to reproductive age and BLUP-EBV were calculated. Progeny testing was included for males in the oldest age class ($k = c_{max}$), by including information on n_{prg} progeny in their EBV. Those progeny were assumed to be born outside the nucleus, so their dams did not enter the breeding value estimation. Females did not have progeny information. Within categories, individuals were ranked on their EBV, and each year the highest-ranking n_k individuals were selected from the kth category to produce the next cohort. The number selected from each age class, n_k, was either set to a fixed value in advance or determined by truncation selection on EBV across age classes. Note that the maximum age, c_{max}, was assumed to be equal for both sexes, but different maximum ages can be included by using $n_k = 0$ for the sex-age classes that do not contribute parents. The total number of male and female parents of each newborn cohort was $N_m = \sum_{k=1}^{c_{max}} n_k$ and $N_f = \sum_{k=c_{max}+1}^{2c_{max}} n_k$, respectively. Each sire was mated at random to d dams

$(d = N_f / N_m)$, and each dam produced a fixed number, n_o, of offspring ($\frac{1}{2}n_o$ of each sex), so that the total number of selection candidates born in a cohort was $T = \frac{1}{2}n_oN_f$ for each sex. The unit of age (i.e., the interval between consecutive age classes) was 1 yr. Genetic contributions and rates of inbreeding per year were therefore equal to genetic contributions and rates of inbreeding per cohort.

Pseudo-BLUP Selection Index

A selection index analogy of the BLUP procedure was developed by extending the pseudo-BLUP selection index of Wray and Hill (1989) to populations with overlapping generations. Because part of the selection candidates may have progeny information, two pseudo-BLUP indices were distinguished. First, *index1*, without progeny information, was used for male selection candidates in categories 1 to $c_{max} - 1$ and for all female selection candidates. Second, *index2*, with progeny information, was used only for males in category c_{max}.

For the ith candidate, *index2* was $\hat{A}_{2,i} = \mathbf{b}_2^{\mathrm{T}}\mathbf{x}_{2,i}$, where superscript T denotes the transpose, $\hat{A}_{2,i}$ is the EBV, $\mathbf{b}_2$ is a 7×1 vector of index weights, and $\mathbf{x}_{2,i}$ is a 7×1 vector of information sources for the ith candidate. Information sources in $\mathbf{x}_{2,i}$ were 1) $\hat{A}_m$, the EBV of the sire of i; 2) $\hat{A}_f$, the EBV of the dam of i; 3) $\bar{\hat{A}}_f$, the average EBV of the d dams mated to the sire; 4) $\bar{P}_{HS}$, the phenotypic average of the n_od half-sibs of i (including i and its full-sibs); 5) $\bar{P}_{FS}$, the phenotypic average of the n_o full-sibs of i (including i); 6) P_i, the phenotype of candidate i; and 7) $\bar{P}_{prg}$, which is the phenotypic average of the n_{prg} offspring of a male in category c_{max}. *Index1* ($\hat{A}_{1,i}$) was identical to *index2*, except for information source 7, which was omitted.

Using the above pseudo-BLUP indices, selection and mating were iterated until equilibrium genetic parameters were reached (Bulmer, 1971). Iterative equations for calculating index weights, the accuracy of selection (ρ), and equilibrium parameters are given in Appendix A. Rates of inbreeding were predicted using equilibrium genetic parameters.

Prediction of Rates of Inbreeding

General. The prediction method is based on the concept of long-term genetic contributions (James and McBride, 1958; Wray and Thompson, 1990). The long-term genetic contribution (r_i) of ancestor i in generation t_1 is defined as the proportion of genes from i that are present in individuals in generation t_2 deriving by descent from i, where $(t_2 - t_1) \rightarrow \infty$ (Woolliams et al., 1993). In other words, the long-term genetic contribution of an individual is its proportional contribution to the genetic make-up of the population in the long term. In the remainder of this article, long-term genetic contributions are referred to as "genetic contributions," or simply as "contributions."

Wray and Thompson (1990) showed that rates of inbreeding per generation are proportional to the sum of squared contributions, $E(\Delta F) = \frac{1}{4}\sum r_i^2$, where r_i is the realized genetic contribution of individual i and the sum is taken over all parents in a generation. Recently, Woolliams and Bijma (2000) showed that rates of inbreeding can be expressed in terms of *expected* contributions:

$$E(\Delta F) = \frac{1}{2}\sum_s n_s E(u_{i,s}^2) + \frac{1}{8}\sum_s n_s\delta_s \qquad [1]$$

where n_s is the number of parents selected from lifetime category s, $u_{i,s}$ is the expected lifetime contribution of individual i in lifetime category s conditional on its selective advantage, and δ is a correction factor for deviations of the variance of family size ($\mathbf{V}_n$) from a Poisson variance. When $\mathbf{V}_n$ deviates from Poisson, $\frac{1}{4}\sum r_i^2$ differs from $\frac{1}{2}\sum_s n_s E(u_{i,s}^2)$ and the difference is accounted for by the second term of Eq. [1]. In this paper, we will predict ΔF using Eq. [1]. Throughout the paper, *family size* refers to the number of selected offspring of a parent, not to the number of candidates. The second term of Eq. [1] will be referred to as the "Poisson correction."

In Eq. [1], it is essential to note that $u_{i,s}$ refers to the *lifetime* contribution of individual i and subscript s denotes the lifetime category to which i belongs. The lifetime contribution of individual i is the sum of all contributions originating from its selection at a specific age, $u_{i,s} = \Sigma u_{ik}$, where the sum is taken over all age classes k in which individual i is selected. Throughout this paper, lifetime categories are indexed by s, which refers to a specific combination of age classes in which the individual is selected, whereas index k refers to one specific sex-age class. This issue will be addressed below (see "Step 2").

Components of Eq. [1] will be calculated in three steps. In the first step we will predict expected genetic contributions of sex-age classes, $u_{i,k}$, using the approach of Woolliams et al. (1999). In the second step, lifetime contributions, $u_{i,s}$, will be expressed as a function of sex-age class contributions, $u_{i,k}$, and, subsequently, $E(u_{i,s}^2)$ will be derived. In the third step we will derive δ_s.

In a selected population, a superior individual is expected to have a higher genetic contribution than an average individual, because its offspring and further descendents have a higher probability of being selected. When predicting genetic contributions, we need to explicitly model this superiority (i.e., we need to define the factors that confer selective advantage to an individual). Before proceeding to the prediction of expected contributions, therefore, we will first define the selective advantage.

Selective Advantage. In principle, the selective advantage should contain all terms that affect the long-term contribution of an individual (i.e., by affecting the selective success of its descendents). In this paper, we will use the breeding value of the individual plus the breeding value of its mate(s). The mate is included because

its breeding value affects the selective success of the offspring, so the breeding value of the mate can be regarded as a component of the selective advantage of an individual. Because random mating is used, the mate affects the expected contribution of an individual in a random manner, which contributes to the variance of the expected contributions (see Eq. [12] and [13]) and, therefore, to the rate of inbreeding. Other models for the selective advantage are possible (e.g., instead of using the true breeding value one may use the EBV together with the prediction error; Woolliams et al., 1999; see Bijma and Woolliams, 2000 for a discussion). For sire i in category k, the selective advantage was as follows:

$$s_{i,k} = (A_{i,k} + \overline{A}_f) - \overline{(A_{i,k} + \overline{A}_f)_k} \quad k = 1, c_{max} \quad [2]$$

where $A_{i,k}$ is the breeding value of sire i in category k, $\overline{A}_f$ is the average breeding value of the d dams mated to sire i, and the second term represents subtraction of the average selective advantage for category k. For dams the selective advantage was as follows:

$$s_{i,k} = (A_{i,k} + A_m) - \overline{(A_{i,k} + A_m)_k} \quad k = c_{max}+1, 2c_{max} \quad [3]$$

where $A_{i,k}$ is the breeding value of dam i in category k and A_m is the breeding value of the sire mated to dam i.

Step 1: Prediction of Expected Contributions. Following Woolliams et al. (1999), expected contributions were predicted by linear regression on the selective advantage. For males, the expected contribution of sire i in category k was as follows:

$$u_{i,k} = \mathrm{E}(r_{i,k}|s_{i,k}) = \alpha_k + \beta_k s_{i,k} + \sum_{j=1}^{d}(\alpha_{cat\,(j)} - \overline{\alpha}_f) \quad [4]$$

$$k = 1, c_{max}$$

where α_k represents the average contribution of an ancestor in category k, $\beta_k s_{i,k}$ represents the deviation of the contribution from this average due to the selective advantage of ancestor i, j denotes the mates of i, $cat(j)$ is the category of mate j, and $\overline{\alpha}_f = \sum_{k=c_{max}+1}^{2c_{max}} (n_k\alpha_k)/N_f$. The last term of Eq. [4] represents the effect of the categories of the mates on the contribution of individual i, which was not accounted for when defining the selective advantage of individual i (i.e., $\overline{A}_f$ in Eq. [2] accounts only for the within-category effect of the mate).

For females, the expected contribution of dam i in category k was as follows:

$$u_{i,k} = \mathrm{E}(r_{i,k}|s_{i,k}) = \alpha_k + \beta_k s_{i,k} + (\alpha_j - \overline{\alpha}_m)/d \quad [5]$$

$$k = c_{max} + 1, 2c_{max}$$

Note that, contrary to Bijma and Woolliams (2000), the effect of the mate is included directly in the model for predicting $u_{i,k}$. This difference does not affect the re-

sults, but the present approach is more straightforward. Solutions for $u_{i,k}$ are obtained by predicting α_k and β_k. There are two mechanisms determining α_k and β_k. First, superior parents are expected to have more selected offspring, which is modeled by a regression coefficient λ. Second, offspring partly inherit the selective advantage of their parents, which is modeled by a regression coefficient π. By modeling these two mechanisms, Woolliams et al. (1999) show that α_k and β_k can be obtained from

$$\mathbf{N}\boldsymbol{\alpha} = [\mathbf{G}^{\mathrm{T}} + (\mathbf{G}^{\mathrm{T}} * \mathbf{D}^{\mathrm{T}})(\mathbf{I} - \mathbf{G}^{\mathrm{T}} * \boldsymbol{\Pi}^{\mathrm{T}})^{-1}(\mathbf{G}^{\mathrm{T}} * \boldsymbol{\Lambda}^{\mathrm{T}})]\mathbf{N}\boldsymbol{\alpha} \quad [6]$$

$$\mathbf{N}\boldsymbol{\beta} = (\mathbf{I} - \mathbf{G}^{\mathrm{T}} * \boldsymbol{\Pi}^{\mathrm{T}})^{-1}(\mathbf{G}^{\mathrm{T}} * \boldsymbol{\Lambda}^{\mathrm{T}})\mathbf{N}\boldsymbol{\alpha} \quad [7]$$

where * denotes element × element multiplication; $\mathbf{I}$ is the $2c_{max} \times 2c_{max}$ identity matrix; $\mathbf{N}$ is a $2c_{max} \times 2c_{max}$ diagonal matrix containing the numbers of parents selected from each category (n_k), $\boldsymbol{\Pi}$ is a $2c_{max} \times 2c_{max}$ matrix of elements π_{kl}, being the regression coefficient of the selective advantage ($s_{j,k}$) of selected offspring j in category k on the selective advantage ($s_{i,l}$) of parent i in category l; $\boldsymbol{\Lambda}$ is a $2c_{max} \times 2c_{max}$ matrix of elements λ_{kl}, being the regression coefficient of the number of selected offspring in category k on the selective advantage of parent i in category l; $\mathbf{G}$ is a $2c_{max} \times 2c_{max}$ gene flow matrix of elements g_{kl}, specifying the proportional contribution of parent category l to selected offspring in category k (similar to a transition matrix in population genetics); $\mathbf{D}$ is a $2c_{max} \times 2c_{max}$ matrix of elements d_{kl}, being the average selective advantage of selected offspring in category k descending from parents in category l, expressed as a deviation from the mean selective advantage in category k; $\boldsymbol{\alpha}$ is a $2c_{max}$ vector of elements α_l; and $\boldsymbol{\beta}$ is a $2c_{max}$ vector of elements β_l. (See Bijma and Woolliams [1999] for a detailed study and an example of the prediction of expected genetic contributions with overlapping generations.) The above matrices follow the gene flow notation (Hill, 1974), so rows represent offspring categories and columns represent parent categories, and this is the reason that the matrices in Eq. [6] and [7] require the transpose.

Note that the gene flow matrix, $\mathbf{G}$, differs from Hill's (1974) gene flow matrix, which specifies the contribution of parent categories to offspring *before* selection. Here, $\mathbf{G}$ refers to selected offspring. For example, for $c_{max} = 2$, $g_{41} = 0.35$ means that 1-yr-old sires contribute 35% of the genes of 2-yr-old selected females.

Matrices $\mathbf{G}$, $\boldsymbol{\Pi}$, $\boldsymbol{\Lambda}$, and $\mathbf{D}$ were derived following the approach of Woolliams et al. (1999) and Bijma and Woolliams (1999). Resulting equations for the case studied in this paper are listed in Appendix B.

Note that contributions predicted from Eq. [4] through [7] are the contributions of a single cohort (i.e., the group born in a single year, not an entire generation) originating from their selection at the different ages. Rates of inbreeding predicted from these contributions are, therefore, also per year.

Step 2: Derivation of $E(u_{i,s}^2)$. The lifetime contribution is the sum of the contributions originating from selection at a specific age, $u_{i,s} = \sum u_{i,k}$, where $u_{i,k}$ is obtained from Eq. [4] or [5]. To calculate $E(u_{i,s}^2)$, therefore, we have to express contributions of lifetime categories, which are indexed by s, in terms of contributions of sex-age class categories, which are indexed by k. Lifetime categories refer to a specific combination of age classes in which the individual is selected (e.g., males selected only at 1 yr of age, which have $u_{i,s}^2 = u_{i,k=1}^2$; males selected at 1 and 2 yr of age, which have $u_{i,s}^2 = (u_{i,k=1} + u_{i,k=2})^2 = u_{i,k=1}^2 + 2u_{i,k=1}u_{i,k=2} + u_{i,k=2}^2$; and so on). This example shows that the square of the lifetime contribution, $u_{i,s}^2$, can be expressed as a sum of squares, $u_{i,k}^2$, and cross-products, $u_{i,k}u_{i,l}$, of sex-age class contributions.

Instead of explicitly specifying all lifetime categories and deriving the corresponding $E(u_{i,s}^2)$, one can directly express $\sum_s n_s E(u_{i,s}^2)$ as a sum of squares and cross-products in terms of the categories k, which has the advantage that contributions predicted from Eq. [4] and [5] can be used directly (Bijma and Woolliams, 2000). In the following equations for calculating squared expected contributions, terms contributing to $\sum_s n_s E(u_{i,s}^2)$ will be collected separately for males without progeny testing, for progeny-tested males, and for females.

For categories without progeny testing, the ranking of animals remains practically unchanged when they move through the age classes, which is the same situation as with mass selection. For those categories, therefore, Eq. [11] and [12] of Bijma and Woolliams (2000) can be used, so that, for male categories without progeny testing

$$\sum_s n_s E(u_{i,s}^2) = \sum_{k=1}^{c_{max}-1} n_k E(u_{i,k}^2)$$ [8]

$$+ 2 \sum_{k=1}^{c_{max}-2} \sum_{l=k+1}^{c_{max}-1} \min(n_l, n_k) E(u_{i,k}u_{i,l})$$

and for all female categories

$$\sum_s n_s E(u_{i,s}^2) = \sum_{k=c_{max}+1}^{2c_{max}} n_k E(u_{i,k}^2)$$ [9]

$$+ 2 \sum_{k=c_{max}+1}^{2c_{max}-1} \sum_{l=k+1}^{2c_{max}} \min(n_l, n_k) E(u_{i,k}u_{i,l})$$

where $\sum_s$ denotes summation over the relevant lifetime categories and $\min(n_k, n_l)$ denotes the minimum of n_k and n_l (see also the example in Bijma and Woolliams, 2000).

For Eq. [8] and [9], $E(u_{i,k}^2)$ is obtained by squaring Eq. [4] and [5], with $E(s_{i,k}^2) = \sigma_{s,k}^2$ because $E(s_{i,k}) = 0$. For Eq. [8] the result is

$$E(u_{i,k}^2) = \alpha_k^2 + \beta_k^2 \sigma_{s,k}^2 + d(\overline{\alpha_f^2} - \overline{\alpha}_f^2)$$ [10]
$$k = 1, c_{max}-1$$

and for Eq. [9],

$$E(u_{i,k}^2) = \alpha_k^2 + \beta_k^2 \sigma_{s,k}^2 + (\overline{\alpha_m^2} - \overline{\alpha}_m^2)/d^2$$ [11]
$$k = c_{max}+1, 2c_{max}$$

Next, for males, the variance of the selective advantage is (see Eq. [2])

$$\sigma_{s,k}^2 = \sigma_A^2(1 - \kappa_k \rho_k^2)(1 - 1/n_k)$$ [12]

$$+ \frac{\sigma_A^2}{d} \sum_{l=c_{max}+1}^{2c_{max}} \frac{n_l}{N_f}(1 - \kappa_l \rho_l^2)(1 - 1/n_l) \quad k = 1, c_{max}$$

and for females (see Eq. [3])

$$\sigma_{s,k}^2 = \sigma_A^2(1 - \kappa_k \rho_k^2)(1 - 1/n_k)$$ [13]

$$+ \sigma_A^2 \sum_{l=1}^{c_{max}} \frac{n_l}{N_m}(1 - \kappa_l \rho_l^2)(1 - 1/n_l) \quad k = c_{max} + 1, 2c_{max}$$

In Eq. [12] and [13], the first term is due to the individual itself, the second term is due to its mate(s), the term $(1 - \kappa\rho^2)$ accounts for reduced variance because the parents are a selected group, κ is Pearson's variance reduction coefficient, and the $1 - 1/n_k$ accounts for finite sample size.

For Eq. [8] and [9], expectations of cross-products are calculated from Bijma and Woolliams (2000):

$$E(u_{i,k}u_{i,l}) = \alpha_k \alpha_l + [1 - 1/n_{min}]\beta_k \beta_l \sigma_A^2[1 - \kappa_{min}\rho_{min}^2]$$ [14]
$$+ \alpha_{min}\beta_{max}E[\overline{A}_{min} - \overline{A}_{max}]$$

where subscript *min* (*max*) denotes the category with the lower (higher) number of animals and $\overline{A}_k$ is the genetic selection differential in category k, $\overline{A}_k = \iota_k \rho_k \sigma_A$. With random mating, there is no covariance between the selective advantages of two different mates, so that mates do not contribute to the cross-product in Eq. [14].

For progeny-tested males we need to derive the contribution of category c_{max} to the sum of squared lifetime contributions, which is composed of the sum of squared contributions from category c_{max} and the sum of cross-products between category c_{max} and the non-progeny-tested male categories,

$$\sum_s n_s E(u_{i,s}^2) = n_{c_{max}} E(u_{i,c_{max}}^2)$$ [15]

$$+ 2 \sum_{k=1}^{c_{max}-1} n_{k,c_{max}} E(u_{i,k}u_{i,c_{max}})$$

where $E(u_{i,c_{max}}^2)$ follows from Eq. [10] by putting $k = c_{max}$, and $n_{k,c_{max}}$ is the number of parents selected jointly

in category k and category c_{max}. The number selected jointly in both categories can be calculated from the number of candidates and the proportion selected jointly in category k and category c_{max}; $n_{k,c_{max}} = \Phi(\tau_k, \tau_{c_{max}}, \rho_{k,c_{max}})T$, where $\Phi(\tau_k, \tau_{c_{max}}, \rho_{k,c_{max}})$ denotes the bivariate normal proportion above truncation points τ_k and $\tau_{c_{max}}$, which are the EBV truncation points for category k and c_{max}; and $\rho_{k,c_{max}}$ is the correlation between $\hat{A}_{i,k}$ and $\hat{A}_{i,c_{max}}$, $\rho_{k,c_{max}} = \sigma_{\hat{A},k}/\sigma_{\hat{A},c_{max}}$. The bivariate normal proportion, $\Phi(\tau_k, \tau_{c_{max}}, \rho_{k,c_{max}})$, was calculated using Dutt's algorithm (Dutt, 1973; Dutt and Soms, 1976; Ducrocq and Colleau, 1986). For Eq. [15], expectations of cross-products follow from Eq. [2] and [5], with no contribution due to the mates:

$$E(u_{i,k}u_{i,c_{max}}) = \alpha_k\alpha_{c_{max}} + \alpha_k\beta_{c_{max}}[E(A_i) - \overline{A}_{c_{max}}] \quad [16]$$
$$+ \alpha_{c_{max}}\beta_k[E(A_i) - \overline{A}_k] + \beta_k\beta_{c_{max}}E[(A_i - \overline{A}_k)(A_i - \overline{A}_{c_{max}})]$$

where i refers to individuals that are selected both in category k and in category c_{max} and $E(A_i)$ is the expected breeding value of those individuals. The terms $E(A_i)$ and $E[(A_i - \overline{A}_k)(A_i - \overline{A}_{c_{max}})]$ are calculated using a result of Tallis (1961) and are given in Appendix C.

Summarizing, the sum of squared expected lifetime contributions is given by Eq. [8] for males without progeny testing, by Eq. [9] for females, and by Eq. [15] for progeny-tested males. Finally, the first term of Eq. [1] is obtained by summing results from Eq. [8], [9], and [15]. The remaining task is to obtain the second term of Eq. [1], which requires the calculation of δ_k.

Step 3: Calculation of δ_k. The calculation of δ_k is a straightforward analogy of the discrete generation case (Bijma and Woolliams, 2000). Here we will outline the concept; explicit equations are given in Appendix D. In Eq. [1],

$$\delta_k = \boldsymbol{\alpha}^{\mathrm{T}}\Delta\mathbf{V}_{n,k}\boldsymbol{\alpha} \quad [17]$$

where $\Delta\mathbf{V}_{n,k}$ is the $2c_{max} \times 2c_{max}$ matrix of deviations of the variance of family size from a Poisson variance, for a parent in category k, conditional on its selective advantage $s_{i,k}$ (Woolliams and Bijma, 2000). For example, for the full variance of family size conditional on the selective advantage, element $\mathbf{V}_{n,k}(l,l')$ represents the covariance between the number of offspring selected in category l, $n_{i,k}(l)|s_{i,k}$, and the number of offspring selected in category l', $n_{i,k}(l')|s_{i,k}$, of a parent in category k. For diagonal elements, the deviation from a Poisson variance is obtained by subtracting the mean number of selected offspring from the full variance (with Poisson, $\sigma^2 = \mu$). For diagonal elements, therefore, $\Delta\mathbf{V}_{n,k}(l,l) = \mathrm{E}_s\{\mathrm{E}[n_{i,k}^2(l)|s_{i,k}] - \mathrm{E}[n_{i,k}(l)|s_{i,k}]^2 - \mathrm{E}[n_{i,k}(l)|s_{i,k}]\}$, where E_s denotes the expectation with respect to $s_{i,k}$, which gives (see also Bijma and Woolliams, 2000):

$$\Delta\mathbf{V}_{n,k}(l,l) = E_s\{E[n_{i,k}(l)(n_{i,k}(l) - 1)|s_{i,k}]\} \quad [18]$$
$$- E_s\{E[n_{i,k}(l)|s_{i,k}]^2\}$$

For off-diagonal elements, $\Delta\mathbf{V}_{n,k}(l,l') = \mathbf{V}_{n,k}(l,l')$, because, with an independent Poisson distribution for each category, the covariance between $n_{i,k}(l)|s_{i,k}$ and $n_{i,k}(l')|s_{i,k}$ is zero. For off-diagonal elements, therefore,

$$\Delta\mathbf{V}_{n,k}(l,l') = E_s\{E[n_{i,k}(l)n_{i,k}(l')|s_{i,k}]\} \quad [19]$$
$$- E_s\{E[n_{i,k}(l)|s_{i,k}]E[n_{i,k}(l')|s_{i,k}]\}$$

where $s_{i,k}$ is the selective advantage of parent i in category k and $n_{i,k}(l)$ is the number of offspring of parent i selected in category l. The extension of Eq. [18] and [19] to a population with overlapping generations and a hierarchical mating structure is given in Appendix D. In summary, the second term of Eq. [1] is obtained using Eq. [17], where $\Delta\mathbf{V}_{n,k}$ is given by Eq. [D1] through [D14].

Results

Accuracy of Predictions

Table 2 shows rates of inbreeding per year from simulation (ΔF_{sim}) and corresponding prediction errors for populations with two age classes, where EBV for males in age class two include information on 100 progeny. In Table 2, the number of parents selected from each age class is a result of truncation selection on EBV across age classes, showing that selection of males moves toward age class two when h^2 and n_o are low. Predictions are accurate for $n_o = 2$. For $n_o = 4$ or 8, predictions are approximately 8% too low. In spite of the errors, predictions and simulations show the same trend. For example, doubling the number of offspring per parent from $n_o = 4$ to $n_o = 8$ for a scheme with $h^2 = 0.1$ raises ΔF_{sim} by a factor 1.7, whereas the prediction indicates a factor 1.8. Note that the use of $\Delta F = 1/(8N_m) + 1/(8N_f)$ would give the same rate of inbreeding for schemes with $n_o = 4$ vs 8 (i.e., a factor of 1).

Table 3 shows rates of inbreeding from simulation and corresponding prediction errors for populations with three age classes, truncation selection across age classes, and where EBV for males in age class three include information on 100 progeny. Accuracy of predictions in Table 3 is comparable to that in Table 2.

With truncation selection across age classes, the majority of the parents are selected from the youngest age class in most cases. To evaluate the accuracy of predictions for any distribution of parents across age classes, the proportion of parents selected from each age class was set to a fixed value and animals were selected by truncation on EBV within age classes. For a population with two age classes, three alternatives for the proportion of parents selected from the different age classes were considered, $p = (0.75, 0.25; 0.75, 0.25)$, $p = (0.5, 0.5; 0.75, 0.25)$, and $p = (0.25, 0.75; 0.75, 0.25)$. For example, $p = (0.25, 0.75; 0.75, 0.25)$ with $N_m = 40$

Table 2. Rates of inbreeding per year from simulation (ΔF_{sim}) and corresponding prediction errors for schemes with two age classes, 20 sires and 60 dams selected per year and truncation selection on EBV across age classes[a]

n_o^c	$h^{2\ d}$	diag$\mathbf{N}^e$				$\Delta F_{sim}^{\ f}$	Error %[b]
2	0.1	10	10	43	17	0.0089	−3
	0.2	11	9	42	18	0.0095	0
	0.4	12	8	41	19	0.0100	2
	0.6	14	6	40	20	0.0101	2
4	0.1	9	11	54	6	0.0131	−10
	0.2	11	9	53	7	0.0147	−7
	0.4	14	6	52	8	0.0160	−8
	0.6	16	4	51	9	0.0151	−9
8	0.1	11	9	58	2	0.0221	−3
	0.2	13	7	58	2	0.0243	−4
	0.4	16	4	58	2	0.0245	−14
	0.6	18	2	57	3	0.0195	−10

[a]Sires in age class two have information on 100 progeny included in their EBV.
[b]Error % = $100\% \times (\Delta F_{pred} - \Delta F_{sim})/\Delta F_{sim}$.
[c]n_o = number of offspring per dam.
[d]h^2 = heritability.
[e]diag$\mathbf{N}$ = number of parents selected from each age class; first two elements refer to sires, last two elements refer to dams.
[f]Standard errors of simulation results were smaller than 1% of their mean value.

and N_f = 80 gives $\mathbf{N}$ = diag{10, 30, 60, 20}. For each of these distributions, all combinations of schemes were evaluated for N_m = 8, 12, 20, 40, 60, or 80; d = 2, 4, or 8; n_o = 2, 4, or 8; and h^2 = 0.1, 0.2, 0.4, or 0.6, with information on 100 progeny for males in age class two, and the restriction that $N_f \leq 160$ to limit computing time for the stochastic simulations. In total 468 different schemes were evaluated. Within this range, the maximum rate of inbreeding was ΔF_{sim} = 0.0597/yr for $\mathbf{N}$ = diag{6, 2, 12, 4}, n_o = 8 and h^2 = 0.1, with a prediction error of +1%. The minimum rate of inbreeding was,

ΔF_{sim} = 0.0021/yr for $\mathbf{N}$ = diag{20, 60, 120, 40}, n_o = 2, and h^2 = 0.1, with a prediction error of −5%.

Table 4 shows the average error and the standard deviation of the error for the whole range of schemes mentioned above, where schemes are grouped according to the number of sires. Schemes with up to 20 sires show accurate predictions (i.e., the absolute value of the mean error and the standard deviation of the error are below 5%). Schemes with more than 20 sires show a systematic underprediction of approximately 8% and an increasing standard deviation of the error.

Table 3. Rates of inbreeding per year from simulation (ΔF_{sim}) and corresponding prediction errors for schemes with three age classes, 20 sires and 60 dams selected per year and truncation selection on EBV across age classes[a]

n_o^c	$h^{2\ d}$	diag$\mathbf{N}^e$						$\Delta F_{sim}^{\ f}$	Error %[b]
2	0.1	12	2	6	40	17	3	0.0093	−6
	0.2	13	2	5	39	17	4	0.0105	−5
	0.4	14	3	3	38	17	5	0.0117	−3
	0.6	15	3	2	37	18	5	0.0112	−2
4	0.1	14	1	5	52	8	0	0.0156	−12
	0.2	16	1	3	51	9	0	0.0177	−11
	0.4	18	1	1	50	9	1	0.0173	−13
	0.6	18	2	0	50	9	1	0.0150	−9
8	0.1	18	0	2	57	3	0	0.0313	−7
	0.2	19	0	1	57	3	0	0.0305	−10
	0.4	20	0	0	57	3	0	0.0242	−9
	0.6	20	0	0	57	3	0	0.0186	−9

[a]Sires in age class three have information on 100 progeny included in their EBV.
[b]Error % = $100\% \times (\Delta F_{pred} - \Delta F_{sim})/\Delta F_{sim}$.
[c]n_o = number of offspring per dam.
[d]h^2 = heritability.
[e]diag$\mathbf{N}$ = number of parents selected from each age class.
[f]Standard errors of simulation results were smaller than 1% of their mean value.

Table 4. Mean and standard deviation of the prediction error
for a range of schemes with two age classes[a]

| N_m[c] | p = 0.75 0.25 0.75 0.25[b] | | p = 0.5 0.5 0.75 0.25 | | p = 0.25 0.75 0.75 0.25 | |
	Mean	SD	Mean	SD	Mean	SD
8	−2	3	2	3	0	4
12	−4	4	0	3	−1	3
20	−5	5	−1	5	−4	3
40	−5	6	−7	6	−7	5
60	−7	9	−7	8	−9	7
80	−7	10	−9	9	−9	9

[a]Sires in age class two have information on 100 progeny inluded in their EBV.
[b]p denotes the proportions of parents selected from each age class.
[c]N_m = number of sires. For each N_m alternative, results were averaged over the heritability, mating ratio, and number of offspring alternatives (see text).

To give some background information on prediction errors in Table 4, Table 5 shows the typical error trend for schemes with many parents. When selection intensity is low (n_o = 2) predictions are accurate. For higher selection intensities (n_o = 4 or 8), predictions are too low. In spite of the errors, predictions give a good indication of the effect of changing population parameters. For example, doubling the number of offspring per parent increases the rate of inbreeding from simulation by a factor 1.96 (h^2 = 0.1, n_o = 4 vs. 8). For the same scheme, the prediction indicates an increase by a factor 1.82, which gives a good indication of the tremendous effect of selection intensity on ΔF with selection on BLUP-EBV. Bijma and Woolliams (2000) obtained similar error trends for populations with discrete generations and showed that the underprediction is due to the use of simple linear models to predict expected genetic contributions (Eq. [2] through [5]).

Table 6 shows rates of inbreeding and corresponding prediction errors for schemes with four age classes for each sex. Because the potential number of alternative schemes is very large, results are presented for a limited number of schemes with N_m = 20, d = 3, and h^2 = 0.3, where the distribution of parents across age classes was varied. For most schemes in Table 6, predictions are accurate or show some underprediction due to the same reason as mentioned above. There are, however, three

schemes in Table 6 that show a surprising overprediction of 14, 11, and 16%.

Detailed examination of the schemes with 14 and 11% error revealed that the overprediction of ΔF was due to overprediction of the contributions of 1-yr-old sires (α_1), which in turn was due to overprediction of the selection intensity in category one. In category one, only two sires are selected and the intraclass correlation between sibs is relatively high ($\overline{\rho}_{FS,11} \approx 0.73$; $\overline{\rho}_{HS,11} \approx 0.33$), indicating that reduction of selection intensity due to finite numbers and correlations between indices of relatives becomes important. Adjusting selection intensities using the method of Meuwissen (1991) reduced intensities from 2.73 to approximately 2.57 for both schemes, and prediction errors reduced from +14% to −4% for the one scheme and from +11% to −3% for the other scheme. These same schemes required adjustment to the selection intensity to obtain accurate prediction of genetic gain. This indicates that the need for adjusted selection intensities is not a specific feature of the method to predict ΔF but is a general requirement for schemes with few parents and high intraclass correlations between EBV of sibs (Meuwissen, 1991).

Examination of the scheme with 16% error revealed that the overprediction of ΔF was due to overprediction of the variance of family size. When the number of selected parents is small compared to the number of

Table 5. Typical trend of the prediction errors for a scheme with many parents[ab]

| $h^{2 e}$ | n_o = 2[d] | | n_o = 4 | | n_o = 8 | |
	ΔF_{sim}	Error %[c]	ΔF_{sim}	Error %[c]	ΔF_{sim}	Error %[c]
0.1	0.0025	0	0.0046	−13	0.0090	−19
0.2	0.0026	0	0.0047	−13	0.0091	−20
0.3	0.0026	4	0.0046	−11	0.0081	−17
0.4	0.0026	4	0.0042	−7	0.0069	−13

[a]For $\mathbf{N}$ = diag{40, 40, 120, 40}.
[b]Sires in age class two have information on 100 progeny included in their EBV.
[c]Error % = 100% × ($\Delta F_{pred} - \Delta F_{sim}$)/$\Delta F_{sim}$.
[d]n_o = number of offspring per dam.
[e]h^2 = heritability.
[f]ΔF_{sim} = rate of inbreeding per year from simulation.

Table 6. Rates of inbreeding per year from simulation (ΔF_{sim}) and corresponding prediction errors for a scheme with four age classes[ab]

diag**N**[c]	n_o[d]	ΔF_{sim}^{f}	Error %[e]
10 5 3 2 30 15 10 5	2	0.0132	0
	4	0.0227	−4
	8	0.0379	−4
10 5 3 2 5 10 15 30	2	0.0117	−1
	4	0.0214	−7
	8	0.0377	−7
2 3 5 10 30 15 10 5	2	0.0098	1
	4	0.0171	3
	8	0.0302	14
2 3 5 10 5 10 15 30	2	0.0085	4
	4	0.0156	−4
	8	0.0262	11
5 5 5 5 15 15 15 15	2	0.0125	9
	4	0.0224	16
	8	0.0390	3
2 8 8 2 5 25 25 5	2	0.0102	−6
	4	0.0171	−13
	8	0.0272	−7
8 2 2 8 25 5 5 25	2	0.0111	3
	4	0.0208	2
	8	0.0390	−1
2 8 2 8 5 25 5 25	2	0.0086	−1
	4	0.0156	−9
	8	0.0257	1

[a]Sires in age class four have information on 100 progeny included in their EBV.
[b]For $h^2 = 0.3$.
[c]diag**N** = number of parents selected from each age class.
[d]n_o = number of offspring per dam.
[e]Error % = $100\% \times (\Delta F_{pred} - \Delta F_{sim})/\Delta F_{sim}$.
[f]Standard errors of simulation results were smaller than 1% of their mean value.

candidates per family, all parents may be selected from very few families, which can be accounted for by adjusting the selected proportion according to Eq. [D13] (Wray et al., 1990). In the present paper, adjusted selected proportions were used for all schemes where minimum $(n_k; k = 1, c_{max}) \leq 0.75 n_o$ (see Appendix D). For the scheme with **N** = diag{5, 5, 5, 5, 15, 15, 15, 15}, this means that selected proportions were not adjusted for $n_o = 2$ or 4, whereas for $n_o = 8$ selected proportions were adjusted. The scheme with $n_o = 4$ is borderline (i.e., in each male age class the number of parents is small, but just above the threshold of $0.75 n_o$ that was used for adjusting the selected proportion). Adjusting the selected proportion for the scheme with $n_o = 4$ reduced the prediction error from +16% to +4%.

Discussion

This paper shows how the general procedure of Woolliams et al. (1999) and Woolliams and Bijma (2000) for predicting rates of inbreeding in selected populations can be implemented for livestock improvement schemes. Except for methods that ignore selection, there are no other methods available to predict rates of inbreeding for livestock improvement schemes at present. Detailed discussions on theoretical issues of the methods have been included in previous papers (Woolliams et al. [1999], Bijma and Woolliams [2000], Woolliams and Bijma [2000]). This discussion, therefore, addresses topics related to the implementation.

In spite of the prediction errors, the present method is a substantial improvement over other available methods, which ignore the effect of selection on the rate of inbreeding. In the absence of selection, ΔF of a population with overlapping generations is equal to ΔF of a population with discrete generations having the same number of parents entering the population per generation and the same lifetime variance of family size (Hill, 1972, 1979). Following that approach, ΔF was predicted for the schemes in Table 5 with $n_o = 8$, resulting in $\Delta F_{pred} = 0.0022$. (Note that, when ignoring selection, ΔF_{pred} is independent of heritability.) In the worst case, the present method showed an underprediction of 20% (0.0073 vs 0.0091), whereas the prediction ignoring selection gives an error of −76% (0.0022 vs 0.0091). Thus, for the scheme with which the present method performs worst, it still accounts for 80% of the true inbreeding, whereas the method ignoring selection only accounts for 24% of the true inbreeding.

The magnitude and pattern of the prediction errors in the present study are in line with prediction errors encountered by Bijma and Woolliams (2000) for populations with discrete generations. Simulation results in discrete generations indicate that the use of a quadratic model for predicting genetic contributions will remove the underprediction for schemes with many parents (Bijma and Woolliams, 2000). We expect that this conclusion extends to overlapping generations, but implementation of such a model is difficult.

With BLUP selection, prediction of the lifetime contribution (required for "*step 2*") is more complicated than with mass selection. With mass selection, in which the phenotype is recorded only once, the ranking of selection candidates remains unchanged when animals become older, because no additional information is added at older ages. When the ranking of animals remains unchanged over time, the number of animals selected in any combination of age classes can be derived directly from the number of animals selected in each age class. With mass selection, therefore, the lifetime contribution depends directly on the number of animals selected from each age class (Bijma et al., 2000a). With BLUP selection, however, the EBV of selection candidates may change when animals become older, because new information (e.g., progeny) becomes available. This will change the ranking of selection candidates. With different rankings in different age classes, the number of animals selected in a particular combination of age classes depends on the proportion selected jointly in those age classes. In the present paper, such bivariate normal proportions were obtained by numerical integration using Dutt's algorithm (Dutt and Soms, 1976; Ducrocq and Colleau, 1986).

When there are more than two different indices (e.g., when progeny information accumulates gradually so that each age class has a different amount of information available) the lifetime contribution can still be predicted using the bivariate normal distribution because cross-products between any two age classes, $u_{i,k}u_{i,l}$, involve only two categories at a time. With different indices for each age class, therefore, the present method can still be applied but cross-products need to be calculated from Eq. [16] for all age classes.

In the present study, we have assumed random mating of selected sires and dams. In practice, matings between close relatives will be avoided. In populations of practical sizes, avoidance of mating between close relatives (e.g., full-sibs) has only a small effect on the rate of inbreeding, because the probability of accidentally mating two full-sibs is small, so avoidance of those matings induces only small deviations from Hardy-Weinberg equilibrium. When a mating strategy is applied that additionally accounts for more distant relatives (e.g., minimum coancestry mating), the rate of inbreeding will be smaller than the values predicted for random mating (Caballero et al., 1996).

In the present study, selection is for a single trait. Predictions for multitrait selection can be developed using the same methodology. First, a multitrait pseudo-BLUP selection index for populations with overlapping generations has to be developed. This can be done by extending the work of Villanueva et al. (1993) to overlapping generations. With multitrait selection, the selective advantage would consist of the sum of breeding values for the different traits, weighted by their economic value. Subsequently, steps 1 to 3 remain essentially the same, but equations have to be derived based on a multitrait index. For example, in Eq. [6] and [7], the Π-matrix would represent the regression of the aggregate breeding value of a selected offspring on the aggregate breeding value of the parent. All of these steps can be performed using standard selection index theory.

One needs to take care when applying the present methodology to populations with extremely high correlations between estimated breeding values of sibs. For example, in a dairy MOET scheme, selection of young bulls may solely be based on information from relatives, so that between-full-sib family selection is practiced ($\rho_{FS} = 1$). For such cases, the equations for calculating the variance of family size and the linear model for predicting expected genetic contributions may be less accurate.

Until now, optimizing rates of gain and inbreeding in livestock improvement schemes required computationally demanding stochastic simulation, which restricts the number of alternative schemes considered. With the present method, rates of inbreeding in livestock improvement schemes can be predicted within very limited computing time, which shows that the general theory of Woolliams et al. (1999) and Woolliams and Bijma (2000) provides a toolbox for the optimization of breeding schemes considering both rates of genetic gain and inbreeding. Recently, Bijma et al. (2000b) optimized rates of gain and inbreeding for crossbreeding schemes, which illustrates how the method to predict ΔF can be connected to an optimization procedure. In their study, approximately 750 alternative breeding schemes were evaluated within 9 CPU seconds. This shows that, even when connected to a numerical search algorithm, the method can be used interactively. A computer program is available from the corresponding author.

Implications

The common livestock breeding practice of selection on estimated breeding values using Best Linear Unbiased Prediction has enabled increased rates of genetic gain but will also lead to increased rates of inbreeding, and thus endangers selection response and genetic diversity in the long term. In this article, we developed a deterministic method to predict rates of inbreeding for livestock improvement schemes. The method enables one to consider both rates of genetic gain and inbreeding before a breeding scheme commences and is therefore an important aid to design sustainable animal breeding plans. This implies that we no longer have to resort to computationally demanding stochastic simulation to balance short- and long-term response in livestock breeding populations.

Literature Cited

Bijma, P., J. A. M. Van Arendonk, and J. A. Woolliams. 2000a. A general procedure to predict rates of inbreeding in populations undergoing mass selection. Genetics 154:1865–1877.

Bijma, P., and J. A. Woolliams. 1999. Prediction of genetic contributions and generation intervals in populations with overlapping generations under selection. Genetics 151:1197–1210.

Bijma, P., and J. A. Woolliams. 2000. Prediction of rates of inbreeding in populations selected on Best Linear Unbiased Prediction of breeding value. Genetics 156:361–373.

Bijma, P., J. A. Woolliams, and J. A. M. Van Arendonk. 2000b. Genetic gain of pure line selection and combined crossbred purebred selection with constrained inbreeding. Genet. Sel. Evol. (In press).

Bulmer, M. G. 1971. The effect of selection on genetic variability. Am. Nat. 105:201–211.

Caballero, A., E. Santiago, and M. A. Toro. 1996. Systems of mating to reduce inbreeding in selected populations. Anim. Sci. 62:431–442.

Ducrocq, V., and J. J. Colleau. 1986. Interest in quantitative genetics of Dutt's and Deak's methods for numerical computation of multivariate normal probability integrals. Genet. Sel. Evol. 18:447–474.

Dutt, J. E. 1973. A representation of multivariate probability integrals by integral transforms. Biometrica 60:637–645.

Dutt, J. E., and A. P. Soms. 1976. An integral representation technique for calculating general multivariate probabilities with an application to multivariate χ^2. Commun. Stat. 11:2195–2205.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Roy. Soc. (Edinburgh) Trans. 52:321–341.

Hill, W. G. 1972. Effective size of populations with overlapping generations. Theor. Pop. Biol. 3:278–289.

Hill, W. G. 1974. Prediction and evaluation of response to selection with overlapping generations. Anim. Prod. 18:117–139.

Hill, W. G. 1979. A note on effective population size with overlapping generations. Genetics 92:317–322.

James, J. W., and G. McBride. 1958. The spread of genes by natural and artificial selection in a closed poultry flock. J. Genet. 56:55–62.

Meuwissen, T. H. E. 1991. Reduction of selection differentials in finite populations with a nested full-half-sib family structure. Biometrics 47:195–203.

Mendel, N. R., and R. C. Elston. 1974. Multifactorial qualitative traits: Genetic analyses and prediction of recurrence risks. Biometrics 30:41–57.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical Recipes in FORTRAN: The Art of Scientific Computing. Cambridge University Press, Cambridge, U.K.

Tallis, G. M. 1961. The moment generating function of the truncated multi-normal distribution. J. R. Stat. Soc. B. 23:223–229.

Verrier, E., J. J. Colleau, and J. L. Foulley. 1993. Long-term effects of selection based on the animal model BLUP in a finite population. Theor. Appl. Genet. 87:446–454.

Villanueva, B., N. R. Wray, and R. Thompson. 1993. Prediction of asymptotic rates of response from selection on multiple traits using univariate and multivariate best linear unbiased predictors. Anim. Prod. 57:1–13.

Woolliams, J. A., and P. Bijma. 2000. Predicting rates of inbreeding in populations undergoing selection. Genetics 154:1851–1864.

Woolliams, J. A., P. Bijma, and B. Villanueva. 1999. Expected genetic contributions and their impact on gene flow and genetic gain. Genetics 153:1009–1020.

Woolliams, J. A., N. R. Wray, and R. Thompson. 1993. Prediction of long-term contributions and inbreeding in populations undergoing mass selection. Genet. Res. Camb. 62:231–242.

Wray, N. R., and W. G. Hill. 1989. Asymptotic rates of response from index selection. Anim. Prod. 49:217–227.

Wray, N. R., and R. Thompson. 1990. Prediction of rates of inbreeding in selected populations. Genet. Res. Camb. 55:41–54.

Wray, N. R., J. A. Woolliams, and R. Thompson. 1990. Methods for predicting rates of inbreeding in selected populations. Theor. Appl. Genet. 80:503–512.

Wray, N. R., J. A. Woolliams, and R. Thompson. 1994. Prediction of rates of inbreeding in populations undergoing index selection. Theor. Appl. Genet. 87:878–892.

Wright, S. 1969. Evolution and the genetics of populations. vol. 2. The Theory of Gene Frequencies. University of Chicago, Chicago, IL.

Appendix A

Pseudo-BLUP Selection Index

There are two indices, *Index1* ($\hat{A}_{1,i}$) without progeny information and *Index2* ($\hat{A}_{2,i}$) with progeny information. Weights for *Index1* (*Index2*) are $\mathbf{b}_1 = \mathbf{V}_1^{-1}\mathbf{g}_1 (\mathbf{b}_2 = \mathbf{V}_2^{-1}\mathbf{g}_2)$, where $\mathbf{V}_1$ ($\mathbf{V}_2$) is the 6×6 (7×7) covariance matrix of information sources in $\mathbf{x}_{1,i}$ ($\mathbf{x}_{2,i}$) and $\mathbf{g}_1$ ($\mathbf{g}_2$) is the 6×1 (7×1) vector of covariances between information sources in $\mathbf{x}_{1,i}$ ($\mathbf{x}_{2,i}$) and the breeding value of the candidate. *Index1* is identical to the index of Wray and Hill (1989).

Elements of $\mathbf{V}_1$ are, per row, as follows: $\mathbf{V}_1(1,1\text{–}6) = (\sigma_m^2, 0, 0, \tfrac{1}{2}\sigma_m^2, \tfrac{1}{2}\sigma_m^2, \tfrac{1}{2}\sigma_m^2)$; $\mathbf{V}_1(2,1\text{–}6) = (0, \sigma_f^2, \sigma_f^2/d, \tfrac{1}{2}\sigma_f^2/d, \tfrac{1}{2}\sigma_f^2, \tfrac{1}{2}\sigma_f^2)$; $\mathbf{V}_1(3,1\text{–}6) = (0, \sigma_f^2/d, \sigma_f^2/d, \tfrac{1}{2}\sigma_f^2/d, \tfrac{1}{2}\sigma_f^2/d, \tfrac{1}{2}\sigma_f^2/d)$; $\mathbf{V}_1(4,1\text{–}6) = (\tfrac{1}{2}\sigma_m^2, \tfrac{1}{2}\sigma_f^2/d, \tfrac{1}{2}\sigma_f^2/d, \sigma_{HS}^2, \sigma_{HS}^2, \sigma_{HS}^2)$; $\mathbf{V}_1(5,1\text{–}6) = (\tfrac{1}{2}\sigma_m^2, \tfrac{1}{2}\sigma_f^2, \tfrac{1}{2}\sigma_f^2/d, \sigma_{HS}^2, \sigma_{FS}^2, \sigma_{FS}^2)$; and $\mathbf{V}_1(6,1\text{–}6) = (\tfrac{1}{2}\sigma_m^2, \tfrac{1}{2}\sigma_f^2, \tfrac{1}{2}\sigma_f^2/d, \sigma_{HS}^2, \sigma_{FS}^2, \sigma_{self}^2)$, where σ_m^2 and σ_f^2 are the variance of the EBV among selected

sires and dams, respectively; σ_{HS}^2 and σ_{FS}^2 are the between-full-sib family and between-half-sib family variance, which are $\sigma_{HS}^2 = \tfrac{1}{4}\sigma_{A,m}^2 + \tfrac{1}{4}\sigma_{A,f}^2/d + (\tfrac{1}{2}\sigma_{A,0}^2 + \sigma_E^2)/n_o d$ and $\sigma_{FS}^2 = \tfrac{1}{4}\sigma_{A,m}^2 + \tfrac{1}{4}\sigma_{A,f}^2 + (\tfrac{1}{2}\sigma_{A,0}^2 + \sigma_E^2)/n_o$, where $\sigma_{A,m}^2$ ($\sigma_{A,f}^2$) is the genetic variance among the selected sires (dams), σ_E^2 is the environmental variance, $\sigma_{A,0}^2$ is the base generation additive genetic variance and $\sigma_{self}^2 = \sigma_A^2 + \sigma_E^2$, where σ_A^2 is the total additive genetic variance. Furthermore, $\mathbf{g}_1^T = (\tfrac{1}{2}\sigma_m^2, \tfrac{1}{2}\sigma_f^2, \tfrac{1}{2}\sigma_f^2/d, \tfrac{1}{4}\sigma_{A,m}^2 + \tfrac{1}{4}\sigma_{A,f}^2/d + \tfrac{1}{2}\sigma_{A,0}^2/n_o d, \tfrac{1}{4}\sigma_{A,m}^2 + \tfrac{1}{4}\sigma_{A,f}^2 + \tfrac{1}{2}\sigma_{A,0}^2/n_o, \sigma_A^2)$.

Matrix $\mathbf{V}_2$ is identical to $\mathbf{V}_1$ but has additional $\mathbf{V}_2(7,1\text{–}7) = \mathbf{V}_2(1\text{–}7,7) = [\tfrac{1}{4}\sigma_m^2, \tfrac{1}{4}\sigma_f^2, \tfrac{1}{4}\sigma_f^2/d, \text{Cov}(\overline{P}_{HS}, \overline{P}_{prg}), \text{Cov}(\overline{P}_{FS}, \overline{P}_{prg}), \tfrac{1}{2}\sigma_A^2, \text{Var}(\overline{P}_{prg})]$, where $\text{Cov}(\overline{P}_{HS}, \overline{P}_{prg}) = \tfrac{1}{8}\sigma_{A,m}^2 + \tfrac{1}{8}\sigma_{A,f}^2/d + \tfrac{1}{4}\sigma_{A,0}^2/n_o d$, $\text{Cov}(\overline{P}_{FS}, \overline{P}_{prg}) = \tfrac{1}{8}\sigma_{A,m}^2 + \tfrac{1}{8}\sigma_{A,f}^2 + \tfrac{1}{4}\sigma_{A,0}^2/n_o$, and $\text{Var}(\overline{P}_{prg}) = \tfrac{1}{4}\sigma_A^2(1 + 1/d_{prg}) + (\tfrac{1}{2}\sigma_{A,0}^2 + \sigma_E^2)/n_{prg}$, where d_{prg} is number of dams mated to a single sire when producing offspring for the progeny test. Assuming the same litter size as in the nucleus: $d_{prg} = n_{prg}/n_o$. Furthermore, $\mathbf{g}_2^T$ is identical to $\mathbf{g}_1^T$, but has additional, $\mathbf{g}_2^T(7) = \tfrac{1}{2}\sigma_A^2$.

Genetic variance among selected parents of sex x was

$$\sigma_{A,x}^2 = \sum_k \sigma_A^2(1 - \kappa_k\rho_k^2)\frac{n_k}{N_x} + \sum_k(\mu_k - \mu)^2\frac{n_k}{N_x}, \text{ with } k = 1, c_{max}$$

for sires and $k = c_{max}+1, 2c_{max}$ for dams, where κ_k is Pearson's variance reduction coefficient [$\kappa_k = \iota_k(\iota_k - \tau_k)$, where ι is selection intensity and τ is the standardized truncation point], ρ_k is the accuracy of selection, μ_k is the genetic mean of the parents selected from category k expressed as a deviation from an arbitrary base: $\mu_k = \iota_k\rho_k\sigma_A - k\Delta\mathbf{G}_y$, where $\Delta\mathbf{G}_y$ is the rate of genetic gain per year, and μ is the overall mean of the selected group for sex x: $\mu = \sum_k \mu_k\frac{n_k}{N_x}$. In the equation for $\sigma_{A,x}^2$, the first term represents the weighted sum of the within-age-class genetic variance after selection and the second term represents the additional variance due to differences between the means of parents selected from different age classes.

The variance of the EBV among selected parents of sex x, σ_m^2 and σ_f^2, was $\sigma_x^2 = \sum_k \sigma_{A,k}^2(1 - \kappa_k)\frac{n_k}{N_x} + \sum_k(\mu_k - \mu)^2\frac{n_k}{N_x}$, where $\sigma_{A,k}^2$ is the variance of the EBV among selection candidates in category k.

Accuracy of selection for category k was $\rho_k = \sqrt{\mathbf{b}_{index(k)}^T\mathbf{g}_{index(k)}/\sigma_A^2}$, where $index(k) = 1$ for categories without progeny testing and $index(k) = 2$ for categories with progeny testing. The variance of EBV was $\sigma_{A,k}^2 = \rho_k^2\sigma_A^2$. Genetic gain per year was $\Delta G_y = \tfrac{1}{2}\sigma_A \sum_{k=1}^{2c_{max}} \frac{\iota_k\rho_k n_k}{N_{sex(k)}L_0}$, where L_0 is the generation interval based on the average age of parents of newborn offspring: $L_0 = \tfrac{1}{2}\sum_{k=1}^{2c_{max}} age(k)n_k/N_{sex(k)}$.

Each generation, additive genetic variance was calculated from $\sigma_A^2 = \tfrac{1}{4}\sigma_{A,m}^2 + \tfrac{1}{4}\sigma_{A,f}^2 + \tfrac{1}{2}\sigma_{A,0}^2$. The above equations were iterated until equilibrium variances were reached (approximately 10 iterations).

Intraclass Correlations

Three types of intraclass correlations between EBV of sibs can be distinguished: first, between $\hat{A}_{1,i}$ of two sibs with neither progeny-tested; second, between $\hat{A}_{2,i}$ of two progeny-tested sibs; and, finally, between one tested and one not progeny-tested.

Between full sibs in not-progeny-tested categories, $\rho_{FS,kl} = \mathbf{b}_1^T \mathbf{C}_{FS,kl} \mathbf{b}_1 / \sigma_{\hat{A}_1}^2$, where k and l denote the categories of both individuals, $\mathbf{C}_{FS,kl}$ is the 6×6 covariance matrix between the information sources of an individual in category k and the information sources of its full sib in category l. Matrix $\mathbf{C}_{FS,kl}$ is identical to $\mathbf{V}_1$ except for $\mathbf{C}_{FS,kl}(6,6) = \tfrac{1}{4}\sigma_{A,m}^2 + \tfrac{1}{4}\sigma_{A,f}^2$. Between full sibs in progeny-tested categories, $\rho_{FS,kl} = \mathbf{b}_2^T \mathbf{C}_{FS,kl} \mathbf{b}_2 / \sigma_{\hat{A}_2}^2$, where $\mathbf{C}_{FS,kl}$ is a 7×7 covariance matrix that is identical to $\mathbf{V}_2$ except for $\mathbf{C}_{FS,kl}(6,6) = \tfrac{1}{4}\sigma_{A,m}^2 + \tfrac{1}{4}\sigma_{A,f}^2$, $\mathbf{C}_{FS,kl}(6,7) = \mathbf{C}_{FS,kl}(7,6) = \tfrac{1}{8}\sigma_{A,m}^2 + \tfrac{1}{8}\sigma_{A,f}^2$, and $\mathbf{C}_{FS,kl}(7,7) = \tfrac{1}{16}\sigma_{A,m}^2 + \tfrac{1}{16}\sigma_{A,f}^2$. Between a progeny-tested individual and its not-progeny-tested full-sib, $\rho_{FS,kl} = \mathbf{b}_1^T \mathbf{C}_{FS,kl} \mathbf{b}_2 / \sigma_{\hat{A}_1} \sigma_{\hat{A}_2}$, where $\mathbf{C}_{FS,kl}$ is a 6×7 covariance matrix that is identical to the sub-matrix $\mathbf{C}_{FS,kl}(1-6;1-7)$ for the case of progeny testing in both categories.

Between half-sibs in not-progeny-tested categories, $\rho_{HS,kl} = \mathbf{b}_1^T \mathbf{C}_{HS,kl} \mathbf{b}_1 / \sigma_{\hat{A}_1}^2$, where $\mathbf{C}_{HS,kl}$ is identical to $\mathbf{V}_1$ except for $\mathbf{C}_{HS,kl}(5-6,5-6) = \tfrac{1}{4}\sigma_{A,m}^2$, $\mathbf{C}_{HS,kl}(2,5-6) = \mathbf{C}_{HS,kl}(5-6,2) = 0$, and $\mathbf{C}_{HS,kl}(2,2) = 0$. Between half-sibs in progeny-tested categories, $\rho_{HS,kl} = \mathbf{b}_2^T \mathbf{C}_{HS,kl} \mathbf{b}_2 / \sigma_{\hat{A}_2}^2$, where $\mathbf{C}_{HS,kl}$ is identical to $\mathbf{V}_2$ except for $\mathbf{C}_{HS,kl}(5-6,5-6) = \tfrac{1}{4}\sigma_{A,m}^2$, $\mathbf{C}_{HS,kl}(2,5-6) = \mathbf{C}_{HS,kl}(5-6,2) = 0$, and $\mathbf{C}_{HS,kl}(2,2) = 0$, $\mathbf{C}_{HS,kl}(2,7) = \mathbf{C}_{HS,kl}(7,2) = 0$, $\mathbf{C}_{HS,kl}(5-6,7) = \mathbf{C}_{HS,kl}(7,5-6) = \tfrac{1}{8}\sigma_{A,m}^2$, $\mathbf{C}_{HS,kl}(7,7) = \tfrac{1}{16}\sigma_{A,m}^2$.

Between a progeny-tested individual and its not-progeny-tested half-sib, $\rho_{HS,kl} = \mathbf{b}_1^T \mathbf{C}_{HS,kl} \mathbf{b}_2 / \sigma_{\hat{A}_1} \sigma_{\hat{A}_2}$, where $\mathbf{C}_{HS,kl}$ is the (6×7) covariance matrix, which is identical to the sub-matrix $\mathbf{C}_{HS,kl}(1-6;1-7)$ for the case of progeny testing in both categories.

Intraclass correlations were corrected for the number of families being finite, using the empirical correction of Bijma and Woolliams (2000): $\bar{\rho}_{FS,kl} = \rho_{FS,kl} - \rho_{FS,kl}(1 - \rho_{FS,kl})(0.8634/N_m + 0.9540/N_f)$ and $\bar{\rho}_{HS,kl} = \rho_{HS,kl} - \rho_{HS,kl}(1 - \rho_{HS,kl}^2)(1.4075/N_m + 1.4581/N_f)$.

Appendix B

Prediction of Expected Genetic Contributions

The Π and Λ Matrix. Elements of Π and Λ are a multiple category analogy of the discrete generations case (see Bijma and Woolliams, 2000):

$$\pi_{kl} = \tfrac{1}{2} - \mathbf{b}_{index(k)}^T \mathbf{c}_{index(k),l} \kappa_k / \sigma_{s,l}^2 \quad k,l = 1, 2c_{max} \quad [B1]$$

$$\lambda_{kl} = \mathbf{b}_{index(k)}^T \mathbf{c}_{index(k),l} \iota_k / (\sigma_{s,l}^2 \hat{\sigma}_{A,k}) \quad k,l = 1, 2c_{max} \quad [B2]$$

where $index(k) = 1$ for categories without progeny testing and 2 for categories with progeny testing, and $\mathbf{c}_{index(k),l}$ is a vector of covariances between information sources of an offspring in category k and the selective advantage of its parent in category l. For offspring categories without progeny testing, $\mathbf{c}_{1,l} = [\sigma_{A,l}^2(1 - \kappa_l), \overline{\sigma_{A,f}^2}/d, \overline{\sigma_{A,f}^2}/d, \tfrac{1}{2}\sigma_A^2(1 - \kappa_l \rho_l^2) + \tfrac{1}{2}\overline{\sigma_{A,f}^2}/d, \tfrac{1}{2}\sigma_A^2(1 - \kappa_l \rho_l^2) + \tfrac{1}{2}\overline{\sigma_{A,f}^2}/d, \tfrac{1}{2}\sigma_A^2(1 - \kappa_l \rho_l^2) + \tfrac{1}{2}\overline{\sigma_{A,f}^2}/d]$ for $l = 1, c_{max}$ and $\mathbf{c}_{1,l} = [\sigma_{A,m}^2, \sigma_{A,l}^2(1 - \kappa_l), \sigma_{A,l}^2(1 - \kappa_l/d, \tfrac{1}{2}\sigma_A^2(1 - \kappa_l \rho_l^2)/d + \tfrac{1}{2}\overline{\sigma_{A,m}^2}, \tfrac{1}{2}\sigma_A^2(1 - \kappa_l \rho_l^2) + \tfrac{1}{2}\overline{\sigma_{A,m}^2}, \tfrac{1}{2}\sigma_A^2(1 - \kappa_l \rho_l^2) + \tfrac{1}{2}\overline{\sigma_{A,m}^2}]$ for $l = c_{max}+1, 2c_{max}$. For offspring categories with progeny testing, $\mathbf{c}_{2,l} = [\mathbf{c}_{1,l}(1-6), \tfrac{1}{4}\sigma_A^2(1 - \kappa_l \rho_l^2) + \tfrac{1}{4}\overline{\sigma_{A,f}^2}/d]$ for $l = 1, c_{max}$ and $\mathbf{c}_{2,l} = [\mathbf{c}_{1,l}(1-6), \tfrac{1}{4}\sigma_A^2(1 - \kappa_l \rho_l^2) + \tfrac{1}{4}\overline{\sigma_{A,m}^2}]$ for $l = c_{max}+1, 2c_{max}$. Furthermore, $\overline{\sigma_{A,m}^2} = \sigma_A^2 \sum_{l=1}^{c_{max}} \frac{n_l}{N_m}(1 - \kappa_l \rho_l^2)$, $\overline{\sigma_{A,f}^2} = \sigma_A^2 = \sum_{l=c_{max}+1}^{2c_{max}} \frac{n_l}{N_f}(1 - \kappa_l \rho_l^2)$, $\overline{\sigma_{A,m}^2} = \sum_{l=1}^{c_{max}} \frac{n_l}{N_m}\sigma_{Al}^2(1 - \kappa_l)$ and $\overline{\sigma_{A,f}^2} = \sum_{l=c_{max}+1}^{2c_{max}} \frac{n_l}{N_f}\sigma_{A,l}^2(1 - \kappa_l)$.

The G Matrix. Elements of $\mathbf{G}$ are (Bijma and Woolliams, 1999):

$$g_{kl} = \tfrac{1}{2}\frac{p_{kl}}{p_k}\frac{n_l}{N_{sex(l)}} \quad k,l = 1, 2c_{max} \quad [B3]$$

where p_{kl} is the selected proportion among offspring in category k descending from parents in category l. Solutions for p_{kl} were obtained separately for each parent sex, by simultaneously solving the equations $p_k = \sum_l p_{kl} n_l / N_{sex(l)}$ and $p_{kl} = 1 - \Phi[(I_k - \tfrac{1}{2}\mu_l)/\sigma_{\hat{A},k}|l]$ using the algorithm RIDDR_ROOT (Press et al., 1992), where $l = 1, c_{max}$ for sires, $l = c_{max}+1, 2c_{max}$ for dams, I_k is the EBV truncation point common to all offspring in category k, $\sigma_{\hat{A},k}|l$ is the standard deviation of the EBV for offspring in category k descending from parents in category l, μ_l is given in Appendix A, and Φ denotes the cumulative normal distribution function (Bijma and Woolliams, 1999). The standard deviation of the EBV, $\sigma_{\hat{A},k}|l$, was calculated analogous to $\sigma_{\hat{A},k}$ by using the pseudo-BLUP index (Appendix A), but replacing $\sigma_{A,s}^2(\sigma_{A,d}^2)$ by $\sigma_A^2(1 - \kappa_l \rho_l^2)$ and $\sigma_m^2(\sigma_f^2)$ by $\sigma_{A,l}^2(1 - \kappa_l)$ when the parent is a male (female).

The D Matrix. Elements of $\mathbf{D}$ were obtained, separately for each parent sex, from (Bijma and Woolliams, 1999):

$$d_{kl} = \tfrac{1}{2}\mu_l + \iota_{kl}\sigma_{\hat{A},k}|l - \overline{d}_k \quad [B4]$$

where μ_l is given in Appendix A, ι_{kl} is the selection intensity corresponding to p_{kl} (see above) and the second

term represents subtraction of the average: $\bar{d}_k = \sum_l (\frac{1}{2}\mu_l + \iota_{kl}\sigma_{\hat{A},k|l})$, with $l = 1$, c_{max} for sires, $l = c_{max}+1$, $2c_{max}$ for dams.

Appendix C

Step 2: Expectation of Squared Expected Contributions

Equation 16. $E[(A_i - \overline{A}_k)(A_i - \overline{A}_{c_{max}})] = E(A_i^2) - E(A_i\overline{A}_k) - E(A_i\overline{A}_{c_{max}}) + E(\overline{A}_k \overline{A}_{c_{max}})$, where i refers to one of the $n_{k,c_{max}}$ individuals that are selected both in category k and in category c_{max}. When deriving those expectations one has to account for the fact that $\overline{A}_k$ and $\overline{A}_{c_{max}}$ are finite sample means. For example, $E(A_i\overline{A}_k) = E[A_i(A_{1,k} + \ldots + A_{n_k,k})]/n_k = [(n_k - 1)E(A_iA_{j\neq i,k}) + E(A_i^2)]/n_k \approx (n_k - 1)E(A_i)\overline{A}_k/n_k + E(A_i^2)/n_k$. Following this approach, it can be shown that:

$$E[(A_i - \overline{A}_k)(A_i - \overline{A}_{c_{max}})] =$$

$$[(n_k n_{c_{max}} - n_k - n_{c_{max}} + n_{k,c_{max}})/(n_k n_{c_{max}})] \, E(A_i^2)$$

$$- [(n_k - 1)/n_k] \, E(A_i) \, \overline{A}_k - [(n_{c_{max}} - 1)/n_{c_{max}}]E(A_i)\overline{A}_{c_{max}} \quad \text{[C1]}$$

$$+ [1 - n_{k,c_{max}}/(n_k n_{c_{max}})]\overline{A}_k \overline{A}_{c_{max}}$$

where $\overline{A}_k = \iota_k \rho_k \sigma_A$. This equation has the desired property that it reduces to zero when $n_k = n_{c_{max}} = n_{k,c_{max}} = 1$ and that it reduces to $E(A_i^2) - E(A_i) \overline{A}_k - E(A_i)\overline{A}_{c_{max}} + \overline{A}_k \overline{A}_{c_{max}}$ when n_k, $n_{c_{max}}$, $n_{k,c_{max}} \to \infty$.

The terms $E(A_i)$ and $E(A_i^2)$ represent the expectation of the breeding value and the squared breeding value of individuals that are selected both in category k and in category c_{max}, for example, $E(A_i) = E(A_i|\hat{A}_{i,k} > \tau_k, \hat{A}_{i,c_{max}} > \tau_{c_{max}})$, which can be obtained from the moment-generating function of the truncated multivariate normal distribution (Tallis, 1961). From the first and second equations on page 226 of Tallis (1961), with, in the notation of Tallis, $a_1 = A_{\cdot 1} = A_{:1} = -\infty$, it follows that:

$$E(A_i) = \sigma_A[\rho_k\phi(\tau_k)\Phi(\Psi_{k,c_{max}}) \quad \text{[C2]}$$

$$+ \rho_{c_{max}}\phi(\tau_{c_{max}})\Phi(\Psi_{c_{max},k})]p_{k,c_{max}}^{-1}$$

and

$$E(A_i^2) = \sigma_A^2\{p_{k,c_{max}} + \rho_k^2\tau_k\phi(\tau_k)\Phi(\Psi_{k,c_{max}})$$

$$+ \rho_{c_{max}}^2\tau_{c_{max}}\phi(\tau_{c_{max}})\Phi(\Psi_{c_{max}k})$$

$$+ \phi(\tau_k,\tau_{c_{max}},\rho_{k,c_{max}})[\rho_k(\rho_{c_{max}} - \rho_k\rho_{k,c_{max}}) \quad \text{[C3]}$$

$$+ \rho_{c_{max}}(\rho_k - \rho_{c_{max}}\rho_{k,c_{max}})]\}p_{k,c_{max}}^{-1}$$

where ρ_k is the accuracy of selection in category k, $\rho_{k,c_{max}}$ is the correlation between the index of an individual in category k and its index in category c_{max}: $\rho_{k,c_{max}} = \sigma_{\hat{A},k}/\sigma_{\hat{A},c_{max}}$, $\phi(\tau)$ is the univariate normal density function: $\phi(\tau) = (2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}\tau^2}$, $\phi(\tau_k,\tau_l,\rho_{k,l})$ is the bivariate normal density function: $\phi(\tau_k,\tau_l,\rho_{k,l}) = \left(2\pi\sqrt{1 - \rho_{k,l}^2}\right)^{-1}e^{-\frac{1}{2}q}$ with $q = (\tau_k^2 - 2\rho_{k,l}\tau_k\tau_l + \tau_l^2)/(1 - \rho_{k,l}^2)$, and $\Phi(\Psi)$ is the univariate normal upper tail proportion: $\Phi(\Psi) = \int_\psi^\infty \phi(x)dx$ and $\psi_{k,l} = (\tau_l - \rho_{k,l}\tau_k)(1 - \rho_{k,l}^2)^{-\frac{1}{2}}$. Note that $\Psi_{k,l} \neq \Psi_{l,k}$.

Appendix D

Calculation of δ

General equations for calculating $\Delta V_{n,k}$ are given in Appendix E of Woolliams and Bijma (2000) and can also be obtained by extending the equations for BLUP selection with discrete generations (Bijma and Woolliams, 2000) to populations with overlapping generations. To keep notation as short as possible, the number of selected offspring conditional on the selective advantage, $n_{ij,k}(l)|s_{i,k}$, is abbreviated by $n_{ij,k}(l)$.

For sires, $\Delta \mathbf{V}_{n,k}$ is calculated from:

$$\Delta \mathbf{V}_{n,k}(l,l) = E_s\{n_{i*,k}(l)[n_{i*,k}(l) - 1]\} - E_s[\mu_k^2(l)] \quad \text{[D1]}$$

$$k = 1,c_{max}; l = 1, 2c_{max}$$

$$\Delta \mathbf{V}_{n,k}(l,l') = E_s[n_{i*,k}(l)n_{i*,k}(l')] - E_s[\mu_k(l)\mu_k(l')] \quad \text{[D2]}$$

$$k = 1,c_{max}; l,l' = 1, 2c_{max}; l \neq l'$$

and for dams from:

$$\Delta \mathbf{V}_{n,k}(l,l) = E_s\{n_{ij,k}(l)[n_{ij,k}(l) - 1]\} - E_s[\mu_k^2(l)] \quad \text{[D3]}$$

$$k = c_{max}+1, 2c_{max}; l = 1, 2c_{max}$$

$$\Delta \mathbf{V}_{n,k}(l,l') = E_s[n_{ij,k}(l)n_{ij,k}(l')] - E_s[\mu_k(l)\mu_k(l')] \quad \text{[D4]}$$

$$k = c_{max}+1, 2c_{max}; l,l' = 1, 2c_{max}; l \neq l'$$

where $n_{i*,k}(l)$ is the number of selected offspring in category l from the ith sire in category k and $n_{ij,k}(l)$ is the number of selected offspring in category l from the jth dam in category k that is mated to sire i, (i.e., n_{i*} represents the sire family size and n_{ij} represents the dam family size). In Eq. [D1] and [D2], $\mu_k(l) = E[n_{i*,k}(l)|s_{i,k}]$, which is the expected number of selected offspring in category l of sire i in category k, given its selective advantage. In Eq. [D3] and [D4], $\mu_k(l) = E[n_{ij,k}(l)|s_{j,k}]$.

Elements of Eq. [D1] to [D4] are as follows:

$$E_s[\mu_k^2(l)] = 4g_{lk}^2 n_l^2 n_k^{-2}(1 + \lambda_{lk}^2\sigma_{s_k}^2) \quad \text{[D5]}$$

$$k, l = 1, 2c_{max}$$

$$E_s[\mu_k(l)\mu_k(l')] = 4g_{lk}g_{l'k}n_ln_{l'}n_k^{-2}(1 + \lambda_{lk}\lambda_{l'k}\sigma_{s_k}^2) \quad [D6]$$

$$k, l, l' = 1, 2c_{max}; l \neq l'$$

$$E_s\{n_{i*,k}(l)[n_{i*,k}(l) - 1]\} = dE_s\{n_{ij,k}(l)[n_{ij,k}(l) - 1]\}$$
$$+ d(d - 1)E_s\{n_{ij,k}(l)n_{ij',k}(l)\} \quad [D7]$$
$$k = 1, c_{max}; l = 1, 2c_{max}$$

$$E_s[n_{i*,k}(l)n_{i*,k}(l')] = dE_s[n_{ij,k}(l) \ n_{ij,k}(l')]$$
$$+ d(d - 1)E_s[n_{ij,k}(l)n_{ij',k}(l')] \quad [D8]$$
$$k = 1, c_{max}; l, l' = 1, 2c_{max}; l \neq l'$$

where j' is dam other than j,

$$E_s\{n_{ij,k}(l)[n_{ij,k}(l) - 1]\} \quad [D9]$$

$$= \frac{2n_{sex(l)}(n_{sex(l)} - 1)g_{lk}n_l(2g_{lk}n_l - 1)R(p_{lk},p_{lk},\bar{\rho}_{FS,ll})}{n_kN_{sex(k)}^{-1}T_{sex(l)} \ (n_kN_{sex(k)}^{-1}T_{sex(l)} - 1)}$$

$$E_s[n_{ij,k}(l)n_{ij,k}(l')] \quad [D10]$$

$$= \frac{4n_{sex(l)}n_{sex(l')}g_{lk}n_lg_{l'k}n_{l'}R(p_{lk},p_{l'k},\bar{\rho}_{FS,ll'})}{n_k^2N_{sex(l)}^{-2}T_{sex(l)}T_{sex(l')}}$$

$$E_s[n_{ij,k}(l)n_{ij',k}(l)] \quad [D11]$$

$$= \frac{2n_{sex(l)}^2g_{lk}n_l(2g_{lk}n_l - 1)R(p_{lk},p_{lk},\bar{\rho}_{HS,ll})}{n_kN_{sex(k)}^{-1}T_{sex(l)} \ (n_kN_{sex(k)}^{-1}T_{sex(l)} - 1)}$$

$$E_s[n_{ij,k}(l)n_{ij',k}(l')] \quad [D12]$$

$$= \frac{4n_{sex(l)}n_{sex(l')}g_{lk}n_lg_{l'k}n_{l'}R(p_{lk},p_{l'k},\bar{\rho}_{HS,ll'})}{n_k^2N_{sex(k)}^{-2}T_{sex(l)}T_{sex(l')}}$$

where $T_{sex(l)}$ is the total number of candidates of the sex of category l and $n_{sex(l)}$ is the number of offspring of *sex(l)* born per dam ($T_m = T_f = T$ and $n_m = n_f = \frac{1}{2}n_o$ for the current breeding scheme). Note that Eq. [D9] and [D10] are used both in Eq. [D3] and [D4] and in Eq. [D7] and [D8]; they are used in [D3] and [D4] for dam categories ($k = c_{max}+1,2c_{max}$) and in [D7] and [D8] for sire categories ($k = 1, c_{max}$).

Furthermore, from Mendel and Elston (1974), $R(p_{lk}, p_{l'k}, \bar{\rho}_{sibs,ll'}) = p_{lk}/\Phi[(\iota_{l'k}\bar{\rho}_{sibs,ll'} - \tau_{lk})(1 - \kappa_{l'k}\bar{\rho}_{sibs,ll'}^2)^{-\frac{1}{2}}]$, where Φ is the cumulative normal distribution function and τ_{lk} is the standardized truncation point for offspring in category l descending from parents in category k. When $l' \neq l$, the most accurate value is obtained by using l for the category with the smallest selection intensity (Wray et al., 1994).

In situations in which the number of selected parents is small compared to the number of selection candidates per family, Eq. [D9] to [D12] can give substantial bias. For those cases, accuracy of Eq. [D9] to [D12] can be improved by adjusting the selected proportion according to Wray et al. (1990). Adjusted selected proportions for sire categories were (see Appendix B for unadjusted selected proportions) as follows:

$$p_{kl,adj} = (1 - \bar{\rho}_{sibs,kl})p_{kl} + \bar{\rho}_{sibs,kl}\max(p_{kl},1/n_l) \quad [D13]$$

$$l = 1, c_{max}$$

and for dam categories

$$p_{kl,adj} = (1 - \bar{\rho}_{sibs,kl})p_{kl} + \bar{\rho}_{sibs,kl}\max(p_{kl},1/(n_l/d)) \quad [D14]$$

$$l = c_{max}+1, 2c_{max}$$

In the present study, this adjustment was applied to all schemes where, for any of the sire categories, $\min(n_k) \leq 1.5n_o$.

Citations   This article has been cited by 2 HighWire-hosted articles: