



**HAL**  
open science

## **A dynamic deterministic model to optimize a multiple-trait selection scheme**

Anne Devalle Costard, Zulma Vitezica, Carole Moreno-Romieux, Jean Michel J. M. Elsen

### ► **To cite this version:**

Anne Devalle Costard, Zulma Vitezica, Carole Moreno-Romieux, Jean Michel J. M. Elsen. A dynamic deterministic model to optimize a multiple-trait selection scheme. *Journal of Animal Science*, 2009, 87 (3), pp.885-894. <10.2527/jas.2008-0898>. <hal-02665246>

**HAL Id: hal-02665246**

**<https://hal.inrae.fr/hal-02665246v1>**

Submitted on 31 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# JOURNAL OF ANIMAL SCIENCE

*The Premier Journal and Leading Source of New Knowledge and Perspective in Animal Science*

## **A dynamic deterministic model to optimize a multiple-trait selection scheme**

A. D. Costard, Z. G. Vitezica, C. R. Moreno and J.-M. Elsen

*J ANIM SCI* 2009, 87:885-894.

doi: 10.2527/jas.2008-0898 originally published online November 21, 2008

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/87/3/885>



**American Society of Animal Science**

[www.asas.org](http://www.asas.org)

# A dynamic deterministic model to optimize a multiple-trait selection scheme<sup>1</sup>

A. D. Costard,<sup>\*2</sup> Z. G. Vitezica,<sup>†‡</sup> C. R. Moreno,<sup>\*</sup> and J.-M. Elsen<sup>\*</sup>

<sup>\*</sup>l'Institut National de la Recherche Agronomique, SAGA station d'amélioration génétique animale, 31326 Castanet-Tolosan, France; <sup>†</sup>Université de Toulouse; Institut National Polytechnique de Toulouse, Ecole Nationale Vétérinaire de Toulouse; Unité Mixte de Recherche, 1289 Tandem, Tissus Animaux, Nutrition, Digestion, Ecosystème et Métabolisme; Ecole Nationale Supérieure Agronomique de Toulouse, F-31326 Castanet-Tolosan Cedex, France; and <sup>‡</sup>l'Institut National de la Recherche Agronomique, Unité Mixte de Recherche, 1289 Tandem, Tissus Animaux, Nutrition, Digestion, Ecosystème et Métabolisme; Chemin de Borde-Rouge, Auzeville, F-31326 Castanet-Tolosan, France

**ABSTRACT:** A mathematical approach was developed to model and optimize simultaneous selection on 2 traits, a quantitative trait with underlying polygenic variation and a monogenic trait (e.g., resistance to a disease). A deterministic model allows global optimization of the selection scheme to maximize the frequency of the desired genotype for the monogenic trait, while minimizing the loss of genetic progress on the polygenic trait. An additive QTL or gene was considered. Breeding programs with overlapping generations, different se-

lection strategies for males and females, and assortative mating were modeled. A genetic algorithm was used to solve this optimization problem. This modeling approach may easily be adapted to a variety of underlying genetic models and selection schemes. This model was applied to an example where selection on the Prp gene for scrapie resistance was introduced as an additional selection criterion in an already existing dairy sheep selection scheme.

**Key words:** genetic algorithm, global optimization, marker-assisted selection, selection scheme

©2009 American Society of Animal Science. All rights reserved.

J. Anim. Sci. 2009. 87:885–894  
doi:10.2527/jas.2008-0898

## INTRODUCTION

Gene or QTL information could improve selection efficiency and allow the implementation of marker-assisted selection. Marker-assisted selection has been proposed to improve classical selection schemes for complex traits with low heritability or that are difficult or expensive to measure (Lande and Thompson, 1990; Meuwissen and van Arendonk, 1992; Dekkers and van Arendonk, 1998). Strategies to use QTL or gene information in marker-assisted selection are mainly based on indices combining the EBV for QTL and the EBV for the residual polygenic effect. Marker-assisted selection increases the QTL frequency in the short-term, but can result in less response in the long-term (Lande and Thompson, 1990). Several studies (Gibson, 1994;

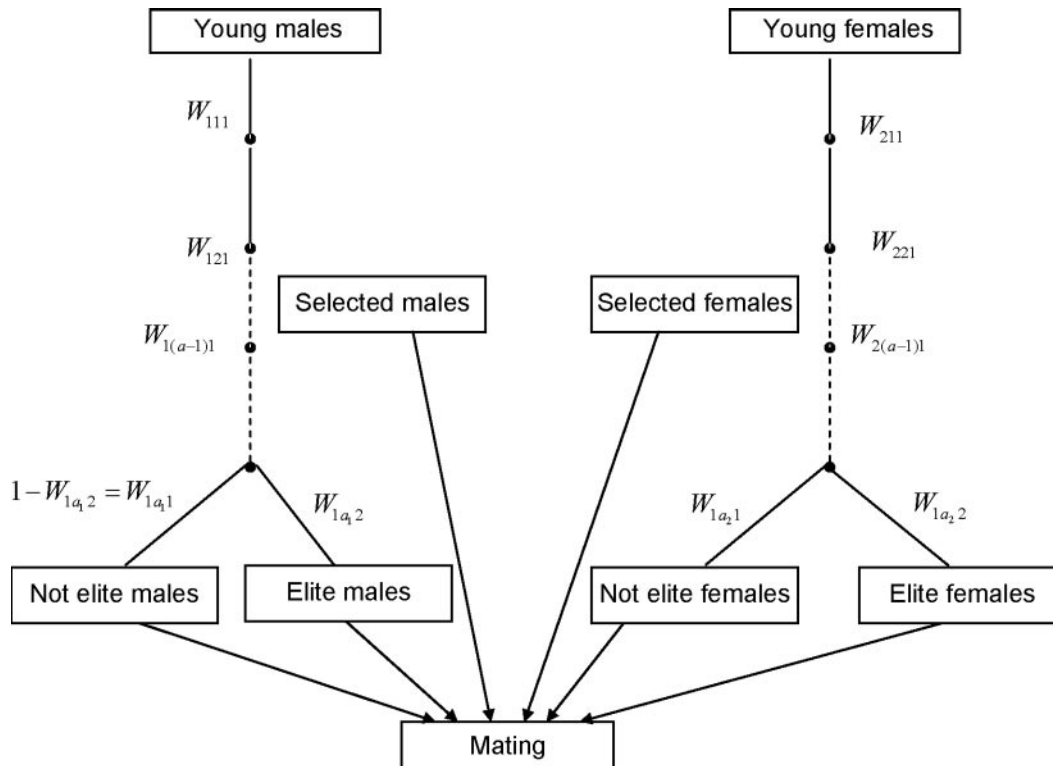
Larzul et al., 1997b; Pong-Wong and Woolliams, 1998) have explored the interactions between selection methods and time horizon. Dekkers and van Arendonk (1998) optimized selection on an identified QTL in a simple selection scheme (phenotypic selection, random mating) using an optimal control approach (Lewis, 1986). The method was extended by Chakraborty et al. (2002). The benefit of optimal selection on a single identified QTL and the differences between optimal and standard QTL selection were evaluated by Dekkers and Chakraborty (2001). Manfredi et al. (1998) optimized selection and mating on a QTL for a sex-limited trait using sequential quadratic programming. In these studies, selection was for a single quantitative trait affected by 1 identified QTL of meaningful size as well as polygenes. The objective of this study is to optimize simultaneous selection on 2 traits, a monogenic trait controlled by an identified QTL or gene and a second independent polygenic trait that is unaffected by the major locus with a genetic algorithm (Holland, 1973; Goldberg, 1989; Mongeau, 2003). A deterministic and dynamic model is presented for breeding programs including overlapping generations, different selection strategies for males and females, and nonrandom mating.

<sup>1</sup>This work was supported by the European Union project QLRT-2000-01733 "Scrapiefreesheep" and by the EU Network of Excellence, Neuroprion.

<sup>2</sup>Corresponding author: Anne.Devalle@toulouse.inra.fr

Received January 23, 2008.

Accepted November 18, 2008.



**Figure 1.** Selection scheme description with individual classes (rectangles), selection steps (●), and selection proportion  $W_{sac}$  (sex  $s = 1, 2$ ; age  $a = 1, \dots, a_s$ ; category  $c = 1, 2$ ).

## MATERIALS AND METHODS

Animal Care and Use Committee approval was not obtained for this study because no animals were used.

### Genetic Model

A population of infinite size with overlapping generations and different classes of animals defined by sex  $s$  ( $s = 1$  for males and  $s = 2$  for females), age  $a$ , and EBV categories  $c$  (elite, not elite) was considered. Animals of identical sex and age may be classified in different subsets according to their EBV mean and their use as reproducers (e.g., the number of progeny per year). The QTL or gene has 2 alleles (B and b). Genotypes BB, Bb/bB, and bb are denoted as  $g1$ ,  $g2$ , and  $g3$ , respectively. The polygenic effects are assumed to follow the infinitesimal genetic model with variance  $\sigma^2$ . The identified QTL or gene is unlinked with genes that influence the quantitative trait. For generation  $t$ , each class of animals has a QTL or gene effect with a genotype frequency of  $f_{sactg}$ , and a polygenic effect denoted by the within class average EBV  $\mu_{sactg}$ . Selection can involve the EBV of the quantitative trait, the identified QTL or gene, or both of them. Random mating and assortative mating according to genotype or to EBV categories can be used. The deterministic and dynamic model includes different equations that correspond to a general scheme presented in Figure 1. The phenomena underlying the evolution of the population are aging,

selection, and matings. The equations describing the scheme are as follows:

Without selection, after year  $t = 1$ , the genotype frequency and the average EBV for an animal of sex  $s$ , category  $c$ , and genotype  $g$  do not change between  $t$  and  $t-1$ .

$$f_{sactg} = f_{s(a-1)c(t-1)g}$$

$$\mu_{sactg} = \mu_{s(a-1)c(t-1)g}$$

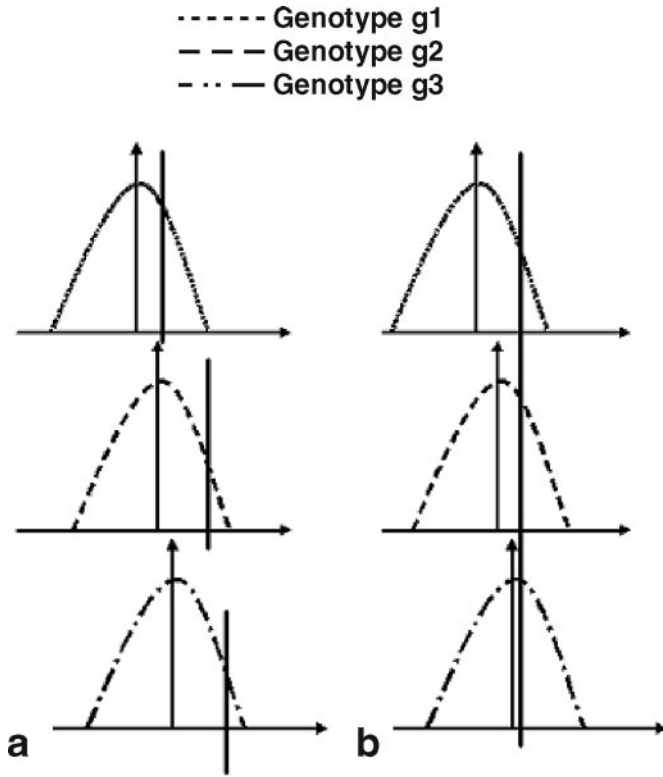
**Selection.** Different selection steps after year  $t = 1$  can be considered for males and females (Figure 1). For the last selection step, males and females are distributed in 2 categories, elite and not elite. The selection rate  $W_{sac}$  can vary at each step for males and females. The expression for  $W_{sac}$  is

$$W_{sac} = \sum_g w_{sactg} f_{s(a-1)c(t-1)g},$$

where  $w_{sactg}$  is the proportion of individuals with sex  $s$ , age  $a$ , and category  $c$  at generation  $t$  with genotype  $g$  that are selected to become breeding animals. At each step, candidates for selection can be evaluated through their phenotype, genotype, or both.

Assuming that EBV are normally distributed,  $w_{sactg}$  is equal to

$$w_{sactg} = \phi(\beta_{2g} - \mu_{s(a-1)c(t-1)g}) - \phi(\beta_{1g} - \mu_{s(a-1)c(t-1)g}),$$



**Figure 2.** Selection threshold for genotype g1 (· · ·), g2 (— — —), and g3 (—●—) under combined genotypic and phenotypic selection (a) and phenotypic selection (b).

where  $\phi$  is the normal cumulative distribution function for variance  $\sigma^2$ , and  $\beta_{1g}$  and  $\beta_{2g}$  are lower and upper thresholds for genotype  $g$  ( $\beta_{1g}$  may be  $-\infty$  and  $\beta_{2g}$  may be  $+\infty$ ) defining the interval selected.

When individuals are selected according to EBV and genotype information, the selection thresholds may vary according to the genotype (Figure 2a), and the state variables  $f_{sactg}$  and  $\mu_{sactg}$  are given by

$$f_{sactg} = \frac{w_{sactg}}{W_{sac}} f_{s(a-1)c(t-1)g}; \quad [1]$$

$$\mu_{sactg} = \mu_{s(a-1)c(t-1)g} + \frac{[\varphi(\beta_{1g} - \mu_{s(a-1)c(t-1)g}) - \varphi(\beta_{2g} - \mu_{s(a-1)c(t-1)g})]}{w_{sactg}} \rho\sigma, \quad [2]$$

where  $\varphi(x)$  is the normal density function,  $\rho$  is the accuracy or the correlation between true and polygenic EBV, and  $\sigma$  the genetic SD. Phenotypic selection is a special case of these equations, where  $\beta_1$  and  $\beta_2$  are the same for all genotypes (i.e.,  $\beta_{11} = \beta_{12} = \beta_{13} = \beta_1$  and  $\beta_{21} = \beta_{22} = \beta_{23} = \beta_2$ ; Figure 2b).

In the case of selection for a monogenic trait, the genotypes are ranked because some genotypes are better than others from a biological point of view (e.g., favorable or unfavorable genotypes for a disease). Let  $g[k]$  be the  $k$ th ranked genotype. These constraints must be verified:

$$\sum_{k=1}^{r-1} f_{s(a-1)c(t-1)g[k]} < W_{sac} \quad \text{and}$$

$$\sum_{k=1}^r f_{s(a-1)c(t-1)g[k]} \geq W_{sac}$$

because there cannot be more individuals in those classes than in the whole population.

The selection proportions are

$$w_{sactg[k]} = 1 \text{ if } k \leq r - 1;$$

$$w_{sactg[k]} = 0 \text{ if } k > r; \text{ and}$$

$$w_{sactg[r]} = \frac{W_{sac} - \sum_{k=1}^{r-1} f_{s(a-1)c(t-1)g[k]}}{f_{s(a-1)c(t-1)g[r]}} \text{ otherwise.}$$

Selection within the  $r$ th genotype was random. The genotype frequency  $f_{sactg}$  and the EBV mean  $\mu_{sactg}$  of the population in generation  $t$  are obtained from Eq. [1] and [2].

**Mating.** Mating between different classes (selected, elite, and not elite) of females and males can be considered (Figure 1). The probability for a progeny to be of genotype  $g$ , given that its parents carry the genotypes  $h$  (sire) and  $k$  (dam), is  $\tau_{ghk}$ . The probability that a parent of sex  $s$  has genotype  $g$  is  $F_{s(t-1)g}$  and is defined as

$$F_{s(t-1)g} = \sum_{a,c} \delta_{sac} f_{sac(t-1)g},$$

where  $\delta_{sac}$  is the proportion of individuals of age  $a$  and class  $c$  among the parents with sex  $s$ . This quantity depends on demographic constraints (e.g., reproductive age) and breeder decisions (e.g., relative use of progeny tested males). From the parameters  $\tau_{ghk}$  and  $F_{s(t-1)g}$ , different matings (random or assortative) can be considered.

For random mating, none of the characteristics of the parents (e.g., EBV or genotype) are considered in matings. The genotype frequency and the EBV mean for the newborn animals of age 1 and category 1 are

$$f_{s11tg} = \sum_{h,k} \tau_{ghk} F_{1h(t-1)} F_{2k(t-1)};$$

$$\mu_{s11tg} = \frac{1}{2f_{s11tg}} \sum_{h,k} \tau_{ghk} \sum_{a_1,c_1} \sum_{a_2,c_2} \frac{\delta_{1a_1c_1} f_{1a_1c_1(t-1)h}}{F_{1h(t-1)}} \frac{\delta_{2a_2c_2} f_{2a_2c_2(t-1)k}}{F_{2k(t-1)}} (\mu_{1a_1c_1(t-1)h} + \mu_{2a_2c_2(t-1)k}),$$

where  $s = 1$  for sire,  $s = 2$  for dam,  $a_1$  and  $a_2$ , and  $c_1$  and  $c_2$  are the age and category for the sire and dam.

For genotypic assortative mating, the mating may be organized according to the genotypes of the parents. The optimization of these matings may be achieved by choosing a value for  $\alpha_{hkt}$ , which is the proportion of newborn animals in generation  $t$  from a sire with genotype  $h$  and a dam with genotype  $k$ . This proportion is constrained by

$$\sum_k \alpha_{hkt} = F_{1h(t-1)};$$

$$\sum_h \alpha_{hkt} = F_{2k(t-1)}.$$

In this situation, the genotype frequency and the EBV mean are given by

$$f_{s11tg} = \sum_{h,k} \tau_{ghk} \alpha_{hkt};$$

$$\mu_{s11tg} = \frac{1}{2f_{s11tg}} \sum_{h,k} \tau_{ghk} \alpha_{hkt} \sum_{a_1, c_1} \sum_{a_2, c_2} \frac{\delta_{1a_1 c_1} f_{1a_1 c_1(t-1)h}}{F_{1h(t-1)}} \frac{\delta_{2a_2 c_2} f_{2a_2 c_2(t-1)k}}{F_{2k(t-1)}} (\mu_{1a_1 c_1(t-1)h} + \mu_{2a_2 c_2(t-1)k}).$$

Finally, other assortative matings can be considered. Most often the structure of the parent population includes different groups involved in mating such as males on progeny testing, elite males or females, selected non elite males or females. For instance, males in progeny testing may be mated exclusively to nonelite females. To include assortative matings in the model, reproducers are classified according to mating groups  $m$ .  $F_{msc(t-1)g}$  is the probability that a parent of group  $m$ , and sex  $s$  has genotype  $g$ . It is defined as

$$F_{msc(t-1)g} = \sum_{a,c} \delta_{msac} f_{sac(t-1)g},$$

where  $\delta_{msac}$  is the proportion of animals of age  $a$ , class  $c$ , and sex  $s$  within the  $m$ th group. When mating groups are considered, the equations of genotype frequency and EBV mean are modified accordingly. For instance in genotypic random mating:

$$f_{s11tg} = \sum_{h,k} \tau_{ghk} \sum_{m_1} \sum_{m_2} F_{m_1 1h(t-1)} F_{m_2 2k(t-1)};$$

$$\mu_{s11tg} = \frac{1}{2f_{s11tg}} \sum_{h,k} \tau_{ghk} \sum_{m_1} \sum_{m_2} \sum_{a_1, c_1} \sum_{a_2, c_2} \frac{\delta_{m_1 1a_1 c_1} f_{1a_1 c_1(t-1)h}}{F_{m_1 1h(t-1)}} \frac{\delta_{m_2 2a_2 c_2} f_{2a_2 c_2(t-1)k}}{F_{m_2 2h(t-1)}} (\mu_{1a_1 c_1(t-1)h} + \mu_{2a_2 c_2(t-1)k}).$$

## Optimization

The deterministic model described above allows global optimization of the selection scheme to maximize the frequency of the desired genotype for a monogenic trait, while minimizing the loss of genetic progress on a polygenic trait. These objectives may be achieved by taking into account the whole population (male and female) or a part of the population (female or male) for a given age. The decrease in genetic progress for the polygenic trait may be considered as a penalty term, which may be defined as a strict or as a progressive penalty. Under a strict penalty, the objective function is put to 0 when the genetic progress is below a threshold. Under a progressive penalty, this term is proportional to the decrease in genetic progress.

The frequencies of the favorable genotype and the average EBV in the population may be considered during all periods of selection (e.g., in the use of discounting) or only in the last year of selection.

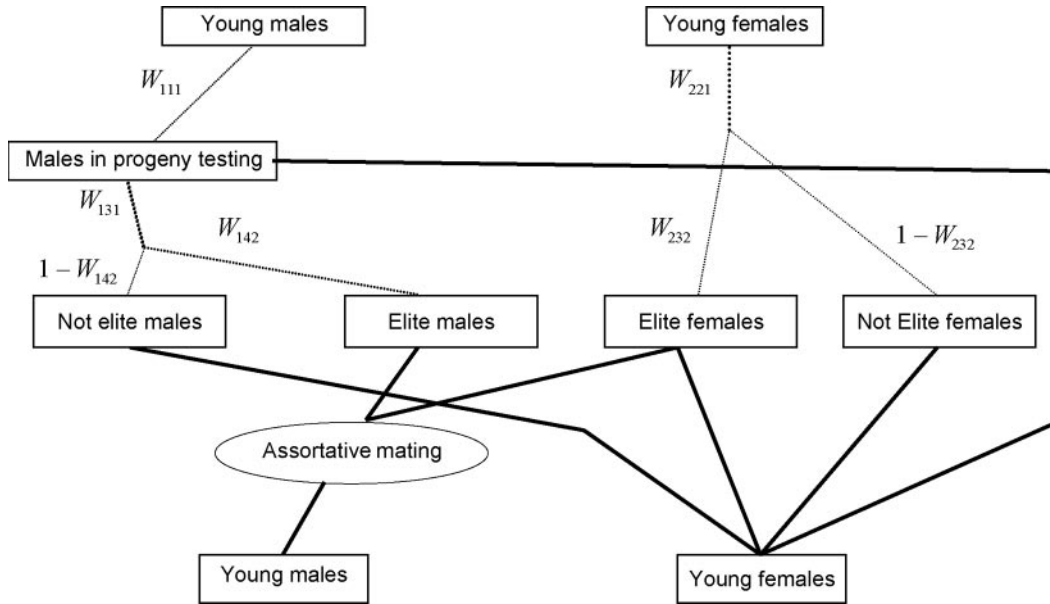
The decision variables  $w_{sactg}$  and  $\alpha_{hkt}$ , describing the evolution of the population, have to be optimized. In our approach, we did not calculate the derivatives of the objective function, differing from the approaches used by Dekkers and van Arendonk (1998) and Manfredi et al. (1998). A genetic algorithm (**GA**) was used to find the optimum solutions (Goldberg, 1989). A GA is an optimization technique based on an evolution of a population of artificial individuals. Each individual is a point in the research space and is characterized by its “chromosome”  $C_k$ , which determines the individual fitness  $h(C_k)$ ;  $k = 1, \dots, n$ ;  $n$  is a population size. The evolution process consists of successive generations. At each generation, individuals with high fitness are selected; the “chromosomes” of selected individuals are recombined and subjected to small mutations. A GA implemented as in Carroll (2001) was used in our study.

## APPLICATION

This model was applied to an example, where selection for scrapie resistance, which is controlled by a single gene (the *PrP* gene), was introduced as an additional selection criterion in an already existing dairy sheep selection scheme. It is expected that genetic progress in the polygenic trait (e.g., milk production) will be reduced under this selection objective.

Scrapie is a spongiform encephalopathy in sheep similar to bovine spongiform encephalopathy in bovine and Creutzfeldt-Jakob disease in humans. In sheep, genetic variability of the resistance to scrapie has been demonstrated and is mainly determined by the PrP polymorphisms (Elsen et al., 1998; Fabre et al., 2006). Thus, selection on the PrP locus was proposed to improve scrapie resistance in sheep populations. Equations corresponding to this example are given in the Appendix.

The population (Figure 3) was composed of animals, males and females, aged from 0 to 6 yr. Females were



**Figure 3.** Population description with 7 categories (rectangles), 7 mating paths (thick lines), selection steps (dotted lines), and 5 selection proportions  $W_{sac}$  (sex  $s=1, 2$ ; age  $a=1, \dots, a_s$ ; category  $c=1,2$ ).

selected on their own performance (with a selection rate  $W_{221}$ ) after their first lactation between 2 and 3 yr of age (Figure 3). The selected females were distributed in 2 categories, elite ( $W_{232}$ ) and nonelite ( $1 - W_{231}$ ). Among the young males born from matings between elite reproducers, a fraction  $W_{111}$  was mated to nonelite females and their EBV for the polygenic trait were estimated after a progeny test. A fraction  $W_{131}$  of these males was selected 2 yr later. Among the selected males, a proportion  $W_{142}$  was mated to the elite females to produce young male candidates for selection in the next generation. The replacement of females resulted from mating between nonelite females and elite, non-elite males or males in progeny testing; and between elite females and elite males. Two alleles (favorable vs. unfavorable) were considered for the Prp gene; thus, 3 genotypes exist ( $g = 1$  favorable genotype,  $g = 2$  intermediary genotype, and  $g = 3$  unfavorable genotype). The PrP genotypes were known for females and males. Only males were selected on their PrP genotype and on their EBV value before and after progeny testing. The proportion  $\delta$  of males in progeny test among reproductive males was equal to 0.2. The proportion  $\lambda_a$  of females with age  $a$  (4, 5, and 6 yr) among reproductive females was 0.4, 0.3, and 0.3, respectively. This selection scheme was applied during 15 yr ( $t = 15$ ) with the following parameter values:  $W_{221} = 0.8$ ;  $W_{232} = 0.1$ ;  $W_{111} = 0.225$ ;  $W_{131} = 0.2$ ;  $W_{142} = 0.3$ ;  $\sigma = 36$ . The accuracy for females ( $\rho_f$ ) was 0.73. The accuracy for males before and after progeny test was 0.56 ( $\rho_{m1}$ ) and 0.77 ( $\rho_{m2}$ ), respectively. The accuracy for elite males ( $\rho_{m3}$ ) was 0.89. The GA parameters were as follows: population size = 100; number of generations = 1,000; creep mutation probability = 0.02; crossover probability = 0.5; jump mutation probability = 0.01.

The aim was to maximize scrapie resistance, minimizing the loss of genetic progress in milk production. The

objective function chosen was the frequency of homozygous favorable females ( $s = 2, g = 1$ ) with a penalty on the genetic progress, constraining it not to be less than a percentage  $\pi$  of the progress obtained without scrapie selection ( $\Delta G_T$ ). The objective function can be written as follows:

$$Fobj = \left( \frac{\sum_{a=1}^6 \sum_{c=1}^2 \frac{1}{a} f_{2acT1}}{\sum_{a=1}^6 \frac{1}{a}} \right) - 100 \times \frac{\pi \Delta G_T - \sum_{g=1}^3 f_{211Tg} \mu_{211Tg}}{\pi \Delta G_T}.$$

The 1st term is the mean of homozygous favorable females of the elite ( $c = 1$ ) and non elite ( $c = 2$ ) group weighted by the coefficient  $\frac{1}{a}$  ( $a$  is the age), which gives more importance to younger females. All of the females were considered in the objective function because all of them will participate in the renewal. The second term is proportional to the difference in genetic progress obtained in this selection scheme and without selection on scrapie resistance.

First, the model was used with different initial frequencies of the favorable genotype without any constraint on genetic progress ( $\pi = 100\%$ ). Three initial frequencies of the favorable genotype were considered (0.03, 0.3, and 0.69). Then, situations with different constraints on genetic progress were considered while fixing the initial frequencies. Finally, simplifications of the strategy were explored to make the optimized selection procedures more practicable for farms.

**Effect of Different Initial Frequencies Without Constraint on Genetic Progress**

Whatever the initial frequency, the unfavorable genotype disappeared after 15 yr of selection (Table 1). The

**Table 1.** Final genotype frequencies for young females, objective function, and loss of genetic progress for 3 initial frequencies (genotype 1, favorable genotype; genotype 2, intermediary genotype; genotype 3, unfavorable genotype) without any constraint on loss of genetic progress

| Initial frequency |            |            | Final frequency for young females |            |            |                    | Loss of genetic progress, % |
|-------------------|------------|------------|-----------------------------------|------------|------------|--------------------|-----------------------------|
| Genotype 1        | Genotype 2 | Genotype 3 | Genotype 1                        | Genotype 2 | Genotype 3 | Objective function |                             |
| 0.03              | 0.28       | 0.69       | 0.77                              | 0.23       | 0.00       | 0.68               | 18.10                       |
| 0.30              | 0.50       | 0.20       | 0.91                              | 0.09       | 0.00       | 0.88               | 6.40                        |
| 0.699             | 0.28       | 0.03       | 0.97                              | 0.03       | 0.00       | 0.96               | 2.00                        |

loss of genetic progress was more important (18.1%) with a low initial frequency of the favorable genotype than with a high initial frequency (2.0%). This loss is much greater in the first few years, and impacts in the first few years are very important if counting discounted returns. The optimal selection proportions and mating rules evolve during the 15 yr process. The favorable genotype frequency increased, keeping a high selection proportion of the homozygous-resistant males before and after their progeny test. When the initial frequency of the favorable genotype (0.03 or 0.3) was low, the selection proportions of males carrying this genotype were high (0.9 and 0.7) during the first years to rapidly fix this genotype in the sires. These selection proportions decreased progressively when all sires became resistant. When the initial frequency of the favorable genotype (0.69) was high, the selection proportions for the favorable genotype were less, in particular before progeny testing (0.32 instead of 0.9 and 0.7). Moreover, the selection proportions after the progeny test were constant from the beginning, whereas they were constant only after yr 9 and 4 for the case of low frequencies (0.03 and 0.3). Whatever the initial frequencies of the favorable genotype, none of the susceptible males were selected, eliminating all of them from the beginning. A fully favorable population of sires was reached more rapidly with a high initial frequency of the favorable genotype (yr 3) than with a low frequency (yr 8).

The reduction of the unfavorable genotype frequencies in the dam population occurred very slowly because (i) dams were selected just once, (ii) the selection threshold was the same across genotype classes, and (iii) the selection pressure was low. With a high initial frequency of the favorable genotype, the objective was reached and the dam population was fully resistant. However, with a low frequency, only sires were resistant at the end of yr 15, with the unfavorable genotype still segregating in the dam population.

### *Effects of Different Constraints on the Loss of Genetic Progress*

In contrast to the situation described above in which no constraint was put on genetic progress, a 10% maximum loss of genetic progress at the end of the process was allowed ( $\pi = 90\%$ ) for the situation with the decreased initial frequency (0.03). The case with the ini-

tial low frequency of the favorable genotype was studied because the loss of genetic progress with this frequency was the most important (18.1%). In this situation, the optimal selection proportions of males with a favorable genotype were less and fluctuated between high and low values (Figure 4). These fluctuations can be explained by the fact that the genetic progress must be kept over a threshold. Each year, the selection intensity varies to compensate for the loss in genetic progress. Contrary to the previous case, the selection proportions of favorable and heterozygous genotypes were high and similar (0.7). The selection proportions for homozygous unfavorable sires were null. After the elimination of unfavorable genotype carriers, the selection proportions of the favorable genotype (20 to 30%) were greater than for the heterozygous genotype (5 to 20%). The heterozygous genotype was conserved for a longer time than without any constraint on the genetic progress. The objective of achieving a fully resistant population of sires was reached less rapidly with 10% of maximum loss of genetic progress (yr 11) than without any constraint on genetic progress (yr 8). In the dam population, proportions of favorable homozygous and heterozygous genotypes were similar. Dams with an unfavorable genotype were almost entirely eliminated, being kept at less than 5%. Thus, the optimal strategy of selection depends on the constraints put on genetic progress. If the priority is to reach a population fully resistant to the disease, the loss of genetic progress may not be essential; the selection on genotype information is drastic at the beginning of the process and the animals are quickly fully resistant. If a more important penalty is put on the loss of genetic progress, the selection changes, with cycles of high and low selection proportions for the favorable genotype. Eventually, the unfavorable genotype disappears and the population is divided between favorable homozygous and heterozygous genotypes.

In the previous case (Figure 4a), the selection proportions decreased gradually as the population became progressively resistant. With a constraint on genetic progress, the fluctuation of selection proportions was larger (Figure 4b). The selection proportions of the favorable genotype were high during the first years, inducing an increase in resistance during this period; they were less than with previous cases to maintain acceptable genetic progress. This optimal strategy would not be practicable for breeders because of the annual varia-

tions of the selection proportions. However, a detailed analysis of selection proportions and assortative mating during the 15 selection years suggests some simplifications.

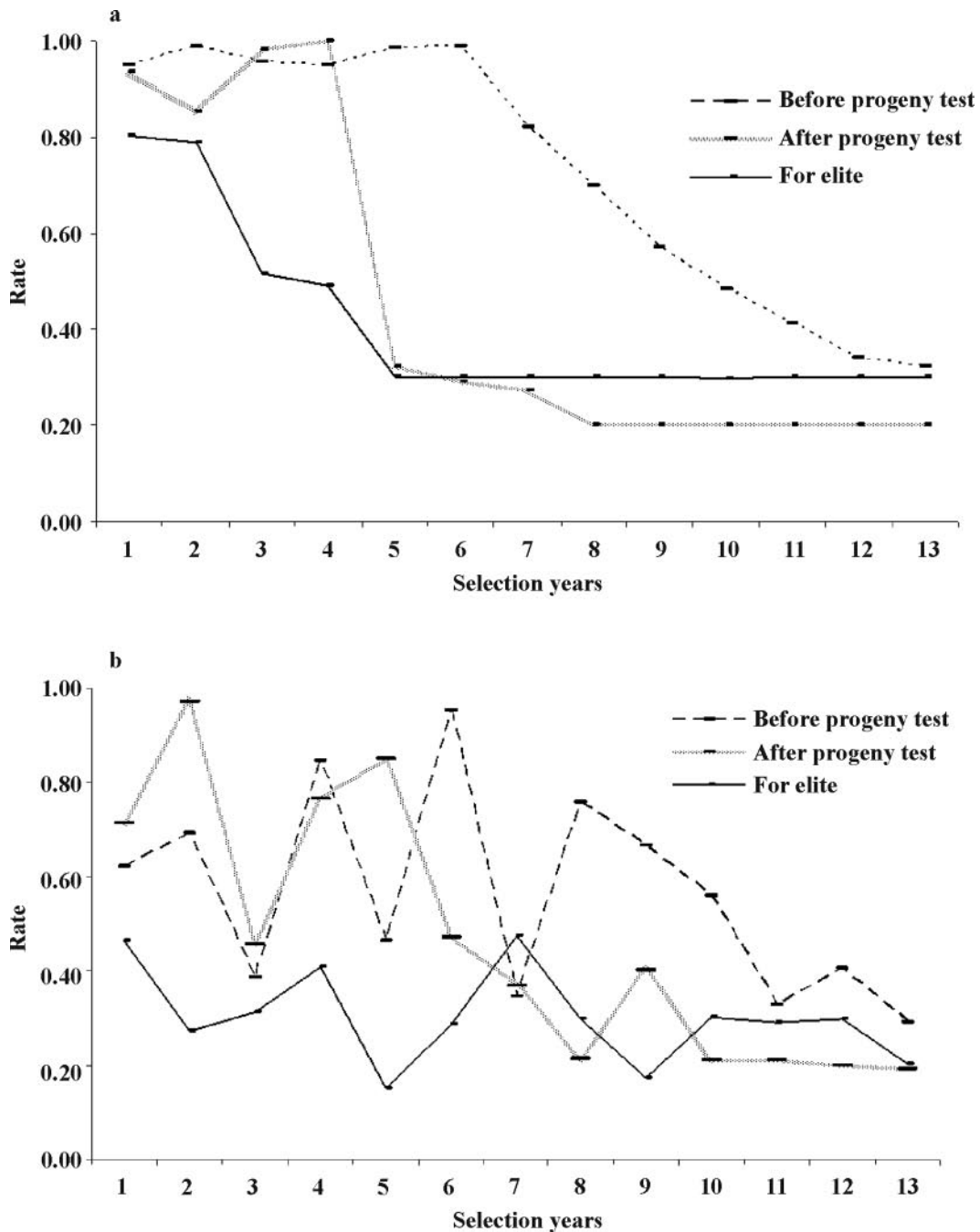
**Simplifications of the Optimal Strategy**

Two cases were analyzed: (1) a low initial frequency of the favorable genotype and a maximum 10% loss of genetic progress accepted after 15 selection years and (2) a low initial frequency of the favorable genotype without any constraint on the loss of genetic progress. The results are given in Table 2. The first case shows

the evolution of the objective function without any penalty on loss of genetic progress. The second case shows the evolution of the objective function with a maximum loss of 10% in the genetic progress.

In the optimal strategy, the selection proportions of the males before progeny testing ( $W_{111}$ , Figure 3) varied between years and thus needed to be optimized. On the contrary, some selection proportions were constant after a few years of selection, suggesting some simplifications.

1. Within genotype selection proportions of males after progeny testing: The stabilization of these



**Figure 4.** Selection rate evolution for the favorable genotype ( $g = 1$ ) without (a) and with (b) constraint in loss of genetic progress before progeny test (-), after progeny test (····), and for elite males (—).

**Table 2.** Value of the objective function with some simplifications in 2 cases without constraint on loss of genetic progress (L-0) and with 10% of loss of genetic progress (L-10)

| Item  | L-0  | L-10 |
|---|------|------|
| Fully optimized scheme                                  | 0.68 | 0.53 |
| Constant selection rate for elite males                 | 0.67 | 0.53 |
| Constant selection rate after progeny testing           | 0.56 | 0.48 |
| Random mating   | 0.68 | 0.52 |
| Constant selection rate for elite males + random mating | 0.67 | 0.52 |

selection proportions is late, after the 11th selection year. However, the fixation of these selection proportions was tested. The values of the objective function were less than in the optimal strategy. Thus, these selection proportions have to be optimized.

2. Within genotype selection proportions for elite males: The results of the previous optimization suggest a suboptimal strategy where these selection proportions are constant. The objective function value obtained with this simplification was similar to the optimal one (Table 2). Consequently, it appears that it is not critical for within genotype selection proportions for elite males to be optimized every year.

Another question was to check if assortative mating based on the genotypes was necessary or not. The slight difference in the objective function value between the fully optimized scheme and the simplified one shows that mating optimization was not necessary.

Globally, 2 simplifications were thus possible: fixed selection proportions for elite males and random mating. The case combining simplifications gave a value of the objective function similar to the optimal one. Moreover, the penalty weight of the loss on genetic progress was identical with or without simplifications. Implementing other simplifications could give a reduced value of the objective function than the one obtained in the fully optimized scheme, but these were not considered here.

In our example, we proposed that the objective was to maximize the genotypic frequency of the favorable female genotypes with a penalty on genetic progress. Avoiding loss of genetic progress increases the time needed to reach fixation of the favorable genotype. For example, sires are all resistant at the end of yr 8 or after yr 11 without any constraint or with a maximum loss of 10% of genetic progress, respectively. Other objective functions are possible and must be adapted to the aim of the breeder. For example, the frequency of males may be considered instead of the frequencies of females. Males have a more important impact on the selection scheme. In this example, the objective was to keep the homozygous genotype; however, in the case of breeds with a risk of extinction, it could be preferable to keep heterozygote genotypes. Another option could

be to consider the whole process, rather than only the final state of the population, possibly using a discounting factor if sooner is better.

## DISCUSSION

In this paper, a method was developed to optimize selection on an identified QTL or gene and a quantitative trait over multiple generations. The breeding program modeled assumed assortative mating and overlapping generations. This modeling may easily be adapted to a variety of situations concerning the underlying genetic model and the selection scheme. In this study, 2 alleles were considered, but the methodology can consider a multi-allelic case. However, the number of alleles (and genotypes) should be limited. The number of optimized parameters increases in an exponential manner with the number of alleles considered. For instance, when 3 genotypes (2 alleles) are considered, 10 parameters have to be optimized, and if 10 genotypes (4 alleles) are considered, the optimization problem would involve 108 parameters. Only the case with an additive QTL or gene was considered here, but dominance could be included. The method can also be extended to multiple QTL or genes with any type of linkage between them (epistasis).

Different methodologies to solve multiple-generation selection problems in animal breeding are possible. In this study, a GA was used to solve the optimization problem. Genetic algorithms are more flexible when additional constraints have to be added than optimal control theory, the method used by Dekkers and van Arendonk (1998), because to optimize selection the derivative of the selection criterion has to be calculated with optimal control theory. Use of a GA bypasses this difficulty.

Additive variance was assumed to be constant over generations, and genotyping errors and relationship errors were assumed to be nonexistent. The reduction of the additive variance by selection could be considered in the model.

The strategy developed here consisted of selecting animals that were resistant to a particular disease to decrease the number of infected animals. The aim was to reduce the incidence of the disease, while keeping the genetic progress on the traditional production traits.

In this model, inbreeding was not considered. A possibility for taking this into account would be to add the notion of family for males. Males of one family participate for the renewal of the males for their family and males of different families participate for the renewal of the females. This possibility would allow a certain genetic variability to be maintained.

Domestic animal diseases might represent a risk to human health. Thus, identification of genes involved in disease resistance is an important genetic research subject. The scrapie case is an example where a national selection program for scrapie resistance was implemented as a precautionary measure. This dynamic and deterministic model was applied here in the scrapie case; however, it can be extended to other genes and selection schemes. For example, genes affecting important performance traits such as adiposity or meat quality in pigs (Le Roy et al., 1990; Larzul et al., 1997a; Hamilton et al., 2000) or prolificacy in sheep (Davis, 2005; Fabre et al., 2006) could use this model. Moreover, this deterministic model could evolve to a stochastic model for a limited number of breeding animals in a small population. Our model would need to be simplified for use in stochastic simulation because stochastic simulations with a large number of choice variables can be impractical.

## LITERATURE CITED

- Carroll, D. L. 2001. The genetic algorithm (ga) driver. <http://cuaerospace.com/carroll/ga.html> Accessed Aug. 25, 2008.
- Chakraborty, R., L. Moreau, and J. C. M. Dekkers. 2002. A method to optimize selection on multiple identified quantitative trait loci. *Genet. Sel. Evol.* 34:145–170.
- Davis, G. H. 2005. Major genes affecting ovulation rate in sheep. *Genet. Sel. Evol.* 37(Suppl. 1):S11–S23.
- Dekkers, J. C. M., and R. Chakraborty. 2001. Potential gain from optimizing multigeneration selection on an identified quantitative trait locus. *J. Anim. Sci.* 79:2975–2990.
- Dekkers, J. C. M., and J. A. M. van Arendonk. 1998. Optimizing selection for quantitative traits with information on an identified locus in outbred populations. *Genet. Resour. Crop Evol.* 71:257–275.
- Elsen, J. M., J. Vu Tien Khang, F. Schelcher, Y. Amigues, J. P. Poivey, F. Eychenne, F. Lantier, and J. L. Laplanche. 1998. A scrapie epidemic in a closed flock of Romanov. *Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia.* 27:269–272.
- Fabre, S., A. Pierre, P. Mulsant, L. Bodin, E. Di Pasquale, L. Persani, P. Monget, and D. Monniaux. 2006. Regulation of ovulation rate in mammals: Contribution of sheep genetic models. *Reprod. Biol. Endocrinol.* 4:20.
- Gibson, J. P. 1994. Short term gain at the expense of long term response with selection of identified loci. *Proc. 5th World Congr. Genet. Appl. Livest. Prod. Univ. Guelph, Ontario, Canada.* 21:201–204.
- Goldberg, D. E. 1989. *Genetics algorithms in search, optimization and machine learning.* Addison-Wesley Publ. Co. Inc., Boston, MA.
- Hamilton, D. N., M. Ellis, K. D. Miller, F. K. McKeith, and D. F. Parrett. 2000. The effect of the Halothane and Rendement Napoleon genes on carcass and meat quality characteristics of pigs. *J. Anim. Sci.* 78:2862–2867.
- Holland, J. H. 1973. Genetic algorithms and the optimal allocation of trials. *SIAM J. Comput.* 220:88–105.
- Lande, R., and R. Thompson. 1990. Efficiency of marker assisted selection in the improvement of quantitative traits. *Genetics* 124:743–756.
- Larzul, C., L. Lefaucheur, P. Ecolan, J. Gogue, A. Talmant, P. Sellier, P. Le Roy, and G. Monin. 1997a. Phenotypic and genetic parameters for longissimus muscle fiber characteristics in relation to growth, carcass, and meat quality traits in Large White pigs. *J. Anim. Sci.* 75:3126–3137.
- Larzul, R., E. Manfredi, and J. M. Elsen. 1997b. Potential gain from including major gene information in breeding value estimation. *Genet. Sel. Evol.* 29:161–184.
- Le Roy, P., J. Naveau, J. M. Elsen, and P. Sellier. 1990. Evidence for a new major gene influencing meat quality in pigs. *Genet. Res.* 55:33–40.
- Lewis, F. L. 1986. *Optimal Control.* Wiley, New York, NY.
- Manfredi, E., M. Barbieri, F. Fournet, and J. M. Elsen. 1998. A dynamic deterministic model to evaluate breeding strategies under mixed inheritance. *Genet. Sel. Evol.* 30:127–148.
- Meuwissen, T. H., and J. A. van Arendonk. 1992. Potential improvements in rate of genetic gain from marker-assisted selection in dairy cattle breeding schemes. *J. Dairy Sci.* 75:1651–1659.
- Mongeau, M. 2003. *Exploitation de la structure en optimisation globale, Laboratoire Mathématiques pour l'Industrie et la Physique.* Handbook. UMR 5640. Toulouse, France.
- Pong-Wong, R., and J. A. Woolliams. 1998. Response to mass selection when an identified major gene is segregating. *Genet. Sel. Evol.* 30:313–337.

## APPENDIX: EQUATIONS OF THE EXAMPLE

Two-year females:

$$f_{221tg} = f_{211(t-1)g};$$

$$\mu_{221tg} = \mu_{211(t-1)g}.$$

Three-year nonelite females:

$$f_{232tg} = f_{221(t-1)g} \frac{w_{221tg}(1 - w_{232tg})}{W_{221}(1 - W_{232})}.$$

$$\mu_{232tg} = \mu_{221(t-1)g} + \frac{i(w_{221tg}) - w_{232tg} \ i(w_{221tg} w_{232tg})}{(1 - w_{232tg})} \rho_f \sigma,$$

where  $i()$  is the selection intensity.

Three-year elite females:

$$f_{231tg} = f_{221(t-1)g} \frac{w_{221tg} w_{232tg}}{W_{221} W_{232}};$$

$$\mu_{231tg} = \mu_{221(t-1)g} + \frac{i(w_{221tg}) - w_{221tg} \ i(w_{221tg} w_{232tg})}{w_{232tg}} \rho_f \sigma.$$

Four- to six-year females:

$$\mu_{2a1tg} = \mu_{2(a-1)1(t-1)g};$$

$$f_{2a1tg} = f_{2(a-1)1(t-1)g};$$

$$\mu_{2a2tg} = \mu_{2(a-1)2(t-1)g}; \text{ and}$$

$$f_{2a2tg} = f_{2(a-1)2(t-1)g}.$$

Two-year males before progeny testing:

$$f_{121tg} = f_{111(t-1)g} \frac{w_{111tg}}{W_{111}};$$

$$\mu_{121tg} = \mu_{111(t-1)g} + i(w_{111tg})\rho_{m1}\sigma.$$

Two-year males after progeny testing:

$$f_{131tg} = f_{121(t-1)g};$$

$$\mu_{131tg} = \mu_{121(t-1)g}.$$

Four-year nonelite males:

$$f_{142tg} = f_{131(t-1)g} \frac{w_{131tg}w_{142tg}}{W_{131}W_{142}};$$

$$\mu_{142tg} = \mu_{131(t-1)g} + i(w_{131tg}w_{142tg})\rho_{m3}\sigma.$$

Four-year elite males:

$$f_{141tg} = f_{131(t-1)g} \frac{w_{131tg}(1-w_{142tg})}{W_{131}(1-W_{142})};$$

$$\mu_{141tg} = \mu_{131(t-1)g} + \frac{i(w_{131tg}) - w_{142tg}i(w_{131tg}w_{142tg})}{1-w_{142tg}}\rho_{m2}\sigma.$$

Renewal.

Newborn males:

$$f_{111(t+1)g} = \sum_{h=1}^3 \sum_{k=1}^3 \tau_{ghk} \alpha_{hkt};$$

$$\begin{aligned} \mu_{111(t+1)g} &= \frac{1}{2} \frac{1}{f_{111(t+1)g}} \\ &= \frac{1}{2} \frac{1}{\sum_{h=1}^3 \sum_{k=1}^3 \tau_{ghk} \alpha_{hkt} \sum_{a_2=3}^6 \frac{f_{2a_2 2tk} \lambda_{a_2}}{\sum_{a_2=3}^6 f_{2a_2 2tk} \lambda_{a_2}}} (\mu_{142th} + \mu_{2a_2 2tk}). \end{aligned}$$

Newborn females:

$$\begin{aligned} f_{211(t+1)g} &= \sum_{h=1}^3 \sum_{k=1}^3 \tau_{ghk} (\alpha_{hkt} W_{232} \\ &+ \sum_{a_2=3}^6 \frac{f_{2a_2 1tk} \lambda_{a_2}}{(1-\delta)m_e W_{142} + (1-\delta)m_{ne}(1-W_{142}) + \delta m_T} \\ &+ \sum_{a_2=3}^6 \frac{f_{142th} ((1-\delta)m_e W_{142} - W_{f2} ((1-\delta)m_e W_{142} + (1-\delta)m_{ne}(1-W_{142}) + \delta m_T))}{(1-\delta)m_e W_{142} + (1-\delta)m_{ne}(1-W_{142}) + \delta m_T} \\ &+ f_{141th} (1-\delta)m_{ne}(1-W_{142}) + \delta m_T f_{121th}); \end{aligned}$$

$$\begin{aligned} \mu_{211(t+1)g} &= \frac{1}{2} \frac{1}{f_{211(t+1)g}} \\ &= \frac{1}{2} \frac{1}{\sum_{h=1}^3 \sum_{k=1}^3 \tau_{ghk} (\alpha_{hkt} W_{232} + \sum_{a_2=3}^6 \frac{f_{2a_2 2tk} \lambda_{a_2}}{\sum_{a_2=3}^6 f_{2a_2 2tk} \lambda_{a_2}}} (\mu_{142th} + \mu_{2a_2 2tk}) \end{aligned}$$

$$\begin{aligned} &+ \sum_{a_2=3}^6 f_{142th} f_{2a_2 1tk} \lambda_{a_2} \\ &\times \left( \frac{(1-\delta)m_e W_{142}}{(1-\delta)m_e W_{142} + (1-\delta)m_{ne}(1-W_{142}) + \delta m_T} - W_{232} \right) \\ &\times (\mu_{142th} + \mu_{2a_2 1tk}) \end{aligned}$$

$$\begin{aligned} &+ \sum_{a_2=3}^6 f_{141th} f_{2a_2 1tk} \lambda_{a_2} \\ &\times \left( \frac{(1-\delta)m_{ne}(1-W_{142})}{(1-\delta)m_e W_{142} + (1-\delta)m_{ne}(1-W_{142}) + \delta m_T} \right) \\ &\times (\mu_{141th} + \mu_{2a_2 1tk}) \end{aligned}$$

$$\begin{aligned} &+ \sum_{a_2=3}^6 f_{2a_2 1tk} \lambda_{a_2} f_{121th} \\ &\times \left( \frac{\delta m_T}{(1-\delta)m_e W_{142} + (1-\delta)m_{ne}(1-W_{142}) + \delta m_T} \right) \\ &\times (\mu_{121th} + \mu_{2a_2 1tk}). \end{aligned}$$

## References

This article cites 16 articles, 4 of which you can access for free at:  
<http://www.journalofanimalscience.org/content/87/3/885#BIBL>