Mate Selection Accounting for Connectedness

R. Carvalheiro[*], B. Kinghorn[†] and S.A. Queiroz[‡]

Introduction

Mate selection refers to selection and mating decisions being performed simultaneously. This one step approach is suitable to accommodate different key issues faced by animal breeders (Kinghorn *et al*., 1999). To implement mate selection an objective function needs to be defined and maximized (Shepherd and Kinghorn, 1999). Kinghorn *et al*. (1999), for example, applied mate selection maximizing a function considering expected merit and inbreeding coefficient of the future progeny and coancestry among selected parents. This strategy, compared with procedures that adopt the selection and mating decisions in separate steps, permitted a higher genetic progress and a better control of the inbreeding.

Degree of connectedness among contemporary groups (CGs) is another key issue that could be considered in mate selection, aiming to increase accuracy of EBVs an their contrasts. Kennedy and Trus (1993) argued that the most appropriate measure of connectedness is the average prediction error variance of differences (PEVd) in EBVs of animals from different groups. This statistic is, however, extremely time consuming and usually not feasible for practical applications. Roso *et al*. (2004) observed that the number of genetic links among CGs is a good, and ease to calculate, measure to predict PEVd.

This study was carried out aiming to develop a program to perform mate selection, and to empirically evaluate the consequences of considering different measures of connectedness in the optimization of the objective functions.

Material and methods

Objective function. The objective functions (OF) that were optimized can be represented in a general form as:

$$OF = b_1x'EBV + b_2F + b_3x'Ax + b_4f(C),$$

where x'EBV is the expected merit of the (future) progeny; F is the predicted progeny mean inbreeding coefficient, for the mating set under consideration; x'Ax is the weighted mean coancestry of selected parents; *f*(C) is a function of connectedness (described latter); b_1 to b_4 are the corresponding weighting factors; and x is the vector of genetic contributions for each

[*] GenSys Associated Consultants, Porto Alegre - RS, 90680-000, Brazil.
[†] University of New England, Armidale – NSW, Australia.
[‡] Sao Paulo State University, Jaboticabal - SP, Brazil.

candidate (to be optimized as well). The connectedness functions considered were based on the number of genetic links among CGs due to common sires. Three functions were investigated: arithmetic mean (A_gl); harmonic mean (H_gl); and the proportion of CGs with more than 500 genetic links (P_gl), which seems to be a threshold to distinguish well and poor connected CGs (Roso *et al.*, 2004).

Mate selection algorithm. An evolutionary algorithm based on differential evolution (DE) (Storn and Price, 1997) was developed in FORTRAN language and applied to optimize the objective functions. The mating sets were defined in the evolutionary process following the problem representation suggested by Gondro and Kinghorn (2008). To enhance the speed and the feasibility of the program, the Colleau (2002) indirect approach to compute coancestry was adopted, and linked lists (Knuth, 1975) were used for the storage and calculations involving sparse matrices.

Data set. A real data set of a Nelore herd was used. The full pedigree contained 4,296 animals, which were born from 1967 to 2007. The proportion of inbreed animals and their average inbreeding coefficient were equal to 33% and 0.037, respectively. Sires and dams considered as candidates for mate selection were those that had progeny in the last crop season (46 sires and 384 dams). It was possible than to compare the realized results with those suggested by the optimization processes.

Table 1 presents the weighting factors and the restrictions put on each component of the objective functions that were optimized. The weightings were empirically defined to evaluate the consequences of considering connectedness functions and to show the pattern of results from varying the emphasis on each component.

Table 1. Weighting factors and restrictions for each objective function (OF1 to OF8)

	OF1	OF2	OF3	OF4	OF5	OF6	OF7	OF8
$\mathbf{b_1}$	0	0	0	0	1	0	0	1
$\mathbf{b_2}$	0	0	0	0	0	-1	0	-0.5
$\mathbf{b_3}$	0	0	0	0	0	0	-1	-30
$\mathbf{b_4}$	1	1	1	-1	0	0	0	0.005
restriction$^{\alpha}$	r1	r2; r3						

$^{\alpha}$ r1: vector of genetic contributions (x) equal to the realized x (i.e. x was not optimized), mimicking a situation where matings would be defined in a second step.
r2: x'EBV greater or equal than realized x'EBV.
r3: F and x'Ax less or equal than realized F and x'Ax, respectively.

For objective functions 3 to 8 (OF3 to OF8), the maximum number of allowed mates per sire was equal to 38 (10% of available dams), and it was assumed that each available dam would produce a fixed number of one offspring, i.e. the female contribution was not optimized in the mate selection process.

Runs of the DE program were carried out using the following sets of operational parameters: 'population size' = 100; 'crossover rate' (CR) = 0.5 (or CR = 0.1 if 'generation' > 2,500); and 'mutation factor' (F) = 0.1 (or F = 1.0 every 4 'generations'). As proposed by Gondro and Kinghorn (2008), proportional and absolute 'mutation' was performed for approximately 3% of the 'loci'. Convergence of the evolutionary process was assumed when the range and the mean absolute deviation of the 'fitness', considering all the 'individuals' of a 'generation' of the DE algorithm, were lower than $1x10^{-6}$.

Results and discussion

Main results of the realized and the optimized matings are shown in Table 2. The realized matings resulted in an expected merit of the future progeny equal to 0.4785 standard deviation units and an average F and coancestry equal to 0.0235 and 0.0321, respectively. Optimization of OF5, OF6 and OF7 revealed, respectively, the maximum expected merit (1.1121) and the minimum inbreeding (0.0060) and coancestry (0.0191) that could be achieved considering the available candidates. Optimization of OF5 corresponds to BLUP selection and it was also used to validate the developed program (the top 10 sires contributed with 38 mates each and the eleventh with 4 mates).

Table 2. Expected merit of the progeny (x'EBV), predicted progeny mean inbreeding coefficient (F), weighted mean coancestry of selected parents (x'Ax), harmonic mean (H_gl), arithmetic mean (A_gl) and proportion of CGs with more than 500 genetic links (P_gl) and number of selected sires (n_sires), according to realized matings or those optimized considering different objective functions$^{\beta}$

	Realized	OF1	OF2	OF3	OF4	OF5	OF6	OF7	OF8
x'EBV	0.4785	0.4785	0.4804	0.2380	0.5848	**1.1121**	0.5773	0.4834	0.9193
F	0.0235	0.0257	0.0227	0.0294	0.0271	0.0249	**0.0060**	0.0130	0.0152
x'Ax	0.0321	0.0321	0.0319	0.0370	0.0377	0.0364	0.0208	**0.0191**	0.0256
H_gl	524	**531**	**633**	**639**	**90**	468	578	558	584
A_gl	734	734	876	890	680	787	798	768	803
P_gl	71	69	80	81	67	78	79	78	79
n_sires	46	46	18	16	21	11	20	27	18

$^{\beta}$ Objective functions (OF1 to OF8) were defined according to the weighting factors and restrictions presented in Table 1.

The optimization of the objective functions OF1 to OF3, which focused on maximizing H_gl, showed (in comparison to realized matings) that: 1) defining matings as a second step (OF1), after fixing genetic contributions, didn't improve connectedness; 2) when mate selection was performed in a single step, under constrained (OF2) and unconstrained (OF3) values of expected merit, inbreeding and coancestry, an expressive improvement of the connectedness functions was observed; 3) it was possible to achieve better connectedness under the same level of expected merit, inbreeding and coancestry (OF2). The same pattern was observed with the maximization of A_gl and P_gl (results not shown).

The application of OF4, as expected, reduced the H_gl. This result was associated with a small reduction of average accuracy of the EBVs, which was equal to 0,68 for OF4 and 0,70 for all the other cases. This association between accuracy and (functions of) genetic links seems to be stronger for datasets with lower degrees of connectedness (Roso *et al.*, 2004).

Some attempts were made to investigate the optimization of other measures of connectedness but they were aborted for being extremely time consuming and not feasible to be applied in the DE developed program.

The empirical weightings used in OF8 resulted, in comparison to the realized matings, in better parameters for all the components considered in the objective function, highlighting the potential of the program to improve genetic progress and to control inbreeding.

It is important to note that no attempt was made in the present study to derive the optimum weighting factors. Gondro and Kinghorn (2008) suggested an iterative procedure to determine the weighting factors during the evolutionary process. According to the authors, this strategy would give great flexibility to learn about the potential outcomes and to optimally balance them, without having to rely on theoretical calculations about what weighting factors to be used *a priori*. This approach was not investigated in the present study because the developed program lacked a suitable interface to do that.

Conclusion

Results show that some connectedness measures can be accommodated in the objective function to perform mate selection, in addition to other usual/relevant components related to genetic merit and inbreeding. Although no great impact was observed in the present study, the improvement of the degree of connectedness is expected to increase accuracy of EBVs and their contrasts, what would ultimately permit a higher genetic progress.

References

Colleau, J. J. (2002) *Genet. Sel. Evol.*, 34:409-421.

Gondro, C. and Kinghorn, B. (2008) Application of Evolutionary Algorithms to solve complex problems in Quantitative Genetics and Bioinformatics. CGIL, Univ. Guelph, Canada (course notes).

Kennedy, B.W. and Trus, D. (1993) *J. Anim. Sci.*, 71:2341–2352.

Kinghorn, B., Shepherd, R. K. and Woolliams, J. A. (1999) In.: *Proc. AAABG*, 13:412-415.

Knuth, D.E. (1975) The art of computer programming, vol.1: fundamental algorithms, 2nd ed. Reading: Addison-Wesley.

Roso, V. M., Schenkel, F. S. and Miller, S. P. (2004) *Can. J. Anim. Sci.*, 84:37-47.

Shepherd, R. K. and Kinghorn, B. (1999) In: *Proc. AAABG*, l3:126-129.

Storn, R., and Price, K. (1997) *J. Global Optimization*, 11:341-359.