A general procedure for predicting rates of inbreeding in populations undergoing mass selection

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Genetics

Publisher Rights Statement:
Copyright 2000 by the Genetics Society of America

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
A General Procedure for Predicting Rates of Inbreeding in Populations Undergoing Mass Selection

Piter Bijma,* Johan A. M. Van Arendonk* and John A. Woolliams1

*Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen Agricultural University, 6700 AH Wageningen, The Netherlands and 1Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, United Kingdom

ABSTRACT

Predictions of rates of inbreeding ($\Delta F$), based on the concept of long-term genetic contributions assuming the infinitesimal model, are developed for populations with discrete or overlapping generations undergoing mass selection. Phenotypes of individuals are assumed to be recorded prior to reproductive age and to remain constant over time. The prediction method accounts for inheritance of selective advantage both within and between age classes and for changing selection intensities with age. Terms corresponding to previous methods that assume constant selection intensity with age are identified. Predictions are accurate (relative errors $<8\%$), except for cases with extreme selection intensities in females in combination with high heritability. With overlapping generations $\Delta F$ reaches a maximum when parents are equally distributed over age classes, which is mainly due to selection of the same individuals in consecutive years. $\Delta F/\text{year}$ decreases much more slowly compared to $\Delta F/\text{generation}$ as the number of younger individuals increases, whereas the decrease is more similar as the number of older individuals increases. The minimum $\Delta F$ (per year or per generation) is obtained when most parents were in the later age classes, which is mainly due to an increased number of parents per generation. With overlapping generations, the relationship between heritability and $\Delta F$ is dependent on the age structure of the population.

In the absence of selection and with a Poisson distribution of family size, expected rates of inbreeding are related directly to the number of parents: $E(\Delta F) = 1/8N_m + 1/8N_f$ (Wright 1969, p. 212). In selected populations, however, superior families contribute more offspring to the next generation than average families. This increases the rate of inbreeding of a selected population compared to an unselected population. Prediction of rates of inbreeding in selected populations is difficult, because selection decisions are correlated over generations due to the inheritance of selective advantage. Methods accounting for only one or two generations of selection (e.g., Burrows 1984a,b) therefore generally underestimate the rate of inbreeding (Wray et al. 1990; see Caballero 1994 for a review).

Two approaches to prediction of rates of inbreeding for selected populations can be distinguished. First, rates of inbreeding can be predicted on the basis of the variance of allele frequency, using the idea of accumulation of selective advantages over generations (Robertson 1961). Using this approach and equilibrium genetic variances, Santiago and Caballero (1995) obtained accurate predictions for populations with discrete generations under mass selection. Nomura (1996) extended that method to populations with overlapping generations and equal numbers of parents per sex in every age class. Second, rates of inbreeding can be predicted using the concept of long-term genetic contributions. Rates of inbreeding are proportional to the sum of squared long-term genetic contributions of ancestors (Wray and Thompson 1990). Wray and Thompson (1990) obtained accurate predictions of rates of inbreeding for populations with discrete generations under mass selection, using iterative regression methods. For discrete generations and mass selection a closed form expression was obtained by Woolliams et al. (1993). For more complicated selection schemes, however, predictions became unmanageable due to the recursive nature of the procedure and the need for predicting the variance of long-term genetic contributions (Wray et al. 1994).

Recently, Woolliams and Bijma (2000) showed that the variance of long-term genetic contributions is related to their squared expectation, making a separate prediction of the variance redundant. Furthermore, Woolliams et al. (1999) obtained general predictions of expected genetic contributions using equilibrium genetic variances instead of second generation genetic variances (Woolliams et al. 1993). Using the approach of Woolliams et al. (1999), Bijma and Woolliams (1999) obtained accurate predictions of genetic contributions for populations with overlapping generations under mass or sib-index selection. However, they did not develop predictions for rates of inbreeding for those schemes.

The aim of this article is twofold. First, explicit predic-
tion equations for rates of inbreeding in populations with discrete or overlapping generations under mass selection are developed, on the basis of the theory of Woolliams et al. (1999) and Woolliams and Bijma (2000). These predictions are valid for any distribution of parents across age classes, overcoming the restriction of Nomura (1996), to give a general and practical method for mass selection with overlapping generations. These methods are compared to methods of Santiago and Caballero (1995) for discrete generations, and to methods of Nomura (1996) for the special case of equal numbers of parents per age class with overlapping generations. The accuracy of predictions is examined using simulation. Second, relationships between rates of inbreeding and genetic or population parameters are examined, and differences between populations with discrete and overlapping generations are presented and discussed.

**DERIVATION OF EXPRESSIONS**

**Population model:** This section describes the population and the selection procedures for which rates of inbreeding are predicted. This model is also used in the simulation. The trait considered is assumed to be determined by an infinite number of additive loci, each having an infinitesimal effect (infinitesimal model; Fisher 1918). Phenotypic values are the sum of additive genetic values (breeding values) and environmental values, \( P = A + E \). A population with either discrete or overlapping generations under mass selection is modeled. With parents up to a maximum age of \( c_{\text{max}} \) there are \( 2c_{\text{max}} \) categories, one for each sex and age of parent. Categories are indexed by \( k \) or by \( l \), so \( k = 1 \ldots c_{\text{max}} \) are males, and \( k = c_{\text{max}} + 1 \ldots 2c_{\text{max}} \) are females. With discrete generations, there are only two categories: males and females that are indexed by \( s = m \) or \( f \). Before reproductive age, phenotypes of individuals are recorded and remain unchanged over time, so that ranking of individuals within categories is constant over time. Within categories, individuals are ranked on their phenotype and each year the highest-ranking \( n_i \) individuals are selected from the \( k \)th category, to produce the next cohort. The total number of male and female parents of each cohort is \( N_m = \sum_{i=1}^{c_{\text{max}}} n_i \), and \( N_f = \sum_{i=c_{\text{max}}+1}^{2c_{\text{max}}} n_i \), respectively. Each sire is mated at random to \( d \) dams (\( d = N_f / N_m \)), and each dam produces a fixed number, \( n_o \), of offspring (\( \frac{1}{2} n_o \) of each sex), so that for each sex the total number of offspring born in a cohort is \( N = \frac{1}{2} n_o N_f \). The unit of age, i.e., the interval between consecutive age classes, was 1 year. Genetic contributions and rates of inbreeding per year therefore will be equal to genetic contributions and rates of inbreeding per cohort. The difference between the current prediction and the method of Woolliams et al. (1993) is (1) the current prediction is based on equilibrium genetic variances, which simplifies the prediction of \( u_{ij} \) (Woolliams et al. 1999); (2) the variance of genetic contributions is not predicted separately, since it is related to the mean (Woolliams and Bijma 2000).

**Discrete generations**

**Step 1: prediction of expected long-term genetic contributions:** Expected genetic contributions of ancestors are obtained from the linear model (Bijma and Woolliams 1999),

\[
E(\Delta F) = \frac{1}{2} \mathbf{1}' \mathbf{N} E(\mathbf{u}) + \frac{1}{2} \mathbf{1}' \mathbf{N} \delta,
\]
where \( s \) denotes males or females, \( \alpha_i \) is the expected contribution for an average ancestor of sex \( s \), and \( \beta_i \) is the regression coefficient of the contribution on the breeding value (\( A_{i,s} \)) of the ancestor as a deviation from the average of the selected group (\( \bar{A}_s \)) for sex \( s \). In discrete generations, \( \alpha_i = 1/(2N_p) \) and \( \beta_i = \alpha_i \lambda / (1 - \pi) \), where \( \lambda = \frac{1}{2} \sigma^2_p \) is the average regression coefficient of the number of selected male and female offspring on the breeding value of the parent, and \( \pi = \frac{1}{2} (1 - k \lambda^2) \) is the average regression coefficient of the breeding value of selected male and female offspring on the breeding value of the parent (Bijma and Woolliams 1999). Here, \( i = \frac{1}{2} (i_m + i_f) \) is selection intensity, \( k = \frac{1}{2} (k_m + k_f) \) is Pearson’s (1903) variance reduction coefficient, and \( h^2 = \sigma^2_A / \sigma^2_S \), where \( \sigma^2_A \) and \( \sigma^2_S \) are Bulmer’s (1971) equilibrium genetic and phenotypic variances.

**Step 2: derivation of \( E(\hat{u}_i) \):** Substituting Equation 2 and with terms added for the mate,

\[
E(\hat{u}_{ij}^2) = \alpha_i^2 + \beta_i^2 E((A_{ij} - \bar{A}_m)^2)
\]

\[
+ \sum_{j=1}^{d} \beta_j^2 E((A_{ij} - \bar{A}_m)^2)
\]

\[
E(\hat{u}_{ij}) = \alpha_i + \beta_i E((A_{ij} - \bar{A}_m))
\]

\[
+ \frac{1}{d} \beta_{n,b} E((A_{ij,m} - \bar{A}_m)^2),
\]

where \( j \) denotes the mate and

\[
E((A_{ij} - \bar{A}_m)^2) = (1 - 1/\ N_s) \sigma^2_A (1 - k h^2).
\]

From Equation 1, ignoring the second term, \( E(\Delta F) = \frac{1}{2} \left( N_m E(\hat{u}_{ij}^2) + N_s E(\hat{u}_{ij}^2) \right) \). From Equation 3 and 4 and the equations for \( \beta, \lambda, \) and \( \pi \), predicted \( \Delta F \) (see appendix a) is

\[
E(\Delta F) = \frac{1}{8N_m} + \frac{1}{8N_s}
\]

\[
+ \left[ \frac{h^2}{4(1 + k h^2)} \right] \left( 1 - \frac{1}{N_m} \right) \left( 1 - \frac{1}{2N_m} \right) \left( 1 - \frac{1}{2N_s} \right)
\]

\[
+ \left( 1 - \frac{k h^2}{1 - \frac{1}{N_m}} \right) \left( 1 - \frac{1}{N_s} \right)
\]

For \( N_m = N_s = \frac{1}{2} N \), the result simplifies to

\[
E(\Delta F) = \frac{1}{2 N} + \frac{1}{N} \left[ \frac{h^2}{1 - \frac{1}{2 N}} \right].
\]

The assumption for Equations 6 and 7 is that, conditional on \( (A_{i,j} - \bar{A}_m) \) in mass selection [i.e., conditional on \( (A_{i,s} - \bar{A}_s) \) in mass selection] family size follows a Poisson distribution (Woolliams and Bijma 2000), which is approximately the case with mass selection when \( n_o \sim \text{Poisson} \). A numerical example is in appendix a.

**Step 3: Correction of \( E(\Delta F) \) for deviations of \( V_x \) from Poisson variances:** With fixed \( n_o \), family size follows a hypergeometric distribution (Burrows 1984b) and a correction is required according to the second term of Equation 1. In this article, the hypergeometric variance is approximated by a binomial variance, which simplifies the prediction. For more complicated selection strategies, e.g., index selection, a hypergeometric variance may be required (Woolliams and Bijma 2000).

With discrete generations, the second term of Equation 1 reduces to \( \frac{1}{2} \left( N_m \delta_{i} + N_s \delta_{j} \right) \), where \( \delta_i = \alpha_i^2 \sqrt{V_{\text{dev},i}} \), \( \delta_j = \alpha_j^2 \sqrt{V_{\text{dev},j}} \), and \( V_{\text{dev},i} \) is the 2 \times 2 matrix of deviations of the (co)variance of family size from Poisson variances for sex \( s \) (Woolliams and Bijma 2000). Diagonal elements of \( V_{\text{dev},i} \) are obtained as \( V_{\text{dev},i} = V_{\text{dev}} - \text{Poisson} \), which are of the form \( np(1 - p) - np = -np^2 \), where \( n \) is the number of candidates and \( p \) is the selected proportion. Off-diagonal elements of \( V_{\text{dev},i} \) are zero. For discrete generations the total correction (appendix a) equals

\[
\frac{1}{2} \left( 1 - \frac{1}{N} \right) \delta = \frac{-1}{8}. \]

**Relation to Santiago and Caballero (1995):** The prediction equation of Santiago and Caballero (1995) can be related directly to the current prediction. With random mating and assuming \( \alpha_i = \alpha_j = 0 \) (see Santiago and Caballero 1995 for notation), Equations 21 and 36 of Santiago and Caballero (1995) reduce to \( \frac{1}{2} \left( E(\hat{u}_{ij}) \right)^2 \left( 1 - \frac{1}{2N_m} \right) \left( 1 - \frac{1}{2N_s} \right) \) and \( Q = 1/\left( 1 - \frac{1}{2k h^2} \right) \), which is identical to our 1/\left( 1 - \frac{1}{2k h^2} \right). Furthermore, they use \( C_{ij} = \frac{1}{2} (1 - k h^2) \), which is identical to our \( 2 \lambda E(\left( A_{ij} - \bar{A}_m \right)^2) \), where the 2 accounts for the mate.

The correction for deviations of \( V_x \) from Poisson variances can also be related to Equation 36 of Santiago and Caballero (1995). They use \( V_{\text{dev}} \) as \( \frac{n_0}{N} \left( 1 - \frac{1}{N} \right) \left( 1 - \frac{1}{N} \right) \) (see Santiago and Caballero 1995, Equation 30 and ignore the term \( C_{ij}^2 \), where \( n_i \) is the number of selection candidates per sex of a parent of sex \( s \) \( h_{ij} = \frac{n_i}{\sqrt{n_o d f}}, h_{ij} = \frac{n_i}{\sqrt{n_d s}} \) and \( s' \) denotes the sex of the offspring. This is a binomial variance. The deviation from a Poisson variance (i.e., \( N_s/\ N \)) equals \( V_{\text{dev}} - \text{Poisson} \). From Equation 36 of Santiago and Caballero (1995), the total correction of \( \Delta F \) equals \( -\gamma^2 \), which is identical to Equation 8 (Santiago et al. 1999). Therefore, Equations 21, 30, and 36 of Santiago and Caballero (1995) are identical to the current prediction for mass selection. A numerical difference between both methods exists because Santiago and Caballero (1995) omit the correction for a finite number of parents when calculating their \( C_{ij} \), which would be equivalent to omitting the \( \frac{1}{N} \) in Equation 5 of the current prediction.
Overlapping generations

Step 1: prediction of expected long-term genetic contributions: Genetic contributions are predicted using Equation 2 again, but now categories refer to sex-age class combinations, which are indexed by \( k \) instead of \( s \), so that \( k = 1 \ldots 2c_{\text{max}} \) and \( u_{i,k} \) is the expected genetic contribution of individual \( i \) originating from its selection in category \( k \). Solutions for \( \alpha \) and \( \beta \) are obtained from Woolliams et al. (1999),

\[
N\alpha = [G^T + (G^T \ast D^T)(1 - G^T \ast \Pi^T)^{-1}(G^T \ast \Lambda^T)]N\alpha
\]

(9)

\[
N\beta = (1 - G^T \ast \Pi^T)^{-1}(G^T \ast \Lambda^T)(N\alpha)
\]

(10)

where \( \ast \) denotes element-by-element multiplication, \( T \) denotes the transpose of matrices, \( I \) is the \( 2c_{\text{max}} \times 2c_{\text{max}} \) identity matrix, \( N \) is a \( 2c_{\text{max}} \times 2c_{\text{max}} \) diagonal matrix containing the numbers of parents selected from every category \( (n_{ik}) \), \( \Pi \) is a \( 2c_{\text{max}} \times 2c_{\text{max}} \) matrix with each element \( \mu_{ik} \), being the regression coefficient of the breeding value of a selected offspring in category \( k \) on the breeding value of the parent in category \( l \). \( \Lambda \) is a \( 2c_{\text{max}} \times 2c_{\text{max}} \) matrix with each element \( \lambda_{ik} \), being the regression coefficient of the number of selected offspring in category \( k \) on the breeding value of the parent in category \( l \). \( G \) is a \( 2c_{\text{max}} \times 2c_{\text{max}} \) modified gene flow matrix connecting selected offspring to parental categories, \( D \) is a \( 2c_{\text{max}} \times 2c_{\text{max}} \) matrix of deviations of breeding values from the mean of the selected group, \( \alpha \) is a \( 2c_{\text{max}} \) vector of elements \( \alpha_{ik} \), and \( \beta \) is a \( 2c_{\text{max}} \) vector of elements \( \beta_{ik} \). Generation interval (\( L \)) was calculated as the time interval in which genetic contributions sum to \( 1: L = 1/[\Sigma_{k=1}^{2c_{\text{max}}} k\mu_{ik}] \) (Woolliams et al. 1999). More details and a numerical example are in Bijma and Woolliams (1999).

Contributions predicted from Equations 9 and 10 are per year; i.e., they are the long-term contribution of a single cohort, not of a total generation. Rates of inbreeding predicted from these contributions therefore are also per year.

Step 2: derivation of \( E(u_{ik}^2) \): For the calculation of \( E(u_{ik}^2) \) one needs to find the lifetime expected genetic contribution; i.e., one has to account for the fact that individuals may be selected in multiple categories. With \( c_{\text{max}} \) age classes per sex and the ranking of individuals within age classes remaining constant, there are \( 2c_{\text{max}} \) exclusive categories, which will be indexed by \( s \), i.e., individuals selected once, twice, up to \( c_{\text{max}} \) times for each sex. Therefore, \( s = 1 \ldots c_{\text{max}} \) denotes males selected 1 through \( c_{\text{max}} \) times and \( s = c_{\text{max}} + 1 \ldots 2c_{\text{max}} \) denotes females selected 1 through \( c_{\text{max}} \) times. The expected lifetime contribution for these categories is \( u_{i,s} = \Sigma u_{i,k} \), where the sum is taken over the age-sex categories \( k \) from which \( i \) is selected. Thus individuals are indexed in two different ways, i.e., by whether or not they were selected at a specific age, denoted by \( k \), and by how many times they were selected throughout their lifetime, denoted by \( s \).

The scalar equivalent of the first term of Equation 1 is

\[
\frac{1}{2} \sum_{s=1}^{c_{\text{max}}} \sum_{k=1}^{c_{\text{max}}} n_{ik}E(u_{ik}^2)
\]

with the first term denoting males and the second females. The summation over exclusive categories \( s \) can be written in terms of the categories \( k \), for males,

\[
\sum_{s=1}^{c_{\text{max}}} \sum_{k=1}^{c_{\text{max}}} n_{ik}E(u_{ik}^2)
\]

\[
+ 2 \sum_{k=1}^{c_{\text{max}}-1} \sum_{s=k+1}^{c_{\text{max}}} \min(n_{ik}, n_{is})E(u_{ik}u_{is}),
\]

(11)

and for females,

\[
\sum_{s=1}^{c_{\text{max}}} \sum_{k=1}^{c_{\text{max}}} n_{ik}E(u_{ik}^2)
\]

\[
+ 2 \sum_{s=1}^{c_{\text{max}}-1} \sum_{k=1}^{c_{\text{max}}-1} \sum_{l=k+1}^{c_{\text{max}}} \min(n_{ik}, n_{il})E(u_{ik}u_{il}),
\]

(12)

where \( \min(n_{ik}, n_{is}) \) denotes the minimum of \( n_{ik} \) and \( n_{is} \) (see also example in appendix b). These summations can be written in matrix form, so that for Poisson family size, the rate of inbreeding per year is

\[
E(\Delta F_t) = \frac{1}{2} I N N_{\text{a}} J,
\]

(13)

where \( I = (1 1 \ldots 1)^T \) and \( N_{\text{a}} \) is similar to \( N \) but has a reordering of age classes within sexes so that they go from large to small according to the number of parents, and \( J \) is a \( 2c_{\text{max}} \times 2c_{\text{max}} \) matrix containing a lower triangular submatrix for each sex (with categories ordered as in \( N_{\text{a}} \)), with \( E(u_{ik}^2) \) on the diagonal and \( 2E(u_{ik}u_{ij}) \) as off-diagonals in the lower triangular submatrices (see example in appendix b). Note that, although Equation 1 uses exclusive categories \( s \), we have expressed \( \Delta F_t \) in terms of the age-sex categories \( k \) in Equation 13. Thus, the expected genetic contributions for the categories \( k \) can be used directly in Equation 13. Rates of inbreeding per generation were calculated as \( E[\Delta F_t] = LE(\Delta F_t) \).

As with discrete generations, \( E(u_{ik}^2) \) has to include terms for the mates. With overlapping generations, the mate term consists of two elements. The first element is due to the category of the mate as a deviation of the average category for the sex of the mate, \( \alpha_{ik} - \overline{\alpha}_{\text{sex}} \). The second term is due to the selective advantage of the mate within its category, \( \beta_{ik}(A_{ik} - \overline{A}_{i}) \). Therefore, for males, \( u_{ik} = \alpha_{ik} + \beta_{ik}(A_{ik} - \overline{A}_{i}) + \Sigma_{j=1}^{c_{\text{max}}-1} \left( \alpha_{ij} - \overline{\alpha}_{\text{sex}} \right) + \beta_{ij}(A_{ij} - \overline{A}_{j}) \); and for females, \( u_{ik} = \alpha_{ik} + \beta_{ik}(A_{ik} - \overline{A}_{i}) + \left( \alpha_{ik} - \overline{\alpha}_{\text{sex}} \right) \beta_{ij}(A_{ij} - \overline{A}_{j})/d \), where \( d \) denotes the mate, \( l \) the category of the mate, and \( s \) the sex of the mate. For Equations 11 and 12, expectations of squared contributions are obtained for males as
E(u_{i,i}) = \alpha_i^2 + (1 - 1/n_i)\beta_i^2\sigma_i^{2}(1 - \kappa_i h_i) + \delta_i^2 - \alpha_i^2 + \sigma_i^2(1 - 1/n_i)\beta_i(1 - \kappa_i h_i)\{1\}

(14)

where \( k = 1 \ldots c_{\text{max}} \), and for females as

E(u_{i,i}) = \alpha_i^2 + (1 - 1/n_i)\beta_i^2\sigma_i^{2}(1 - \kappa_i h_i)

+ \frac{1}{d^2}\delta_i^2 - \alpha_i^2 + \sigma_i^2(1 - 1/n_i)\beta_i(1 - \kappa_i h_i)\{1\}

(15)

where \( k = c_{\text{max}} + 1 \ldots 2c_{\text{max}} \) and bars with subscripts \( m \) or \( f \) denote weighted averages over mate categories.

Cross-products in Equations 11 and 12 arise only from the individuals selected in both categories, which are all the individuals selected from the smallest category [i.e., \( \min(n_f, n_m) \)]. Cross-products are therefore

E(u_{i,j},u_{i,j}) = \alpha_{i,j}\alpha_i + (1 - 1/\min(n_f, n_m))\beta_{i,j}\sigma_{i,j}(1 - \max(\kappa_i, \kappa_j) h_i)

+ \alpha_{i,j}\beta_{i,j}\delta_{i,j}(\bar{\alpha}_{\text{min}} - \bar{\alpha}_{\text{max}}).

(16)

where subscript \( \text{min} \) denotes the category with the lower number of parents and subscript \( \text{max} \) denotes the category with the higher number of parents. (With random mating there is no covariance between different mates of \( i \); therefore, there is no mate term in the cross-product.) A numerical example is in appendix b.

**Step 3: correction of \( E(\Delta F) \) for deviations of \( V_s \) from Poisson variances:** The second term of Equation 1 is \( \frac{1}{2}I^T \text{N}_a \delta_\text{a} \), where \( \delta_\text{a} \) is a \( 2c_{\text{max}} \) vector of elements \( \delta_i = \alpha_i^2 V_{i,i,\text{dev}} \) and \( V_{i,i,\text{dev}} \) is a \( 2c_{\text{max}} \times 2c_{\text{max}} \) matrix with deviations from Poisson variances (Woolliams and Bijma 2000). Similar to the discrete generation case, \( V_{i,i} \) is approximated by a binomial variance. Elements of \( V_{i,i,\text{dev}} \) and a numerical example are in appendix b.

**Relation to Nomura (1996):** Nomura (1996) developed predictions for the special case of equal numbers of parents per sex selected from every age class (denoted \( n_m \) and \( n_f \)), i.e., for constant selection intensity with age. With those schemes every parent is selected in every category (except for categories with zero parents) and there are only two exclusive categories; i.e., males selected always and females selected always. In this respect, schemes with equal numbers of parents selected from every age class are like discrete generations, i.e., only two categories that do not compete for being selected. Bijma et al. (1999) show that Equations 30 and 31 of Nomura (1996) reduce to \( \Delta F = \frac{n_m}{2}\alpha_{\text{a}}^2 + \alpha_{\text{a}}^2 Q_{\text{a}} \frac{c_{\text{a}}}{L} + \alpha_{\text{a}}\frac{\delta}{\alpha_{\text{a}}^2 Q_{\text{a}} c_{\text{a}}} \), which is equivalent to the first term of Equation 1. This result is a rescaling of discrete generations, i.e., with discrete generation \( \alpha_{\text{a}} = 1/(2n_L) \), with overlapping generations and two exclusive categories, each contributing half, \( \alpha_{\text{a}} = 1/(2n_L) \), where \( L \) is the generation interval. Summation of contributions over the number of parents per generation shows that they sum to unity.

\[
\sum_{i=1}^{n_L} \frac{1}{(2N_m L)} + \sum_{i=1}^{n_f} \frac{1}{(2N_f L)} = 1.
\]

Furthermore, Nomura (1996) calculated \( Q \) using \( (1 - P)^{-1} \) [\( P \) is a gene flow matrix identifying the contribution of parental age groups to selected offspring multiplied by the proportion of genetic variance remaining after selection; Nomura (1996, Appendix)], which, for his special case, is equivalent to our \( (1 - G^{*\text{III}})^{-1} \) (see Woolliams et al. 1999, Equation 10). Analogous to Santiago and Caballero (1995), Nomura (1996) calculated \( Q \) omitting the \( (1 - 1/n_f) \). Contrary to Santiago and Caballero (1995) and to the present study, Nomura (1996) included a term \( C_{\text{a}} \) in the calculation of \( V_{i,i,\text{dev}} \) [first term in Equation 22 of Nomura (1996)]. Finally, Nomura (1996) considered only one generation inheritance of selective advantage when he calculated the total contribution of age classes. [See Equation 8 of Nomura (1996), which is equivalent to solving \( \alpha \) from \( N_{\alpha} = G^N_{\alpha} \) instead of using Equation 9 (Woolliams et al. 1999).]

**Stochastic simulation:** To examine the accuracy of the prediction equations, the breeding scheme described in the "population model" section was simulated and rates of inbreeding were calculated from simulated data. The simulation procedure is described in Bijma and Woolliams (1999). In the simulated data, an ancestor cohort \( t_0 \) and a descendant cohort \( t_1 \) were chosen (Bijma and Woolliams 1999). Inbreeding coefficients of individuals in cohorts \( t_0 \) and \( t_1 \) were calculated from the simulated pedigree, using the algorithm of Meuwissen and Luo (1992). Rates of inbreeding per year were calculated as \( \Delta F = \frac{1}{2} \left[ (1 - F_{t_0})/(1 - F_{t_0}) \right]^{t_1 - t_0} - 1 \), where \( F_{t_0} \) and \( F_{t_1} \) are the average inbreeding coefficients in cohorts \( t_0 \) and \( t_1 \), respectively. Rates of inbreeding per generation were calculated as \( \Delta F = L \Delta F_g \). Results were averaged over 500 replicates.

**RESULTS**

**Discrete generations:** For examination of the accuracy of predictions and to identify the origin of prediction errors, Table 1 shows simulated and predicted \( \Delta F \). Two types of predictions are in Table 1: \( \Delta F_{\text{pred}} \) is the prediction using \( \alpha \) and \( \beta \) estimated from simulation, and \( \Delta F_{\text{pred}} \) is the full deterministic prediction using \( \alpha \) and \( \beta \) from Equations 9 and 10. Differences between \( \Delta F_{\text{pred}} \) and \( \Delta F_{\text{pred}} \) reflect prediction errors originating from the prediction of \( \beta \) [in discrete generations, \( \alpha_{\text{a}} = 1/(2n_L) \) is known]. Differences between \( \Delta F_{\text{sim}} \) and \( \Delta F_{\text{pred}} \) reflect errors in Equation 1.

Generally, errors of the full prediction in Table 1 are small, most errors are below 5%, maximum errors are up to 8.1%, and trends agree well between simulations and predictions. Though errors are small, some trends can be observed. Most errors are positive and errors tend to be highest for \( N_m = 10 \), but errors tend to be
negative for \( n_o = 8 \) and \( N_m = 100 \). Prediction errors are partly due to errors in the prediction of \( \beta \); i.e., \( \Delta F_{\text{pred}}^* \) is generally more accurate than \( \Delta F_{\text{pred}} \). Because we have approximated the hypergeometric variance of family size by a binomial variance, positive errors for small numbers of parents were expected. The correction for hypergeometric variances becomes larger with fewer parents (Burrows 1984b), whereas a binomial correction is unaffected by the number of parents. Because the correction is a negative value, a binomial correction results in an overprediction for small numbers of parents. The current prediction was compared to the prediction of Santiago and Caballero (1995). As expected from the close agreement between equations of both methods, both methods gave very similar results (Bijma et al. 1999).

Figure 1 shows the relationship between \( \Delta F \) and heritability (\( h^2 \)), for \( N_m = N_f = 20 \) and for three selection intensities (\( n_o = 2, 8, \) or \( 32 \rightarrow m = 0, 1.271, \) or \( 1.967 \)). Though relationships of \( \Delta F \) with heritability and selection intensity can be inferred from other studies (e.g., Wright and Thompson 1990), they have never been explored in detail.

Figure 1 shows that \( \Delta F \) has a maximum for intermediate heritabilities (except for \( n_o = 2 \)), and changes in \( \Delta F \) are more pronounced with greater selection intensity. The maximum of \( \Delta F \) for intermediate \( h^2 \) is due to the Bulmer effect. When the Bulmer effect is ignored in Equation 7 (i.e., \( \kappa = 0 \)) the rate of inbreeding increases with \( h^2 \) over the whole range. The logic behind this is that with increasing \( h^2 \) the reduction of between-family variance increases, reducing the importance of the family component in the phenotype. [Note also that the intraclass correlation between full sibs \( (\rho = \frac{1}{2}h^2(1 - \kappa h^2)) \) has a maximum for \( h^2_{\text{max}} = 1/(2\kappa) \), which for a common value of \( \kappa = 0.8 \) equals \( h^2_{\text{max}} = 0.625 \).]

## Table 1

Rates of inbreeding from simulation (\( \Delta F_{\text{sim}} \)) and from prediction (\( \Delta F_{\text{pred}}^*, \Delta F_{\text{pred}} \)) for populations with discrete generations

<table>
<thead>
<tr>
<th>( N_m )</th>
<th>d</th>
<th>( n_o = 4 )</th>
<th>( n_o = 8 )</th>
<th>( n_o = 4 )</th>
<th>( n_o = 8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \Delta F_{\text{sim}} )</td>
<td>( \Delta F_{\text{pred}}^* )</td>
<td>( \Delta F_{\text{pred}} )</td>
<td>( \Delta F_{\text{sim}} )</td>
<td>( \Delta F_{\text{pred}}^* )</td>
</tr>
<tr>
<td>10</td>
<td>0.222</td>
<td>0.0225</td>
<td>0.0225</td>
<td>0.029</td>
<td>0.030</td>
</tr>
<tr>
<td>20</td>
<td>0.186</td>
<td>0.0195</td>
<td>0.0195</td>
<td>0.0226</td>
<td>0.0230</td>
</tr>
<tr>
<td>5</td>
<td>0.0166</td>
<td>0.0173</td>
<td>0.0174</td>
<td>0.0191</td>
<td>0.0196</td>
</tr>
<tr>
<td>40</td>
<td>0.0056</td>
<td>0.0057</td>
<td>0.0057</td>
<td>0.0077</td>
<td>0.0078</td>
</tr>
<tr>
<td>5</td>
<td>0.0047</td>
<td>0.0049</td>
<td>0.0049</td>
<td>0.0062</td>
<td>0.0061</td>
</tr>
<tr>
<td>100</td>
<td>0.0022</td>
<td>0.0023</td>
<td>0.0023</td>
<td>0.0031</td>
<td>0.0031</td>
</tr>
<tr>
<td>2</td>
<td>0.0019</td>
<td>0.0020</td>
<td>0.0020</td>
<td>0.0025</td>
<td>0.0025</td>
</tr>
<tr>
<td>5</td>
<td>0.0017</td>
<td>0.0018</td>
<td>0.0018</td>
<td>0.0022</td>
<td>0.0021</td>
</tr>
</tbody>
</table>

\( \Delta F_{\text{pred}}^* \), prediction using \( \alpha \) and \( \beta \) from simulation; \( \Delta F_{\text{pred}} \), full prediction; \( h^2 \), base generation heritability; \( N_m \), no. of selected sires; \( d \), mating ratio; \( n_o \), number of offspring per dam. Standard errors of simulation result were \(-1\%\) of the estimate.

### Figure 1

Relation of predicted (lines) and simulated (symbols) rates of inbreeding (\( \Delta F \)) with heritability (\( h^2 \)) for populations with discrete generations, with 20 sires and 20 dams and varying number of offspring per dam (\( n_o \)). \( n_o = 2: \square \), \( \Delta F_{\text{sim}} \); \( \triangle \), \( \Delta F_{\text{pred}} \); \( n_o = 8: \circ \), \( \Delta F_{\text{sim}} \); \( \cdot \), \( \Delta F_{\text{pred}} \); \( n_o = 32: \bigcirc \), \( \Delta F_{\text{sim}} \); \( \cdot \cdot \), \( \Delta F_{\text{pred}} \).
from Poisson variances. Without this correction, ΔF_{pred} is equal to a situation with h^2 = 0 and Poisson family size, resulting in ΔF_{pred} = 0.0125. The correction halves the prediction to 0.00625. This is an established result (Fairchild and Mackay 1996, p. 69). In the absence of variance of family size (which can only be achieved for d = 1), effective population size equals twice the actual population size: N_e = 2(N_m + N_f) and E(ΔF) = 1\sqrt{N_e} = 1/(4^{4}) = 0.00625.

With higher selection intensities (n_a = 8 or 32), ΔF increases considerably with heritability. For example, for h^2 = 0.6, ΔF increases by 54% compared to random selection (i.e., h^2 = 0) for n_a = 8, and by 105% for n_a = 32. The large increase of ΔF with selection intensity originates from the regression of the number of selected offspring on the breeding value of the parent, which is linear in i (λ = \frac{1}{2}h^2σ^2), giving a quadratic term in ΔF (Equation 7). Large values of λ indicate that the population descends for a large proportion from only a few ancestors.

For practical selection intensities (n_a = 2, 8), there is close agreement between ΔF_{pred} and ΔF_{sim}. For large selection intensities errors are larger (e.g., for n_a = 200, N_m = N_f = 40 and h^2 = 0.4, an error of ~18% was found). Large errors for extreme selection intensities do not undermine the general theory, i.e., Equation 1 is still valid, but the linear model (Equation 2) may be insufficient to predict expected genetic contributions (Woolliams and Bijma 2000).

**Overlapping generations**: Table 2 shows simulated and predicted rates of inbreeding per generation and generation intervals. Predictions of ΔF using α and β from simulation (such as ΔF_{pred} in Table 1) are not included, because standard errors on β were too large to draw conclusive inferences. Because the potential number of alternative schemes is very large with overlapping generations, a wide range of schemes was evaluated. Only schemes 1, 3, 5, 6, and 7 are within the scope of Nomura (1996). Schemes 1–5 represent a situation with two age classes with gradually increasing ages of females. Scheme 6 is similar to scheme 5 but with a mating ratio of two. Scheme 7 has equal numbers of parents in all categories. With schemes 8 and 9, parents were ranked on estimated breeding value [EBV_{j} = h^2(P_{ij} - P_{i})] across age classes and the highest ranking N_m males and N_f females were selected across age classes, which gives the highest genetic level of the offspring in the next cohort (James 1987). This strategy resulted in N = diag{19, 1, 38, 2} for n_a = 8 and N = diag{8, 2, 33, 7} for n_a = 4 (for both h^2 = 0.2 and 0.5). Furthermore, some arbitrary schemes with three and four age classes were evaluated to show that predictions are also accurate for more than two age classes. Prediction errors of ΔF were small, with most <5%. The maximum error was 6.6% and most errors were positive. Similar to the case with discrete generations, positive errors for small numbers of parents were expected due to the binomial approximation for the variance of family size.

**Generation intervals** are systematically underestimated in Table 2 (except for schemes with only one reproductive category per sex in which case L is fixed; schemes 1, 5, and 6). The underprediction is entirely explained by the way L_{sim} is calculated, i.e., L_{sim} = 1/n_{rep} \sum_{k=1}^{n_{rep}} L_{k}, where L_k = 1\sum_{i=1}^{n_a} n_{ai}; i.e., the generation interval is calculated per replicate as the time in which genetic contributions sum to unity and subsequently averaged over replicates (Bijma and Woolliams 1999). However, if α was averaged over replicates

### Table 2
Rates of inbreeding per generation from simulation (ΔF_{sim}) and from prediction (ΔF_{pred}) and generation intervals from simulation (L_{sim}) and prediction (L_{pred}) for populations with overlapping generations

<table>
<thead>
<tr>
<th>diag N</th>
<th>ΔF_{sim}</th>
<th>ΔF_{pred}</th>
<th>L_{sim}</th>
<th>L_{pred}</th>
<th>ΔF_{sim}</th>
<th>ΔF_{pred}</th>
<th>L_{sim}</th>
<th>L_{pred}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. {20, 0, 20, 0}</td>
<td>0.0150</td>
<td>0.0156</td>
<td>1.00</td>
<td>1.00</td>
<td>0.0169</td>
<td>0.0173</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2. {20, 0, 15, 5}</td>
<td>0.0178</td>
<td>0.0185</td>
<td>1.11</td>
<td>1.11</td>
<td>0.0200</td>
<td>0.0207</td>
<td>1.10</td>
<td>1.10</td>
</tr>
<tr>
<td>3. {20, 0, 10, 10}</td>
<td>0.0194</td>
<td>0.0206</td>
<td>1.23</td>
<td>1.22</td>
<td>0.0229</td>
<td>0.0235</td>
<td>1.20</td>
<td>1.18</td>
</tr>
<tr>
<td>4. {20, 0, 5, 15}</td>
<td>0.0158</td>
<td>0.0165</td>
<td>1.35</td>
<td>1.34</td>
<td>0.0205</td>
<td>0.0209</td>
<td>1.32</td>
<td>1.30</td>
</tr>
<tr>
<td>5. {20, 0, 20}</td>
<td>0.0097</td>
<td>0.0104</td>
<td>1.50</td>
<td>1.50</td>
<td>0.0112</td>
<td>0.0115</td>
<td>1.50</td>
<td>1.50</td>
</tr>
<tr>
<td>6. {20, 0, 40}</td>
<td>0.0080</td>
<td>0.0083</td>
<td>1.50</td>
<td>1.50</td>
<td>0.0091</td>
<td>0.0091</td>
<td>1.50</td>
<td>1.50</td>
</tr>
<tr>
<td>7. {20, 10, 10, 10}</td>
<td>0.0237</td>
<td>0.0247</td>
<td>1.48</td>
<td>1.43</td>
<td>0.0285</td>
<td>0.0290</td>
<td>1.41</td>
<td>1.36</td>
</tr>
<tr>
<td>8. {9, 1, 38, 2}</td>
<td>0.0132</td>
<td>0.0133</td>
<td>1.05</td>
<td>1.05</td>
<td>0.0152</td>
<td>0.0148</td>
<td>1.04</td>
<td>1.04</td>
</tr>
<tr>
<td>9. {8, 2, 33, 7}</td>
<td>0.0117</td>
<td>0.0123</td>
<td>1.14</td>
<td>1.13</td>
<td>0.0133</td>
<td>0.0137</td>
<td>1.13</td>
<td>1.13</td>
</tr>
<tr>
<td>10. {20, 0, 0, 10, 0, 30}</td>
<td>0.0103</td>
<td>0.0106</td>
<td>1.63</td>
<td>1.62</td>
<td>0.0163</td>
<td>0.0159</td>
<td>1.48</td>
<td>1.43</td>
</tr>
<tr>
<td>11. {20, 0, 0, 30, 0, 10}</td>
<td>0.0125</td>
<td>0.0130</td>
<td>1.17</td>
<td>1.16</td>
<td>0.0153</td>
<td>0.0150</td>
<td>1.10</td>
<td>1.10</td>
</tr>
<tr>
<td>12. {20, 5, 5, 0, 20, 10, 5, 5}</td>
<td>0.0217</td>
<td>0.0217</td>
<td>1.66</td>
<td>1.58</td>
<td>0.0267</td>
<td>0.0254</td>
<td>1.46</td>
<td>1.41</td>
</tr>
</tbody>
</table>

For n_a = 8, N is the distribution of parents over age classes and h^2 is the base generation heritability. Standard errors of ΔF_{sim} were ~1% of the estimate.

* n_a = 4 for this scheme.
and $L_{\alpha m}$ was calculated from the average, i.e., $L_{\alpha m} = \frac{1}{\sum_{i=1}^{n_{\alpha}} n_i \alpha_i}$, then $L_{\text{pred}}$ and $L_{\text{sim}}$ were in very close agreement (results not shown). This result was expected from the nonlinear relationship between $L$ and $\alpha$, so that $E[L \alpha]$ differs from $1/\sum_{i=1}^{n_{\alpha}} n_i E[\alpha_i]$.

Results from the current prediction were compared to results from the prediction of Nomura (1996) for the special case of equal numbers of parents selected from every age class. [A comparison was made for all schemes presented by Nomura (1996).] As expected from theory, results from both methods were similar (Bijma et al. 1999).

Relationship between $\Delta F$ and distribution of parents over age classes: Figure 2 shows the relationship between the rate of inbreeding (per year and per generation) and the proportion of parents selected from the second age class ($p_2$), for a population with two age classes, $N = N_t = 20$, $h_0 = 0.4$, and $n_0 = 10$. With the exception of $p_2 = 0, 0.5$, and 1.0, these schemes are beyond the scope of Nomura (1996). Generation interval was also included in Figure 2. On the horizontal axis, parents are shifted from all parents in the first age class ($p_2 = 0$, $N = \text{diag}(20, 0, 20, 0)$) to all parents in the second age class ($p_2 = 1$, $N = \text{diag}(0, 20, 20, 0)$). For $p_2 = 1$, no predictions are presented (i.e., no lines, only symbols), because in this scheme there are two distinct subpopulations that do not mix; i.e., individuals born in odd-numbered cohorts are one population and individuals born in even-numbered cohorts are the other population. This scheme violates the assumption of one random mating population in the derivation of $\Delta F_{\text{pred}}$. Therefore, the populations should be treated separately, which resulted in accurate predictions. Despite the complex relationship between $\Delta F$ and $p_2$ in Figure 2, where, for example, $\Delta F_i$ is nearly constant before declining sharply, accurate predictions were obtained throughout. The rate of inbreeding per year has a flat curve with a maximum for $p_2 = 0.5$, because the increase of $\Delta F_i$ with $p_2$ is counteracted by an increase in the generation interval, and as a result, $\Delta F_i = \Delta F_i / L$ shows only slight increase before $p_2 = 0.5$ and steep decrease after $p_2 = 0.5$.

For random selection, Hill (1979) showed that the rate of inbreeding in overlapping generations is related to the lifetime variance of family size and the number of parents entering the population per generation. The same pattern can be observed in Figure 2, which shows that $\Delta F_i$ has a maximum when parents are equally distributed over age classes, i.e., for $N = \text{diag}(0, 10, 10, 10)$ where the 10 parents selected in age class 1 the first year are the same as the 10 parents selected in age class 2 the next year. Thus only 10 distinct parents are selected from every cohort for this scheme, and with $L = 1.41$ the number of parents entering the population per generation equals only 14.1. For $N = \text{diag}(20, 0, 20, 0)$ 20 distinct parents are selected from every cohort and with $L = 1$, 20 parents enter the population per generation. The rate of inbreeding per generation reaches a minimum for $p_2 = 0.95$ ($N = \text{diag}(1, 19, 1, 19)$). At first glance, this result is counterintuitive; i.e., one might expect approximately equal rates of inbreeding per generation for $N = \text{diag}(19, 1, 19, 1)$ and for $N = \text{diag}(1, 19, 1, 19)$, however, for $N = \text{diag}(0, 19, 1, 19)$ 19 distinct individuals are selected from every cohort and, with $L = 1.90$, the number of parents per generation equals 36.1.

Line subdivision and migration: As mentioned earlier, the scheme with $N = \text{diag}(0, 20, 20, 0)$ has two nonmixing lines. Changing this scheme to $N = \text{diag}(0, 19, 1, 19)$ is equivalent to allowing some migration between both lines. Figure 3 shows a comparison between full line subdivision, line subdivision with migration, and one single line for schemes with 2 or 3 age classes. Note that the total number of parents per year is equal per comparison. The comparison shows that allowing some migration between lines substantially reduces $\Delta F_i$ (i.e., 0.0104 vs. 0.0141 and 0.0075 vs. 0.0141). The smallest $\Delta F$ is obtained when lines are joined together (40, 40) with a cohort interval of 2 years and (60, 60) with a cohort interval of 3 years. When comparing these rates of inbreeding, it must be realized, however, that the schemes with full line subdivision accumulate a between-line genetic variance equal to $2(1 - 1/\text{line}) \sigma_{\text{line}}^2$, where the $(1 - 1/\text{line})$ accounts for the fact that the mean is estimated from the sample; i.e., the variance is the observed variance in the sample ( Falconer and Mackay 1996, p. 265). The total genetic variance at time $t$, i.e., $\sigma_t^2 = \sigma_{\text{between}}^2 + \sigma_{\text{within}}^2$, equals $\sigma_t^2$ for $N = \text{diag}(0, 20, 20, 0)$ and $\sigma_t^2(1 + \frac{1}{2} F_i)$ for $N = \text{diag}(0, 0$, 0,
20, 0, 0, 20) and therefore the genetic variance is larger with full line subdivision.

Relationship between $\Delta F$ and heritability: Figure 4 shows the relationship between $h_0^2$ and $\Delta F$, for two breeding schemes. The first scheme ($S_1$) has most parents in the first age class, $N = \text{diag}(16, 4, 16, 4)$ whereas the second scheme ($S_2$) has most parents in the second age class, $N = \text{diag}(4, 16, 4, 16)$. With $S_1$, $\Delta F$ has a maximum for $h_0^2 = 0.5$–0.6, similar to the discrete generation case (see Figure 1). With $S_2$, however, $\Delta F$ increases with heritability over the whole range. The increase of $\Delta F$ with $h_0^2$ for $S_1$ is mainly due to an increased contribution of parents in age class 1 at high heritabilities. With high heritability, genetic gain is large, which gives offspring of 1-year-old parents an increased selective advantage. This increases the contribution of parents in age class 1 relative to the contribution of parents in age class 2. For example, with $S_2$ and $h_0^2 = 0.5$, expected genetic contributions of average parents are $\alpha_1^T = [0.027, 0.012, 0.027, 0.012]$, whereas for $h_0^2 = 0.9$, expected genetic contributions of average parents are $\alpha_1^T = [0.040, 0.011, 0.040, 0.011]$. This result shows that with increasing $h_0^2$ the genetic contributions become distributed more unequally over parents, resulting in a higher sum of squared contributions and therefore in an increased $\Delta F$. Furthermore, with $S_2$, $\beta$ increases with heritability, resulting in increased differences between genetic contributions of different parents selected from the same category, which further increases $\Delta F$.

Rates of inbreeding per year can be obtained from
Figure 4 as $\Delta F_T = \Delta F_i / L$, which shows the same trends with $h^2_i$ as $\Delta F_i$. In conclusion, results from Figure 4 show that in contrast to the case of discrete generations, no general pattern can be observed in the relationship between $\Delta F$ and $h^2_i$ with overlapping generations.

**DISCUSSION**

Explicit prediction equations for rates of inbreeding in populations with either discrete or overlapping generations under mass selection were developed, on the basis of the approach of Woolliams and Bijma (2000) and Woolliams et al. (1999). Except for extreme selection intensities in females, predictions were accurate for discrete as well as for overlapping generations. Though based on a different approach, the current method extends the method of Nomura (1996) to populations with overlapping generations and an arbitrary distribution of parents across age classes, removing the stringent restriction of Nomura (1996). Relationships between rates of inbreeding and genetic and population parameters were also presented. General relationships apparent in discrete generations could not be extended to overlapping generations. For the prediction of rates of inbreeding in overlapping generations it is crucial to account for the inheritance of selective advantage both between and within categories. For discrete generations with only two categories (males and females), which do not compete for selection, only competition between selection candidates within categories is relevant.

The current method was compared to methods based upon the proportion of genetic variance transmitted to the offspring, which showed that with random mating, the equations of both Santiago and Caballero (1995) and Nomura (1996) can be reduced to simple expressions in terms of expected genetic contributions. Santiago and Caballero (1995) suggested that the differences between their results using the reduced genetic variance and those of Woolliams et al. (1993) using long-term contributions were due to the difference in approach. The present results show that the differences obtained previously were most likely due to errors in the derivations involving complex pathways over multiple generations that were needed by Woolliams et al. (1993). These complexities were avoided by Santiago and Caballero (1995). However, Woolliams and Bijma (2000) were able to derive the present results using long-term contributions by modeling the transfer of selective advantages in a single generation by assuming an equilibrium. The idea of basing the prediction on Bulmer’s equilibrium variances was introduced by Santiago and Caballero (1995). However, their approach to modeling the inheritance of selected advantage by the proportion of genetic variance retained is correct only for mass selection [see Woolliams et al. (1999) for a general approach].

Prediction errors became large when the number of selection candidates per dam became extremely large (Figure 2), but these situations are out of the range of most artificial selection programs. Certain species (e.g., fish or chicken) are able to produce many offspring per dam, but the number of selection candidates per dam is generally lower. High selection intensities in males can easily be obtained with a limited number of selection candidates per dam when the mating ratio is large. For these situations predictions were accurate (see Table 1, schemes with $d = 5$, $n_0 = 8$ – $l = 2.063$). The errors with large $n_0$ were not present for low $h^2$ (results not shown), which indicates that the current method is also applicable to species with a large number of offspring when natural directional selection acts on a trait with low heritability.

In this article, equations for predicting rates of inbreeding were developed assuming a model of truncation selection on a normally distributed trait controlled by an infinitesimal model of gene effects. The predicted rate of inbreeding relates to homozygosity (by descent) at a neutral locus, unlinked to genes affecting the trait under selection (Woolliams and Bijma 2000). When the infinitesimal model does not hold, and the number of genes affecting the trait is large, or when the number of chromosomes is small, it is questionable whether neutral and unlinked loci exist at all. When loci are nonneutral, or linked to nonneutral loci, predicted rates of inbreeding cannot be related directly to the homozygosity at the locus, because a covariance between the genetic contribution and the gene frequency will arise due to selection (Woolliams and Bijma 1999). However, the rate of inbreeding can still be related to rates of inbreeding obtained by analyzing pedigrees using Wright’s (1922) path coefficient method, or Malecot’s (1948) coefficient of kinship, and also to estimates of inbreeding depression based on inbreeding levels calculated from the pedigree. Recently, Santiago and Caballero (1998) extended prediction methods for effective population size to populations with linked loci undergoing mass selection but for discrete generations only.

In general, to obtain accurate predictions of $\Delta F$ one needs to account for more than one generation of inheritance of selective advantage between categories. It was sufficient for Nomura (1996) to account for only a single generation because of the special case of equal numbers of parents per age class. In that case, shifting contributions between age classes has only a minor effect on $\Delta F$ because the contributions will remain with the same individuals with the same relative fitness, because every individual is selected in every category. Therefore the lifetime contribution will not be affected. For schemes where the number of parents differs between age classes, shifting of contributions between categories means shifting to other individuals (at least partly), which will affect the lifetime contribution. Consider, for example, scheme 10 in Table 2 with $h^2_i = 0.5$: 

$\Delta F_T = \Delta F_i / L$
accounting for only one generation of inheritance (i.e., calculating $\alpha$ from $N\alpha = G^2 N\alpha$; Woolliams et al. 1999) gives $\Delta F_{\text{pred}} = 0.0128$, an error of $-21$%; whereas using Equation 9 gives $\Delta F_{\text{pred}} = 0.0159$, an error of only $-2$%.

The use of the concept of long-term genetic contributions to predict rates of inbreeding has several appealing properties. First, the derivation of the relationship between rates of inbreeding and genetic contributions is based directly on the probability of alleles being identical by descent, which enhances the intuitive understanding (Woolliams and Bijma 2000). Furthermore, rates of genetic gain can easily be obtained from the covariance between the genetic contribution and the Mendelian sampling component of the breeding value (Woolliams and Thompson 1994; Woolliams et al. 1999), which integrates methods for predicting genetic gain and rates of inbreeding. Finally, the prediction procedure for genetic contributions describes mechanisms determining the impact of current individuals on future populations and the turnover rate of genes and gives therefore an understanding of the mechanisms determining the development of the pedigree (Woolliams et al. 1999; Bijma and Woolliams 1999). Because the approach is general, it is clear how prediction equations can be extended to other situations.

With a fixed total number of parents selected per year, populations showed maximum rates of inbreeding (per year and per generation) when the number of parents entering the populations per generation was least, which occurred with an equal number of parents in every age class. Rates of inbreeding were smallest when most parents were in the older age classes, because those schemes had the largest number of parents entering the population per generation. This result broadly resembles the results of Hill (1974) for random selection in overlapping generations, although selected populations have an additional component of inbreeding arising from the expected differential contributions within age classes, which will modify this relationship. Schemes with most parents in the later age classes resembled population subdivision with some migration between lines. Because the selective advantage of categories depends on heritability, genetic contributions of categories are strongly affected when heritability changes (Bijma and Woolliams 1999); i.e., contributions generally shifted to the younger age classes when heritability increased. Therefore, no general relationship between heritability and rate of inbreeding could be observed with overlapping generations.

In this article, equations were developed to predict rates of inbreeding for diploid populations with two sexes under controlled selection. The results are therefore primarily relevant for populations under artificial selection, for example, in animal breeding or in selection experiments. Though this article focuses on mass selection within age classes, results for mass selection across age classes can easily be accomplished by choosing the appropriate $N$, as in schemes 8 and 9 in Table 2. An extension to a situation where individuals in older age classes have more information, e.g., progeny information, only requires the calculation of probabilities of selecting the same individual on different ages, which can be done using standard index theory. The method can also be extended to other selection strategies and modes of inheritance (e.g., index selection and imprinting), using the key results of Woolliams and Bijma (2000) and Woolliams et al. (1999).

In animal breeding, optimization of breeding programs has focused for a long time on the maximization of genetic gain for the short term, partly because methods to predict long-term response were not available. When rates of inbreeding in selected populations can be predicted, predictions of long-term response under the infinitesimal model become available. This article enables methods for the optimization of breeding schemes on the long term (e.g., Villanueva et al. 1996; Villanueva and Woolliams 1997) to be extended to populations with overlapping generations and mass selection.

J.A.W. gratefully acknowledges the Ministry of Agriculture, Fisheries and Food (United Kingdom) for financial support. Ab Groen is acknowledged for giving useful comments on the manuscript, and Tetsuro Nomura for generously sending us a copy of his programs. This research was financially supported by the Netherlands Technology Foundation (STW) and was coordinated by the Earth and Life Science Foundation (ALW).

**LITERATURE CITED**


Derivation of Equation 8: With a binomial distribution of family size, the deviation from a Poisson variance equals \( np(1 - p) - np = -np^2 \), where \( n \) is the number of candidates (\( \gamma_n \) for sires and \( \gamma_n \) for dams) and \( p \) is the selected proportion \( (1 / (2 \gamma_n d) \) for male offspring and \( 1 / (2 \gamma_n o) \) for female offspring). Elements of \( \mathbf{V}_{n}^{\text{dir}} \) are therefore \( V_{n_{i o}} = \left[ -1 / (2 \gamma_n d), 0, 0, -d / (2 \gamma_n o) \right] \) and \( V_{n_{i d}} = \left[ -1 / (2 \gamma_n d), 0, 0, -d / (2 \gamma_n o) \right] \). From \( \delta_t = \alpha^2 \mathbf{V}_{t}^{\text{dir}} \alpha \) it follows that \( \delta_m = -1 / (2 \gamma_n d) (4 N_h \gamma_n) \) and \( \delta_l = -1 / (2 \gamma_n d) (4 N_l \gamma_n) - 1 / (2 \gamma_n o) \). The total correction equals \( \gamma N_m \delta_m + \gamma N_o \delta_l = -1 / (8 N_l N_o) - 1 / (8 N_o N_o) = -1 / (8 T) \), where \( T = 1 / 2 \gamma_n \).

Example: For \( N_m = 20, N_o = 60, n_o = 8, \) and \( h^2 > 0.4 \), selected proportions, selection intensities, and variance reduction coefficients are \( p_m = 0.083, p_o = 0.250, l_m = 1.839, l_r = 1.271, i = 1.555, k_m = 0.839, k_r = 0.759, \) \( \kappa = 0.799 \). For \( \mu \) = 71971) equilibrium genetic variance and heritability are \( \sigma^2 = 0.314, h^2 = 0.343 \). From Equation 6 \( \Delta F \) for a Poisson variance of family size equals \( \Delta F_{\text{Poisson}} = 0.00625 + 0.00208 + 0.1270(0.02255 + 0.01212) = 0.01276 \). From Equation 8, the correction equals \(-1 / 1920 = -0.00052 \) and the final prediction is \( E[\Delta F] = 0.0122 \).

**APPENDIX B: OVERLAPPING GENERATIONS**

**Corrections for deviations of \( V_{n}^{\text{dir}} \) from Poisson variances:** From Equation 1, the correction equals \( \frac{1}{2} \mathbf{1}^T \mathbf{N} \delta \), where \( \delta \) is a \( 2c_{\text{max}} \) vector of elements \( \delta_t = \alpha^T \mathbf{V}_{t}^{\text{dir}} T \alpha \), and where \( \mathbf{V}_{n}^{\text{dir}} \) is a \( 2c_{\text{max}} \times 2c_{\text{max}} \) matrix with deviations of \( \mathbf{V}_{n} \) from Poisson variances (Woolliams and Bijma 2000). Similar to the case of discrete generations, deviations from Poisson variances are \(-np^2 \), where \( n \) is the number of candidates (\( \gamma_n \) for sires and \( \gamma_n \) for dams) and \( p \) is the selected proportion. The selected proportion in subclass \( kl \), i.e., among offspring in category \( k \) descending from parents in category \( l \), equals \( p_k = p_k / g_k \), where \( p_k \) is the selected proportion in category \( k \) (\( p_k = n_k / T \)), and \( g_k \) and \( g_k \) are elements of the modified gene flow matrix \( \mathbf{G} \) and the conventional gene flow matrix \( \mathbf{G}_o \), respectively. The element \( g_k \) represents the proportion of selected offspring in category \( k \) descending from parents in category \( l \), and \( g_k \) represents the proportion of candidates for selection in category \( k \) descending from parents in category \( l \) (Woolliams et al. 1999).

Therefore, \( \mathbf{V}_{n}^{\text{dir}}(k, k) \) equals \( -\gamma_n d p_k g_k / [g_k] \) when the parent is a male, and \(-\gamma_n d p_k g_k / [g_k] \) when the parent is a female. Off-diagonal elements of \( \mathbf{V}_{n}^{\text{dir}} \) are zero with binomial family size.

Example: For \( N = \text{diag}(2, 8, 15, 25), h_0 = 0.4, \) and \( n_o = 4, \) selected proportions, selection intensities, and variance reduction coefficients are \( p = (0.1500 0.1000 0.1875 0.3125), i = (1.5544 1.7546 1.4357 1.1331), \) \( \kappa = (0.8051 0.8297 0.7877 0.7306) \). Predicted \( \alpha, \beta, \) generation interval and Bulmer’s (1971) equilibrium genetic

**APPENDIX A: DISCRETE GENERATIONS**

**Derivation of Equation 6:** Starting from Equations 3 and 4, and substituting \( \alpha^2 = 1 / 4 N_s^2, \beta^2 = \alpha^2 (1 - \pi)^2 \), \( \lambda = \gamma^2 / \sigma^2, \pi = (1 - k h)^2 \), and \( E[A_{s, o} - \bar{A}^2] = (1 - 1 / N_s) \sigma^2 (1 - k h)^2 \), it follows that

\[
E[u_{m}^{2}] = \frac{1}{4N_m^2} + \frac{1}{4N_m^2} \frac{i^2}{\sigma^2} \frac{(1 - 1/N_m) \sigma^2 (1 - k h)^2}{(1 + k h)^2} \\
+ \frac{d}{4N_m^2} \frac{i^2}{\sigma^2} \frac{(1 - 1/N_m) \sigma^2 (1 - k h)^2}{(1 + k h)^2},
\]

and

\[
E[u_{i}^{2}] = \frac{1}{4N_i^2} + \frac{1}{4N_i^2} \frac{i^2}{\sigma^2} \frac{(1 - 1/N_i) \sigma^2 (1 - k h)^2}{(1 + k h)^2} \\
+ \frac{d^2}{4N_m^2} \frac{i^2}{\sigma^2} \frac{(1 - 1/N_m) \sigma^2 (1 - k h)^2}{(1 + k h)^2}.
\]

Substituting those expressions into \( E[\Delta F] = \frac{1}{2} N_m E[u_{m}^{2}] + \frac{1}{2} N_o E[u_{i}^{2}] \) and using \( \alpha^2 / \sigma^2 = h^2 \) and \( d = N_o / N_m \) gives Equation 6.
variance and heritability are (see Bijma and Woolliams 1999 for an example of the prediction of $\alpha$ and $\beta$)

\[ \alpha^2 = (0.01974, 0.01454, 0.01171, 0.00710), \beta^2 = (0.02228, 0.01829, 0.01251, 0.00904), L = 1.416, \sigma^2_3 = 0.3355, h^2 = 0.3586. \]

The conventional and modified gene flow matrices are (Bijma and Woolliams 1999)

\[
\begin{bmatrix}
0.3 & 0.2 & 0.1875 & 0.3125 \\
1 & 0 & 0 & 0 \\
0.3 & 0.2 & 0.1875 & 0.3125 \\
0 & 0 & 1 & 0
\end{bmatrix}
\]

\[
\begin{bmatrix}
0.3245 & 0.1755 & 0.2291 & 0.2709 \\
0.3276 & 0.1724 & 0.2347 & 0.2653 \\
0.3227 & 0.1773 & 0.2258 & 0.2742 \\
0.3180 & 0.1820 & 0.2175 & 0.2825
\end{bmatrix}
\]

For $N = \text{diag}(12, 8, 15, 25)$ there are four exclusive categories: (1) males selected both at 1 and 2 years of age (i.e., the eight highest-ranking males), for which $E(u^2_{i,1}) = E((u_{i,1} + u_{i,2})^2)$; (2) males selected only at 1 year of age (i.e., males ranking 9–12) for which $E(u^2_{i,1-2}) = E(u^2_{i,1});$ (3) females selected both at 1 and 2 years of age (i.e., the 15 highest-ranking females), for which $E(u^2_{i,2}) = E((u_{i,3} + u_{i,4})^2)$; and (4) females selected only at 2 years of age (i.e., females ranking 16 through 25), for which $E(u^2_{i,2}) = E(u^2_{i,4})$. Summation of expectations of squares and cross-products over categories, $s$, gives, for males, $\Sigma^2_{i=1} n_i E(u^2_{i,j}) = 12 E(u^2_{i,1}) + 8E(u^2_{i,1-2}) + 16E(u_{i,1}u_{i,2})$; and for females; $\Sigma^2_{i=1} n_i E(u^2_{i,j}) = 15E(u^2_{i,3}) + 25E(u_{i,3}^2) + 30E(u_{i,3}u_{i,4})$ (see also Equations 11 and 12).

From Equation 14,

\[
E(u^2_{i,j}) = \alpha^2 + (1 - \frac{1}{n_i}) \beta^2 \sigma^2 (1 - \kappa h^2)
\]

\[ + \sigma^2_3 (1 - \frac{1}{n_i}) \beta^2 \sigma^2 (1 - \kappa h^2) \]

\[ = 0.0003897 + 0.0001086 + 0.0000606 = 0.0005589. \]

(Bars with subscript $f$ denote averages over female categories weighted by the number of dams in the categories, e.g., $\bar{u}^2_f \left( 15 \times 0.01171^2 + 25 \times 0.00710^2 \right) / 40 = 0.829 \times 10^{-4}$. Similarly, for females $E(u^2_{i,j})$ is calculated from Equation 15. From Equation 16, $E(u^2_{i,j}u_{i,k}) = \alpha^2 \alpha^2_3 + (1 - \frac{1}{n_i}) \beta^2 \sigma^2_3 (1 - \kappa h^2) + \alpha^2 \beta^2_3 E(\bar{A} - \bar{X}) = 0.0000831 + 0.0000254 + 0.0000111 = 0.0001196. $)

Using Equation 13 (note the reordering) with $N_c = \text{diag}(2, 8, 25, 15)$ and

\[
U = E
\begin{bmatrix}
\begin{array}{cccc}
0 & 0 & 0 \\
2u_i & u_j & 0 & 0 \\
0 & 0 & u_i & 0 \\
0 & 0 & 2u_i & u_j
\end{array}
\end{bmatrix}
\]

\[ = 10^{-3}
\begin{bmatrix}
0.5589 & 0 & 0 & 0 \\
0.7868 & 0.3408 & 0 & 0 \\
0 & 0 & 0.0946 & 0 \\
0 & 0 & 0.2393 & 0.1972
\end{bmatrix}
\]

the rate of inbreeding with Poisson family size is $E(\Delta F) = \frac{1}{2} N N \cdot U \cdot 1 = 0.0123$. Instead of using the matrix form of Equation 13, one can also use $E(\Delta F) = \frac{1}{2} \Sigma_{i=1}^{n_{max}} n_i E(u^2_{i,j}) + \frac{1}{2} \Sigma_{i=1}^{n_{max}} n_i E(u^2_{i,j})$ with the summations calculated as above.

The correction for deviations of $V_{n(1)}$ from Poisson variances, for sires in age class one to selected male offspring in age class two, is $V_{n(1), \text{dev}}(2, 2) = -\frac{1}{2} n_i d_{ip} g_{i,p} / (g_{i,p})^2 = -0.0477$. The full matrix for sires in age class one equals

\[
V_{n(1), \text{dev}} = \begin{bmatrix}
-0.1053 & 0 & 0 & 0 \\
0 & -0.0477 & 0 & 0 \\
0 & 0 & -0.1627 & 0 \\
0 & 0 & 0 & -0.4389
\end{bmatrix}
\]

The matrices for other age classes are $V_{n(2), \text{dev}} = \text{diag}(-0.0692, -0.0297, -0.1105, -0.3235)$, $V_{n(3), \text{dev}} = \text{diag}(-0.0672, -0.0313, -0.1020, -0.2629)$, $V_{n(4), \text{dev}} = \text{diag}(-0.0338, -0.0144, -0.0541, -0.15957)$, $\delta^2 = (-9.554 - 6.474 - 6.004 - 3.168) \times 10^{-5}$, e.g., $\delta^2 = \alpha^2 \cdot N \cdot \Delta F = -9.554 \times 10^{-5}$. The correction factor is $\frac{1}{2} N N \alpha = 0.0004$, resulting in $\Delta F = 0.0123 - 0.0004 = 0.0119$ and $\Delta F = L \Delta F = 0.0168$. 