Mating and selection schemes when rates of inbreeding are constrained

Anna K. Sonesson and Theo H. E. Meuwissen

DLO - Institute for Animal Science and Health, Box 65, 8200 AB Lelystad, The Netherlands

Abstract

The effect of five mating schemes on genetic gain was compared for populations with discrete generations. Selection and mating were made in two steps. Parents were first selected with the Optimum Contribution selection method, where genetic gain is maximised while the average coancestry of selected animals is constrained. Thereafter, the mating algorithms were applied for the selected animals only. This two-step selection and mating strategy requires less computer time, compared to a one-step selection and mating scheme, where all selection candidates are also included in the mating step. Minimum coancestry (MC), Minimum coancestry with a maximum of one offspring per fullsib family (MC1) and Minimum variance of the relationships of offspring (MVRO) mating schemes resulted in a delay in inbreeding of about two generations compared to Random and Random factorial (R1) mating. In breeding schemes where selection is constraining rates of inbreeding $\Delta F$, any improvement in family structure due to non-random mating increases genetic gain. For schemes with $\Delta F$ constrained to 1.0%, $h^2=0.25$ and 100 selection candidates (50 males and 50 females), genetic gain was 22% higher for the MC1 and MVRO schemes compared to Random mating schemes. For schemes with a less stringent constraint on $\Delta F$ (2.5%) or more selection candidates (200), the superiority of the MC1 and MVRO schemes was smaller (5-6%). In general, MC1 seemed the preferred mating method, since it almost always yielded the highest genetic gain.

1. Introduction

Mating and selection schemes attempt to reduce rates of inbreeding and/or increase genetic gain. Optimum Contribution selection methods are maximising genetic gain while constraining inbreeding by restricting coancestry among selected parents (Wray and Goddard, 1994; Brisbane and Gibson, 1995; Meuwissen, 1997; Meuwissen and Sonesson, 1998; Grundy et al., 1998; Sonesson et al., 1999; Grundy et al., 1999). These authors assumed however random mating among selected parents, whereas non-random mating schemes may provide more opportunities for Optimum Contribution selection to select superior animals in the next generation. Caballero et al. (1996) concluded that minimum coancestry and compensatory mating strategies generally reduce rates of inbreeding and that they have a small effect on the genetic response to selection. Caballero et al. (1996) used phenotypic and BLUP selection. With Optimum Contribution selection schemes, rates of inbreeding are however set to a desired value in the selection scheme and are not expected to change for different mating schemes. Hence, we expect that the Optimum Contribution selection will increase rates of genetic gain by using any mating achieved improvement in structure of the relationships of the selection candidates.

The aim of this paper is to investigate rates of genetic gain and rates of inbreeding for five mating schemes that are combined with Optimum Contribution selection in a two step selection and mating optimisation procedure, i.e. only selected animals are included in the mating step. For all schemes, the number of offspring for each selected animal is given by the selection scheme. The five mating schemes are:
1. Random mating;
2. Random mating with only one offspring per full sib family (i.e. the factorial mating design of Woolliams (1989));
3. Minimum coancestry mating where the least related animals are mated;
4. Minimum coancestry mating with only one offspring per full sib family, and
5. A mating scheme that minimises variance of coancestry among the progeny.

The mating schemes mainly affect the family structure of the selection candidates, either by permitting only one offspring per full-sib family (schemes 2 and 4), by connecting certain families (schemes 3 and 4) or by reducing extreme relationships (scheme 5).

2. Material and Methods

2.1. Selection method

The method of Meuwissen (1997) was used to select animals. This method maximises the genetic level of next generation of animals, \( G_{t+1} = c_t'EBV_t \), where \( c_t \) is a vector of genetic contributions of the selection candidates to generation \( t+1 \) and \( EBV_t \) is a vector of BLUP estimated breeding values of the candidates for selection in generation \( t \). Rates of inbreeding are controlled by constraining the average coancestry to \( \overline{C}_{t+1} = c_t'A_tA_t/2 \), where \( A_t \) is a \((n \times n)\) relationship matrix among the selection candidates, \( \overline{C}_{t+1} = 1 - (1 - \Delta F_d)^4 \), and \( \Delta F_d \) is the desired rates of inbreeding (Grundy et al., 1998). Note that the level of the constraint \( \overline{C}_{t+1} \), can be calculated for every generation, before the breeding scheme commences. For optimisation of \( c_t \), it is also necessary to constrain the sum of the contributions of males (females) to \( 1/2 \), i.e. \( Q'c_t = 1/2 \) where \( Q \) is a \((n \times 2)\) incidence matrix of the sex of the selection candidates (the first column yields ones for males and zeros for females, and the second column yields ones for females and zeros for males) and \( 1/2 \) is a \((2 \times 1)\) vector of halves.

In order to obtain the optimal \( c_t \) that maximises \( G_{t+1} \), LaGrangian multipliers are used, which yields the following quadratic index, \( I_t \) (Meuwissen, 1997):

\[
I_t = c_t'EBV_t - \lambda_0(c_t'A_tA_t - 2\overline{C}_{t+1}) - (c_t'Q^{-1/2})\lambda 
\]  

where \( \lambda_0 \) and \( \lambda \) are LaGrangian multipliers \((\lambda = (2 \times 1)\) vector of LaGrangian multipliers). The objective function, \( c_t'EBV_t \), is maximised for \( c_t \) under two restrictions; the first one is on the average relationship of the selection candidates and the second one is on the contribution per sex. The optimisation procedure was explained in Meuwissen (1997). The number of selection candidates per generation, \( T \), is 100 or 200 in all breeding schemes. The output from the selection method is a vector with genetic contribution for each selection candidate, \( c_t \).

2.2. Random mating (R)

For the R scheme, a sire and a dam are allocated at random for each new born progeny with a probability that is proportional to the genetic contribution that they received from the selection algorithm, \( c_t \). This scheme is not a strictly hierarchical scheme, since dams can be mated to multiple sires. This scheme is therefore expected to have lower inbreeding and/or higher genetic gain than the commonly studied hierarchical breeding schemes.

2.3. Minimum coancestry mating (MC and MC1)

With minimum coancestry mating, MC, the average relationship of sires and
dams and therefore also the inbreeding of their progeny is minimised. A matrix $F$ of size $(N_s \times N_d)$ is set up, where $N_s$ ($N_d$) is the number of selected sires (dams) and element $F_{ij}$ is the coefficient of kinship of selection candidates which is also the inbreeding coefficient of their progeny. MC schemes can result in many full sib offspring, because a sire has only minimum coancestry with one dam. However, many full sib relationships in a scheme result in either less genetic gain or more inbreeding compared to the situation where the full sib relationships are replaced by paternal and maternal half sib relationships (Woolliams, 1989). A higher inbreeding in generation $t$ can lead to problems for the selection algorithm to constrain the inbreeding in generation $t+1$. Hence, for MC1 an additional constraint is imposed, namely that each mating pair should obtain only one offspring. However, this can not always be achieved, e.g. when a sire should obtain more offspring than the number of dams that are selected, i.e. one offspring per dam does not suffice. In the latter case, the sire will obtain more than one offspring per dam.

Minimum coancestry matings were obtained by applying the simulated annealing algorithm (Press et al., 1989). The implementation of the annealing algorithm is described in the Appendix. The alternative solutions that are tried by the annealing algorithm differ from the current optimal solution by replacing mating pairs according to the scheme in Figure A1 of the Appendix, but changes were made at random and an optimal solution was not searched for as for the MC and MVRO schemes.

2.5. Minimum variance of relationship of offspring mating (MVRO)

This mating method reduces the variance of relationships of progeny of selected sires and dams, which reduces the number of extremely high relationships among the progeny. The latter also reduces the probability that two full sibs are made because of their strong relationship. Note that the average relationship of the offspring is calculated by the Optimum Contribution selection algorithm and it can not be changed by the mating algorithm. For MVRO schemes, a symmetric matrix $V$ is set up that yields the term that a pair of offspring contributes to the variance of the relationship of offspring, i.e.

$$V_{ij} = (a_{ij} - \bar{a})^2, \text{ for } i \neq j \text{ and } V_{ii} = 0 \text{ for } i=j,$$

where $a_{ij}$ is the relationship of offspring $i$ and $j$, and $i(j)$ denotes every possible offspring from all $N_s \times N_d$ mating combinations, i.e. $V$ is of size $(N_s \times N_d)$ and $\bar{a}$ is the average relationship of selected animals. The minimisation of terms that are selected from the $V$ matrix, i.e. selecting from the matings that are performed, is again done by the annealing algorithm (see Appendix). The alternative solutions that are tried by the annealing are the same as for MC mating.

2.6. The simulated breeding schemes

The general structure of the breeding schemes was that of a closed nucleus
scheme with 100 or 200 selection candidates. Genotypes, \( g_i \), of the base animals are sampled from the distribution \( N(0, \sigma^2_a) \), where \( \sigma^2_a \) is base generation genetic variance (0.25). Later generations are obtained by simulating offspring genotypes from \( g_i = \frac{1}{2}g_s + \frac{1}{2}g_d + m_i \), where \( s \) denotes the sire and \( d \) the dam of the offspring \( i \), and \( m_i \) is the Mendelian sampling component, which is sampled from \( N(0, \frac{1}{2}(1-F)\sigma^2_a) \), where \( F \) is the average inbreeding coefficient of parents \( s \) and \( d \). Phenotypes are simulated by adding an error term to the genotypes, which was sampled from \( N(0, \sigma^2_e) \). The base generation phenotypic variance \( \sigma^2_p = \sigma^2_a + \sigma^2_e \) is equal to 1. Estimates of breeding values (EBVs) are obtained using the BLUP- breeding value estimation procedure (Henderson, 1984). The simulated trait was recorded on both sexes, before selection. One hundred replicates of the scheme were run for 20 years with an inbreeding constraint of 1 or 2.5%.

3. Results

3.1. Inbreeding trend

For MC1, MC and MVRO, there was a delay in inbreeding of about two generations compared to for the R and R1 schemes. For MC there was somewhat lower F than for MC1. The latter is probably due to the extra restriction of having only one offspring per mating in MC1, which reduces the opportunities for MC1 of mating animals that are least related, i.e. its opportunity to generate offspring with lowest inbreeding. For all schemes, the realised rates of inbreeding were close to the constraint on the rates of inbreeding (Table 1), so a comparison of genetic gain between the different schemes is justified.

3.2. Genetic gain in small schemes with low inbreeding

For schemes where \( \Delta F \) was constrained to 1.0% and the number of selection candidates was 100, the highest genetic level at year 20, \( G_{20} \), was achieved for the MC1 and MVRO schemes and the lowest for the R schemes (Table 1). \( G_{20} \) was 22.3% higher for the MC1 \( (4.01 \sigma^2_p \text{ units}) \) and 22.6% higher for the MVRO schemes \( (4.02 \sigma^2_p \text{ units}) \) than for the R scheme \( (3.28 \sigma^2_p \text{ units}) \). For the R1 and MC schemes, genetic gain was somewhat lower than for the MC1 and MVRO schemes, i.e. \( G_{20} \) was 21.1% higher for the R1 \( (3.98 \sigma^2_p \text{ units}) \) than for the R scheme.

For all breeding schemes, genetic gain showed a linear increase over the years (results are not shown).

3.3. Genetic gain in small schemes with high inbreeding

As expected, a less stringent constraint on \( \Delta F \) of 2.5% resulted in higher \( G_{20} \) (Table 1). With this less stringent constraint on \( \Delta F \), the superiority of the non-random mating schemes was much lower. The highest \( G_{20} \) was achieved by the MC1 and MVRO schemes \( (6.9\% \text{ higher than R}) \) followed by the R1 \( (4.9\% \text{ higher than R}) \) and MC schemes \( (4.3\% \text{ higher than R}) \).

3.4. Genetic gain in a large scheme with low inbreeding

For the larger schemes where \( \Delta F \) was constrained to 1% per generation and the number of selection candidates was 200, the same pattern of \( G_{20} \) over the heritabilities was seen as for schemes where \( \Delta F \) was constrained to 2.5% per generation and the number of selection candidates was 100. As expected, \( G_{20} \) was in general higher than for the smaller schemes. The highest \( G_{20} \) was achieved for the MC1 (on average 6.7% higher than R over the three heritabilities) and MVRO
schemes (6.6% higher than R), followed by the MC schemes (5.8% higher than R) and R1 scheme (5.4% higher).

Table 1. Average rate of inbreeding ($\Delta F$), genetic level at generation 20 ($G_{20}$), number of selected sires and dams ($h^2=0.25$)

<table>
<thead>
<tr>
<th>$\Delta F$/gen</th>
<th>$G_{20}$ (s.e.)</th>
<th>Sel. sires/dams</th>
</tr>
</thead>
<tbody>
<tr>
<td>(%)</td>
<td>(s.e.)</td>
<td>no</td>
</tr>
<tr>
<td>R</td>
<td>1.00</td>
<td>3.28 (.0296)</td>
</tr>
<tr>
<td>R1</td>
<td>1.01</td>
<td>3.98 (.0249)</td>
</tr>
<tr>
<td>MC</td>
<td>1.00</td>
<td>3.98 (.0263)</td>
</tr>
<tr>
<td>MC1</td>
<td>1.00</td>
<td>4.01 (.0266)</td>
</tr>
<tr>
<td>MVRO</td>
<td>1.00</td>
<td>4.02 (.0291)</td>
</tr>
</tbody>
</table>

$\Delta F=1.0\%$, 100 selected candidates

| $\Delta F=2.5\%$, 100 selected candidates |
| R | 2.50 | 4.94 (.0396) | 14.7/14.8 |
| R1 | 2.48 | 5.18 (.0346) | 13.3/13.4 |
| MC | 2.49 | 5.15 (.0391) | 13.5/13.6 |
| MC1 | 2.49 | 5.28 (.0394) | 13.9/13.6 |
| MVRO | 2.48 | 5.28 (.0355) | 13.3/13.4 |

$\Delta F=1.0\%$, 200 selected candidates

| $\Delta F=2.5\%$, 200 selected candidates |
| R | 1.00 | 5.07 (.0279) | 36.5/36.1 |
| R1 | 0.99 | 5.34 (.0229) | 31.8/32.1 |
| MC | 1.00 | 5.33 (.0327) | 31.3/31.6 |
| MC1 | 1.00 | 5.42 (.0266) | 32.5/32.6 |
| MVRO | 1.00 | 5.43 (.0265) | 32.8/33.1 |

R = random mating, R1 = random mating with maximum one full-sib per couple, MC = Minimum coancestry mating, MC1 = minimum coancestry mating with maximum one full-sib per couple, MVRO = minimum variance of relationship of offspring

3.5. Number of selected animals

For schemes where $\Delta F$ was constrained to 1% and the number of selection candidates was 100, the highest number of animals was selected for the R scheme and the lowest number for the MC scheme. The number of animals that were selected for the MVRO and MC1 schemes was somewhat higher than the number of animals that were selected for the MC scheme. The number of animals selected was about halved when $\Delta F$ was constrained to 2.5% per generation, compared to when $\Delta F$ was constrained to 1.0% per generation, although the number of selection candidates was 100 in all schemes. There were about the same number of animals selected for the larger schemes with 200 selection candidates compared to for the smaller schemes, which indicates that the intensity of selection has increased for the larger schemes. Meuwissen and Sonesson (1998) found a similar result for Optimum Contribution selection schemes with overlapping generations.

4. Discussion

4.1. Genetic gain

Non-random mating schemes yielded an increase in genetic gain of up to 22% compared to random mating schemes (Table 1). This increase in genetic gain is probably due to three effects of non-random mating:

1. MC mating connects contributions of unrelated families. Offspring resulting from these matings will have more equal contributions from their ancestors. The latter facilitates their selection, because then selection affects contributions of ancestors less than with random mating. Hence, selection of offspring becomes more independent from that of earlier generations, giving the Optimum Contribution algorithm more opportunity to increase genetic gain without affecting optimum contributions of previous generations.

2. MVRO mating avoids extreme relationships of offspring, which makes the offspring more independent, giving the Optimum Contribution algorithm more opportunity to increase genetic gain within a generation. In a sense it increases the effective number of offspring. The restriction of having only one offspring per full-sib family (factorial mating) also avoids extreme
relationships obtained by mating full sibs.

3. Especially MC mating minimises inbreeding levels of the offspring and thus of the parents of next generation. This mating scheme will therefore achieve larger Mendelian sampling variances, i.e. the term \( \frac{1}{2} (1 - F) \sigma_a^2 \) is larger, resulting in more genetic variance, and thus also in more genetic gain. Since larger Mendelian sampling variances lead to more within family drift, the Optimum Contribution algorithm needs however to restrict the between family drift more in order to achieve the constraint on inbreeding. The latter will decrease genetic gain. Because of these two opposite impacts on genetic gain and because schemes do not differ much in levels of inbreeding, this third effect of non-random mating is probably least important of the three effects.

Although some of the non-random mating schemes are especially designed to achieve one (or two) of the above effects, all non-random mating schemes will achieve all three effects to some extent, because the effects are highly related, e.g. MVRO will also lead to the mating of less related animals. The above effects of non-random mating schemes can be quantified. MC1 mating combines the first and second effect, since it avoids full-sib relationships of the offspring. Hence, the ratio \( \frac{G_{20}(MC) - G_{20}(R)}{G_{20}(MC1) - G_{20}(R)} \) ranges from 0.62 to 1.00 and depicts the proportion of the effect of MC1 that is achieved by minimum coancestry mating. Similarly, the ratio \( \frac{G_{20}(R1) - G_{20}(R)}{G_{20}(MC1) - G_{20}(R)} \) ranges from 0.50 to 0.96 and depicts the proportion of the effect of MC1 that is due to the avoidance of full-sib relationships. Since these two fractions do not add to one, the effects of minimum coancestry mating and avoidance of full-sib relationships are not additive. If minimum coancestry mating is introduced in addition to schemes with the restriction of one offspring per full-sib family, it yields less extra genetic gain compared to when it replaced an R scheme. Similarly, the restriction of one offspring per full-sib family yields less extra genetic gain when introduced in an MC scheme compared to when it is introduced in an R scheme. Because MVRO achieves as much as or less response than MC1, it seems that avoidance of full-sib mating is a sufficient measure to avoid extreme relationships. The fact that MVRO also attempts to avoid other high relationships than full-sibs does not seem to yield extra genetic gain. There are three categories of relationships in a population, namely full-sibs, half-sibs and other less related animals. Full-sib relationships may be avoided at the expense of more half-sib relationships, but half-sib relationships can not be avoided when animals have more than one offspring. The avoidance of strong relationships within the category of other relationships does probably not improve the family structure much, such that avoidance of full-sib matings achieves about all response that is achievable by the second effect of the previous paragraph.

4.2. Effects of the breeding scheme

There was a large difference in how much genetic gain increased between the schemes, mainly depending on the size of the schemes and on the constraint on \( \Delta F \). Benefits of mating strategy were smallest for large schemes and for schemes with high rates of inbreeding. This is probably because there is a relatively higher weight on relationships in the quadratic index (1) when \( \Delta F \) is low and/or the number of candidates is small. This means that an improved family structure has a large effect on genetic gain in schemes with low \( \Delta F \) and/or small numbers of selection candidates.

For larger schemes with 200 selection candidates, the larger families (twice as many animals per generation, but about
the same number of selected animals) resulted in more within family selection, even if selection is for BLUP-EBV. Hence, the weight of the family relationships in the quadratic index (1) is smaller and the structure of selection candidates was less important. Therefore, the effect of mating was also reduced to about the same level as for the schemes with 100 selection candidates and less constrained \(\Delta F\) (2.5%).

4.3. Practical breeding schemes

In most breeding schemes, the control over matings is not as high as we have assumed in the general schemes simulated here. However, in dairy nucleus schemes, ova pick up and in vitro maturation and fertilisation (Kruip et al., 1994) may yield both the required number of offspring per dam and the control over parentage of every offspring.

4.4. Conclusions

For MC, MC1 and MVRO schemes, there was a delay in inbreeding of about two generations compared to the R and R1 schemes, but rates of inbreeding were about the same for all schemes. In breeding schemes where there is selection with a constraint on rates of inbreeding, any improvement in family structure due to non-random mating increases genetic gain. For schemes with \(\Delta F\) constrained to 1.0% and 100 selection candidates, genetic gain was up to 22% higher for the MC1 and MVRO schemes compared to the R scheme. For schemes with a less stringent constraint on \(\Delta F\) or more selection candidates, the superiority of the non-random mating schemes compared to the random schemes was smaller (5-6%). In general, MC1 seemed to be the preferred mating method, since it almost always yielded the highest genetic gain and is simpler and computationally faster than MVRO. R1 and MC increased genetic gains that ranged from 0.50 to 0.96 respectively 0.62 to 1.00 of that of MC1.

References


Woolliams, J.A., 1989. Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and
Appendix

Implementation of the simulated annealing algorithm

Minimum coancestry matings are obtained by applying the simulated annealing algorithm, (Press et al., 1989). In a feasible starting solution, each sire and dam obtains the number offspring that was assigned by the Optimum Contribution selection method (and only one offspring per sire-dam couple for MC1). In order to achieve this, sires and dams were ranked according to their number of offspring and then sires with high rank were mated to dams with low rank.

Given this feasible starting solution, the simulated annealing algorithm optimises the mating scheme as follows:

1. Calculate the sum of coancestry of the mating pairs of start solution and set it to $V_0$. The initial ‘temperature’ is set to $\text{Temp}=1.0$.

2. Consider an alternative set of matings, where the changes of matings are sampled randomly, although the number of progeny per sire and dam must remain constant. Sires 1 and 2 and dam 1 and 2 are randomly chosen, with the restriction that their matings in the current solution are as indicated in Figure A1. Evaluate the alternative value, $V_a = V_0 + \delta$, where $\delta$ is the change of the coancestry if the alternative matings are accepted, i.e.

$$\delta = a_{s1d1} + a_{s2d2} - a_{s1d2} - a_{s2d1},$$

where $a_{s1d2}$ is the coancestry coefficient of sire 1 and dam 2.

When $\delta$ is smaller than zero, replace the current set of matings with the alternative set of matings and set $V_0$ equal to $V_a$. When $\delta$ is larger than zero, replace the current set of matings and $V_0$ with a probability equal to $\exp(-\delta/\text{Temp})$, which decreases when $\delta$ and/or Temp is small.

3. When the current set of matings was replaced $10*\max(N_s,N_d)$ times by an alternative set of matings or when $100*\max(N_s,N_d)$ alternative set of matings have been evaluated, the ‘temperature’, $\text{Temp}$, was decreased by 10%, where $N_s(N_d)$= total number of sires(dams). When there was no accepted alternative set of matings since the last reduction of $\text{Temp}$: finish, because the algorithm did not find an improved distribution and $\text{Temp}$ is too low to accept a reduction of $V_0$. Otherwise go to step 2.

For the MVRO schemes, simulated annealing algorithm is used to minimise the variance of relationship of the offspring that are produced by the matings, i.e.

$$\text{minimise } V_0 = \sum_{i=mating} \sum_{j=mating} V_{ij} \text{ over sets of matings, where } \sum_{i=mating} \text{ denotes summation over all progeny that are produced by the current set of matings and the } V_{ij} \text{ elements are as defined in the main text. The same changes of matings as for MC are evaluated by the simulated annealing algorithm. The alternative values are } V_a = V_0 + \delta, \text{ with}

$$\delta = 2 \sum_{i=mating} \left[ -V_{ij} - V_{i1j} + V_{i1j} + V_{i2j} + V_{i2j} - V_{i1i} \right],$$

where $i_1$ is defined as the progeny that is going to be cancelled with sire $s_1$ and dam $d_1$, $i_2$ is a new progeny by $s_1$ and dam $d_2$, $i_3$ is a new progeny of dam $d_1$ and sire $s_2$ and $i_4$ is the old progeny of $s_2$ and $d_2$, which is cancelled. Except for the above calculation of the $\delta$ values, the simulated annealing algorithm equals that of the MC mating.

For more details on the mechanisms of the annealing algorithm, see Press et al. (1989).
<table>
<thead>
<tr>
<th>Sire</th>
<th>Dam 1</th>
<th>Dam 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 ➞ 0</td>
<td>0 ➞ 1</td>
</tr>
<tr>
<td>Sire 2</td>
<td>0 ➞ 1</td>
<td>1 ➞ 0</td>
</tr>
</tbody>
</table>

Figure A1. Suggested modifications of the current solution for the annealing algorithm. 1 ➞ 0 indicates that the mating conducted in the present solution is not conducted in the alternative solution and 0 ➞ 1 indicates the opposite. All indicated changes have to be made simultaneously to maintain the total number of offspring selected for each sire and dam.