

An indirect approach to the extensive calculation of relationship coefficients

Jean-Jacques COLLEAU

Station de génétique quantitative et appliquée,
Institut national de la recherche agronomique,
78352 Jouy-en-Josas Cedex, France

(Received 12 July 2001; accepted 25 January 2002)

Abstract – A method was described for calculating population statistics on relationship coefficients without using corresponding individual data. It relied on the structure of the inverse of the numerator relationship matrix between individuals under investigation and ancestors. Computation times were observed on simulated populations and were compared to those incurred with a conventional direct approach. The indirect approach turned out to be very efficient for multiplying the relationship matrix corresponding to planned matings (full design) by any vector. Efficiency was generally still good or very good for calculating statistics on these simulated populations. An extreme implementation of the method is the calculation of inbreeding coefficients themselves. Relative performances of the indirect method were good except when many full-sibs during many generations existed in the population.

relationship coefficient / inbreeding coefficient / pedigree

1. INTRODUCTION

Selection has been very well known to increase inbreeding and relationship coefficients, which in turn contribute to the decrease in the ultimate rates of genetic gain after many generations. Consequently, many research works have been devoted to defining selection methods efficient for the long term. For instance, procedures have been proposed for maximizing genetic gains with inbreeding rates constrained at desired values [4, 8] or alternatively, minimizing inbreeding rates with constrained selection differentials [10]. These methods are either analytical (*e.g.*, constraint handling through Lagrange multipliers or linear programming [4, 11, 13]) or Monte-Carlo (such as the annealing algorithm [8]) or a combination of both [8]. Furthermore, the current genetic situation of real populations, often with large sizes, as to inbreeding and coancestry coefficients, has to be monitored first to evaluate the importance

of inbreeding and second to assess the practical efficiency of appropriate new selection methods.

These approaches to the management of breeding programmes share the common characteristic that extensive calculations involving matrices of relationship coefficients are needed. Then, the amount of calculation might become critical when the size of the populations involved becomes larger and larger, although sampling might be resorted to, if reasonable accuracy and not full exactness is only required for practical purposes [12]. Efficient methods for calculating inbreeding coefficients do exist. Quaas [6] proposed a method based on the Cholesky decomposition of the numerator relationship matrix according to columns *i.e.*, processing from ancestors to current individuals. Alternatively, Meuwissen and Luo [3] used the Cholesky decomposition by row *i.e.*, processing from current individuals to ancestors. This procedure was shown to be less computationally demanding when continuous updating was required (a situation likely to occur when dynamic optimization procedures are used). Reasons are that computation time increased only linearly with the number of ancestors and that re-calculating inbreeding coefficients of the previous generations is not necessary. Tier [9] first identifies the only relationship coefficients to be finally calculated recursively and stored, using linked list techniques. This method does not spare storage room but can be run faster than Meuwissen and Luo's method if pedigrees includes many generations of ancestors [3].

These methods can be called direct methods because the relationship matrices involved are calculated element by element. The purpose of the present work was to investigate the potential of an indirect method where groups of elements were obtained simultaneously. This method was basically dedicated to optimizing planned matings. However, it might be employed for providing statistics about the relationship coefficients of existing populations and even for calculating individual inbreeding coefficients.

2. AN INDIRECT METHOD FOR CALCULATING RELATIONSHIP STATISTICS ABOUT PLANNED MATINGS

Let us consider matings between n sires (s_i) and m dams (d_j). Then, the overall number of potential matings is nm . For the sake of simplicity, these matings are sorted by sire *i.e.*, so that the mating sequence is

$$s_1d_1, \dots, s_1d_m, \dots, s_nd_1, \dots, s_nd_m.$$

The relationship matrix between the corresponding dummy individuals is A , of size $nm \times nm$. Let x be the vector of size $nm \times 1$ proportional or equal to mating frequencies *i.e.*, $I'x = \text{constant}$. The expected relationship coefficient

after considering any pair of matings is then proportional to $\mathbf{x}'\mathbf{A}\mathbf{x}$. The kernel of this calculation is vector $\mathbf{A}\mathbf{x}$. Analytical optimization for minimizing this expectation requires the use of derivatives *i.e.*, the calculation of $\mathbf{A}\mathbf{x}$. Now, it can be shown that this vector can be obtained without setting \mathbf{A} explicitly.

Let \mathbf{A}_0 of size $n_0 \times n_0$ be the matrix of relationship coefficients involving the sires, the dams and their ancestors till the base population. Let \mathbf{A}_1 be the matrix of relationship coefficients linking this population and the population of planned matings. If we set

$$\begin{aligned}\mathbf{A}\mathbf{x} &= \mathbf{y} \\ \mathbf{A}_1\mathbf{x} &= \mathbf{z}\end{aligned}$$

then we have

$$\begin{pmatrix} \mathbf{A}_0 & \mathbf{A}_1 \\ \mathbf{A}'_1 & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{z} \\ \mathbf{y} \end{pmatrix}.$$

Let \mathbf{A}^* stand for matrix

$$\begin{pmatrix} \mathbf{A}_0 & \mathbf{A}_1 \\ \mathbf{A}'_1 & \mathbf{A} \end{pmatrix}.$$

Then,

$$\mathbf{A}^{*-1} \begin{pmatrix} \mathbf{z} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{x} \end{pmatrix}.$$

As already shown by Henderson [2] and Quaas [6], matrix \mathbf{A}^{*-1} is a sparse matrix with expression

$$(\mathbf{I} - \mathbf{T})' \mathbf{D}^{-1} (\mathbf{I} - \mathbf{T}).$$

For the sake of simplicity, the base population gathers the individuals with both unknown parents and the single unknown parents, after corresponding recodification. Finally, parents precede progeny and each progeny has two known parents. Then, \mathbf{D} is the diagonal matrix with terms equal to 1 for the base population and terms equal to the within-family segregation variance for the other individuals *i.e.*, $0.50 - 0.25 (F_{\text{sire}} + F_{\text{dam}})$. Inbreeding coefficients are assumed to be available. \mathbf{I} is the identity matrix of size $(n_0 + nm) \times (n_0 + nm)$. \mathbf{T} is a null matrix except for two terms, equal to 0.5, for each row corresponding to non-base individuals, linking them to their parents.

Then, the value of \mathbf{y} can be obtained after successively solving two simple linear systems of equations. When solving the system

$$(\mathbf{I} - \mathbf{T})' \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{y}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{x} \end{pmatrix}$$

the information brought by vector \mathbf{x} concerning planned matings is then merged up to the immediate ancestors and processed recursively and collectively up to

the base. Then, this transformed information is processed down from ancestors to planned individuals, after solving the system

$$(\mathbf{I} - \mathbf{T}) \begin{pmatrix} \mathbf{z} \\ \mathbf{y} \end{pmatrix} = \mathbf{D} \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{y}_1 \end{pmatrix}.$$

It can be noticed that there is no way of skipping the calculation of \mathbf{z} , *i.e.*, $\mathbf{A}_1\mathbf{x}$, which is not used later on. The instantaneous storing capacity needed corresponds to only one vector of size $n_0 + nm$. In the first step, \mathbf{x} is overwritten by \mathbf{y}_1 and \mathbf{z}_1 is built progressively only from \mathbf{y}_1 , due to the special form of the right hand side. In the second step, vector

$$\mathbf{D} \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{y}_1 \end{pmatrix}$$

is overwritten downwards by \mathbf{z} and \mathbf{y} .

Quaas [7] and Mrode and Thompson [5] presented a recursive algorithm showing how to compute vector $\mathbf{L}'\mathbf{r}$ where \mathbf{L} is the lower triangular matrix after the Cholesky decomposition of \mathbf{A} (*i.e.*, $\mathbf{A} = \mathbf{L}\mathbf{L}'$). The algorithm presented here might be considered as a kindred algorithm where vector $\mathbf{r} = \begin{pmatrix} \mathbf{0} \\ \mathbf{x} \end{pmatrix}$ and where sparseness of \mathbf{A}^{-1} is exploited as well.

The first benefit from using this algorithm is that matrix \mathbf{A} no longer has to be calculated and stored. Furthermore, computation time can be saved, especially when repetitive evaluation of function $\mathbf{A}\mathbf{x}$ is required, because the amount of calculations is only linear with the overall number of individuals (planned matings + ancestors) and not quadratic as for the direct approach using matrix \mathbf{A} .

A similar approach for obtaining the variance of coefficients a_{ij} [3] would have required to calculate $\text{Tr}\mathbf{A}\mathbf{D}_x\mathbf{A}\mathbf{D}_x$, where \mathbf{D}_x is the diagonal matrix obtained from \mathbf{x} . The trace could be obtained only after setting matrix $\mathbf{A}\mathbf{D}_x$ column by column, which might be very time-consuming.

3. AN INDIRECT METHOD FOR PROVIDING RELATIONSHIP STATISTICS IN REAL POPULATIONS

Let \mathbf{A} be the full relationship matrix for a list of individuals under investigation and their corresponding ancestors. Then, the previous approach can be used in a simpler way, letting

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{y} \\ (\mathbf{I} - \mathbf{T})' \mathbf{y}_1 &= \mathbf{x} \\ (\mathbf{I} - \mathbf{T})\mathbf{y} &= \mathbf{D}\mathbf{y}_1. \end{aligned}$$

Finally, vectors \mathbf{Ax} and quadratics $\mathbf{x}'\mathbf{Ax}$ can be calculated after a number of operations increasing only linearly with n the number of individuals + ancestors.

3.1. Relationship coefficients within a group

Let \mathbf{x} be a sparse vector except for a series of 1's at the positions pertaining to the m members of the group. A single run of function \mathbf{Ax} allows one to obtain the vectors of average relationship coefficients between each member and the whole group (including self-relationships) and the average pairwise relationship coefficients. If vector \mathbf{p} denotes the positions filled in vector \mathbf{x} , then the first vector corresponds to positions \mathbf{p} of $\frac{1}{m}\mathbf{Ax}$ and the scalar corresponds to $\frac{1}{m^2}\mathbf{x}'\mathbf{Ax}$. If self-relationships have to be excluded, corrections are straightforward because these coefficients are equal to 1 + inbreeding coefficients.

3.2. Relationship coefficients between two groups

In some circumstances, knowing the full relationship matrix between a list of males and females is not needed. For instance, breeders might be interested only in the average relationship coefficient between a given sire and all the females of the population. This could be enough for describing the genetic originality of this sire vs. the female population or for modifying selection index to decrease inbreeding rates [1].

Let vector \mathbf{p}_1 denote the positions filled by the first group in sparse vector \mathbf{x}_1 and let vector \mathbf{p}_2 denote the positions filled by the second group in sparse vector \mathbf{x}_2 . Then, positions \mathbf{p}_1 of vector $\frac{1}{m_2}\mathbf{Ax}_2$ correspond to the vector of average relationships between members of group 1 vs. the whole group 2. In the same run, positions \mathbf{p}_2 correspond to the vector of average relationships between members of group 2 vs. the whole group 2. Finally, after a second run where \mathbf{x}_1 and \mathbf{x}_2 are permuted, complete statistics between and within groups can be obtained.

4. AN INDIRECT METHOD FOR CALCULATING INDIVIDUAL INBREEDING COEFFICIENTS

The indirect method can be used for calculating individual inbreeding coefficients, provided that inbreeding coefficients of ancestors are already known and that parents precede progeny according to the sequential identification number. It consists of running function \mathbf{Ax} for each of the different sires involved. Sires are very often much less numerous than dams but sexes might be interchanged if this can save calculation steps. The different \mathbf{x} involved include a single 1 at positions corresponding to the current sires.

For each sire, the terms corresponding to the dams mated are extracted from the resulting vector, divided by two, and affected to the corresponding progeny. Understandably, the efficiency of this approach in comparison with direct methods is likely to depend on the sparseness of the mating design. Substantial computation time can be saved during each back exploration step because many terms of the working vector are still null. The only terms corresponding to the ancestors of the current sire have to be visited. These algorithms can be implemented vectorwise, if the population is split into sections where no pair parent-progeny occurs within sections. This can be carried out very easily if during the extraction of pedigrees, pseudogeneration numbers ψ are calculated (ψ for progeny = 1 + Max (ψ for parents)) and if population is finally sorted according to these numbers. Then, the indirect method can be processed section after section, calculating the full relationship matrix between the parents of the section and then re-affecting the relevant selected relationship coefficients to the individuals of the section. Finally, inbreeding coefficients are equal to half these relationship coefficients.

5. COMPUTATION EFFICIENCY OF THE INDIRECT METHOD

The correctness of the above theory was checked numerically on various complex populations, with overlapping generations, at any times. Direct methods were either the Quaas' method or Meuwissen and Luo's method. The last one was chosen to provide efficiency bench-marks, focusing only on computation times: storing capacity was then considered to be a factor of decreasing influence, with the fast evolution of hardware.

5.1. Populations investigated

For simplifying presentation, discrete generations were assumed: either 10 or 30 (data not shown here and obtained on real populations followed the general pattern shown here and commented afterwards). Each generation, 10 or 50 males and 100 or 200 females were randomly selected and mated. For each of these four situations, family size was allowed to be 2 or 10, with two alternatives: either one sire per dam or the maximum number of sires per dam. Then, overall, 32 random situations were investigated. In order to see whether comparisons might change due to selection and pedigree concentration, a BLUP (animal model) selection was simulated on the populations with 50 sires and 200 dams, based on a trait of initial h^2 equal to 0.5, observable in each sex at any generation.

5.2. General tasks under comparison

Four tasks were investigated on these simplified populations.

Task T_1 : after matings were planned between all the males and all the females of the last generation, the task consisted of multiplying the corresponding relationship matrix A by a vector x .

Task T_2 : in the same context, the task was to calculate the average relationship of each male with all the females.

Task T_3 : the task consisted of calculating the average pairwise relationship coefficient for all the individuals of the last generation

Task T_4 : the task was to calculate the inbreeding coefficients from the base to the last generation.

5.3. Details on computation operations

5.3.1. Task 1

The direct method executed the multiplication of matrix A by a vector x . Then, the calculation time required for setting matrix A itself was not accounted for. The calculation time was assumed to be equal to the square of the matrix size multiplied by a constant corresponding to the time needed for carrying out a basic multiplication plus a basic addition. On a Unix Risc 6000 Workstation, the computer used throughout, this constant was $6 \cdot 10^{-8}$ s CPU.

The indirect method used existing inbreeding coefficients and calculation times were those being obtained in the repetitive uses incurred with optimisation: they corresponded to the time needed for obtaining the solution of the double linear system but did not include the overheads incurred by extracting the relevant ancestors from the whole simulated population and by recodification. The method was implemented in APL2 language, an uncompiled language but endowed with powerful instructions for group operations (here the generation groups), thus reducing the overhead. They were used as often as possible.

5.3.2. Task 2

The direct method was inspired from Meuwissen and Luo's method [3], which used existing inbreeding coefficients. Their central idea was implemented in APL2 language, using a tabular method for back exploration of pedigrees. Individual tables of ancestors and contributions were stored in core only for sire and dams, and obtained from merging those of their parents. Extensive calculations of relationship coefficients at a given generation were carried out from a repetitive use of the stored tables of parents. Computation time was saved when families of full-sibs existed. If n was the number of sires to be mated to m females, then in reality these sires might come from a lower number n^* of families and these dams might come from m^* families. The relationship coefficient between a male and a female of the same family was quite easy to calculate and involved only three inbreeding coefficients (those of parents and that of the family). Then, the final number of relationship coefficients really

needed was even lower than n^*m^* . This final number (observed) was used during the bench-mark so as to calculate the overall computation time. The average computing time per pair under comparison was based on the observed computation time for a sample of the population (50 sires out of the list of males, mated to the whole list of females).

When using the indirect method, the initial overheads (see above) were included. Computation time could not be saved when the full-sib existed because this situation did not affect the size of the mating design considered by the method.

5.3.3. Task 3

In the direct method, the existence of full-sibs was treated as above, reducing the number of pairs of individuals to be compared. The average computation time per pair was the same as for task 2 because it was implemented on animals of the same generation.

5.3.4. Task 4

In the direct method, only one member of each full-sib family in each new generation was investigated (a procedure used by Meuwissen and Luo, as well). This was not carried out in the indirect method for the reason mentioned above.

6. RESULTS AND DISCUSSION

6.1. Task 1

The results are shown in Table I: the relative efficiency is the computation time needed by the indirect method expressed as % vs. the direct method. The absolute computation times for the indirect method are indicated in s CPU and between brackets. Very clearly, the indirect method was far more efficient than the direct calculation because the computation time needed was lower than 4% and even fell down to 0.01%.

The relative efficiency improved when the size of the mating design increased. In the top half of the table, this size was 10 000 or 40 000 while in the bottom half, the increase of size was substantial (up to 250 000 or 1 000 000). In this bottom half, relative computation times were very small, in the range 0.01–0.06%.

As previously mentioned, calculations involved in the indirect method depended linearly on the number of ancestors of the population to be mated while in the direct method, this dependence was quadratic. This basic fact was of an overwhelming influence, despite the remaining overheads incurred with the indirect method. The direct method was superior only for very small mating designs, due to these overheads (data not shown).

Table I. Relative times (%) of indirect vs. direct method for computing Ax (absolute times in s CPU).

Nb dams	Family size	Nb sires	Hierarchical matings		Factorial matings	
			generation number		generation number	
			10	30	10	30
100	2	10	1.6 (0.1)	4.3 (0.3)	1.8 (0.1)	4.1 (0.3)
100	2	50	2.3 (0.1)	6.2 (0.4)	2.5 (0.2)	6.6 (0.4)
200	2	10	0.29 (0.3)	0.86 (0.8)	0.30 (0.3)	0.50 (0.5)
200	2	50	0.34 (0.3)	0.96 (0.9)	0.33 (0.3)	0.73 (0.7)
200S	2	50S	0.32 (0.3)	0.71 (0.7)	0.80 (0.8)	0.70 (0.7)
100	10	10	0.04 (1.6)	0.06 (2.1)	0.04 (1.6)	0.05 (1.9)
100	10	50	0.04 (1.6)	0.06 (2.2)	0.04 (1.5)	0.06 (2.1)
200	10	10	0.01 (5.4)	0.01 (7.3)	0.01 (5.1)	0.01 (7.5)
200	10	50	0.04 (1.6)	0.06 (2.2)	0.04 (1.5)	0.06 (2.1)
200S*	10	50S	0.01 (5.7)	0.01 (7.4)	0.01 (5.1)	0.01 (7.8)

* S population under BLUP selection.

6.2. Task 2

For the sake of simplicity, the four quarters of Table II were named NW, NE, SW, SE according to their geographical positions. Then, in the NW quarter, the size of the mating design was moderate, family size was small and matings were hierarchical. The NE quarter was similar to quarter NW, except that matings were hierarchical. In the SW quarters (SE), the size of the mating design was large, family size was large and matings were hierarchical (factorial).

Except for the SW quarter, the results obtained resembled very much those of Table I because the range of relative computation times was only 0.02–1.4%. The upper values were met in the NW quarter where matings were hierarchical.

The results obtained in the SW quarter differed markedly so that after 30 generations, the indirect method turned out to be less efficient. The absolute computation times for this method were very similar to those obtained in the SE quarter where matings were factorial and where the relative performance of the method was good. Consequently, its disappointing performance in the SW quarter originated from the fact that it could not take profit of the existence of numerous full-sibs.

6.3. Task 3

In comparison with the previous task, performances of the indirect method improved clearly so that it always was superior to the direct one (Tab. III).

Table II. Relative times (%) of indirect vs. direct method for computing average relationships between males and females (absolute times in s CPU).

Nb dams	Family size	Nb sires	Hierarchical matings		Factorial matings	
			generation number		generation number	
			10	30	10	30
100	2	10	1.2 (0.7)	1.5 (2.2)	0.14 (0.7)	0.13 (2.2)
100	2	50	1.4 (0.8)	1.3 (2.7)	0.12 (0.8)	0.10 (2.8)
200	2	10	0.59 (1.3)	0.75 (4.9)	0.06 (1.3)	0.07 (4.9)
200	2	50	0.65 (1.5)	0.55 (5.7)	0.05 (1.5)	0.05 (6.3)
200S	2	50S	0.68 (1.5)	0.55 (5.7)	0.05 (1.5)	0.04 (5.4)
100	10	10	34 (3.3)	146 (19.7)	0.09 (3.2)	0.20 (18.1)
100	10	50	34 (3.3)	140 (18.7)	0.03 (3.3)	0.04 (18.7)
200	10	10	25 (9.1)	107 (57)	0.07 (9.1)	0.14 (57)
200	10	50	26 (9.5)	120 (63)	0.02 (10.1)	0.03 (68)
200S*	10	50S	25 (9.4)	115 (63)	0.01 (9.4)	0.02 (63)

* S population under BLUP selection.

Two reasons might be invoked to explain this change. First, in comparison to Task 2, the absolute computation times decreased for the indirect method

Table III. Relative times (%) of indirect vs. direct method for computing average relationships (absolute times in s CPU).

Nb dams	Family size	Nb sires	Hierarchical matings		Factorial matings	
			generation number		generation number	
			10	30	10	30
100	2	10	0.8 (0.9)	1.0 (3.1)	0.11 (1.2)	0.08 (2.6)
100	2	50	0.93 (1.2)	0.76 (3.2)	0.09 (1.3)	0.06 (3.1)
200	2	10	0.47 (2.1)	0.46 (5.9)	0.04 (1.6)	0.04 (5.4)
200	2	50	0.47 (2.0)	0.30 (6.2)	0.03 (1.7)	0.02 (6.3)
200S	2	50S	0.52 (2.3)	0.37 (6.8)	0.03 (1.7)	0.2 (6.0)
100	10	10	20.7 (4.0)	74 (19.8)	0.06 (3.8)	0.11 (19.6)
100	10	50	19.9 (3.8)	74 (19.6)	0.01 (3.7)	0.02 (19.9)
200	10	10	14.1 (10.2)	58 (61)	0.04 (9.6)	0.08 (60)
200	10	50	13.7 (10.0)	63 (67)	0.01 (10.1)	0.02 (68)
200S*	10	50S	13.5 (10.3)	60 (67)	0.01 (10.0)	0.01 (67)

* S population under BLUP selection.

due to the decrease of the size of the “mating design” because it fell from $0.25(\text{population size})^2$ to only $\text{population size} + 1$. Second, this time, all the possible pairs of families were involved in the direct method. In the previous task instead, some families were represented in males but not in females and *vice-versa*.

6.4. Task 4

Roughly speaking, the relative computation times of the indirect method ranged from 20% to 80% of those of the direct method (Tab. IV). The worst results were obtained in the SW quarter, especially for many generations, and the better ones in the SE quarter. Once again, the bad performance of the SW quarter could be linked to the impossibility to spare time due to the very numerous full-sibs. This was the dominant factor when matings were hierarchical and when the family size was high.

However, when the direct method was facing the necessity of exploring many different individual pedigrees, possibly very long, then the indirect method could exhibit its main advantage. It allowed one first to lower the exploration frequency of pedigrees and second to consider only the overall population pedigrees. It could be observed that in contrast with the other tasks, the number of sires involved influenced the performances very much. This finding was quite logical because calculations should be re-started for each sire.

Table IV. Relative times (%) of indirect vs. direct method for computing inbreeding coefficients (absolute times in s CPU).

Nb dams	Family size	Nb sires	Hierarchical matings		Factorial matings	
			generation number		generation number	
			10	30	10	30
100	2	10	21 (3.1)	30 (35)	22 (3.5)	22 (34)
100	2	50	53 (9.7)	69 (165)	49 (10.8)	53 (193)
200	2	10	19 (5)	20 (51)	16 (4.8)	13 (49)
200	2	50	43 (14)	42 (255)	36 (13.8)	30 (259)
200S	2	50S	43 (13)	42 (237)	36 (13.8)	29 (252)
100	10	10	43 (6.6)	55 (62)	23 (6.5)	17 (63)
100	10	50	67 (12.4)	81 (175)	31 (13.4)	21 (205)
200	10	10	41 (11.5)	43 (106)	22 (11.5)	13 (108)
200	10	50	56 (20)	53 (281)	24 (20)	41 (292)
200S*	10	50S	55 (17.5)	22 (200)	22 (19)	49 (238)

* S population under BLUP selection.

A question concerning tasks 2 to 4 is whether the comparative results might have been changed substantially after using another reference direct method, such as Tier's method. Direct methods are challenged by tasks 2 and 3 because they involve the calculation of many individual relationship coefficients. For these tasks, Tier's method might have been faster than Meuwissen and Luo's method (after splitting the corresponding relationship matrices into smaller ones to keep within the core storage available, as recommended by this author). Meuwissen and Luo [3] found situations where Tier's method ran twice as fast as theirs. However, it should have run additional times as fast as this method for finally cancelling the efficiency gap displayed by the indirect method. For task 4, superiority of the indirect method was substantial but was not a gap. Then, this superiority might be cancelled by more efficient direct methods.

7. CONCLUSION

The numerical investigations presented above showed that the indirect method was efficient not only for heavy calculations on a planned mating design but also for statistical investigations, especially for calculating average relationship coefficients. The reason of this efficiency was that, due to considering sparse inverses of the relationship matrices, numerous inferences could be obtained from a single or two back explorations of the population pedigree. Inbreeding coefficients were assumed to be available. However, the mere indirect method might be resorted to for calculating these coefficients themselves. The efficiency of such a calculation depended highly on the sparseness of the mating design corresponding to existing individuals.

ACKNOWLEDGEMENTS

The author thanks both reviewers for helpful comments.

REFERENCES

- [1] Brisbane J.R., Gibson J.P., Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions, *Theor. Appl. Genet.* (1995) 421–431.
- [2] Henderson C.R., A simple method for computing the inverse of a numerator relationship matrix used in the prediction of breeding values, *Biometrics* 32 (1976) 69–83.
- [3] Meuwissen T.H.E., Luo Z., Computing inbreeding coefficients in large populations, *Genet. Sel. Evol.* 24 (1992) 305–313.
- [4] Meuwissen T.H.E., Maximizing the response of selection with a predefined rate of inbreeding, *J. Anim. Sci.* (1997) 934–940.

- [5] Mrode R., Thompson R., An alternative algorithm for incorporating the relationships between animals in estimating variance components, *J. Anim. Breed.* (1989) 89–95.
- [6] Quaas R.L., Computing the diagonal elements and inverse of a large numerator relationship matrix, *Biometrics* 32 (1976) 949–953.
- [7] Quaas R.L., Transformed mixed model equations: a recursive algorithm to eliminate A^{-1} , *J. Dairy Sci.* 72 (1989) 1937–1941.
- [8] Sonesson A.K., Meuwissen T.H.E., Mating schemes for optimum contribution selection with constrained rates of inbreeding, *Genet. Sel. Evol.* 32 (2000) 231–248.
- [9] Tier B., Computing inbreeding coefficients quickly, *Genet. Sel. Evol.* 22 (1990) 419–430.
- [10] Toro M.A., Nieto B.M., A simple method for increasing the response to artificial selection, *Genet. Res. Camb.* (1984) 347–349.
- [11] Toro M.A., Perez-Enciso M., Optimisation of selection response under restricted inbreeding, *Genet. Sel. Evol.* 22 (1990) 93–107.
- [12] Van Raden P.M., Smith L.A., Selection and mating considering expected inbreeding of future progeny, *J. Dairy. Sci.* (1999) 2771–2778.
- [13] Wray N.R., Goddard M.E., Increasing long-term response to selection, *Genet. Sel. Evol.* 26 (1994) 431–451.