Predicting the annual effective size of livestock populations

LOUIS OLLIVIER[1]* AND JOHN W. JAMES[2]

[1] *INRA-SGQA, 78352 Jouy-en-Josas cedex, France*
[2] *Reprogen, University of Sydney, NSW 2006, Australia*

(*Received 12 February 2004 and in revised form 11 May 2004*)

Summary

Effective population size (N_e) is an important parameter determining the genetic structure of small populations. In natural populations, the number of adults (N) is usually known and N_e can be estimated on the basis of an assumed ratio N_e/N, usually found to be close to 0·5. In farm animal populations, apart from using pedigrees or genetic marker information, N_e can be estimated from the number N of breeding animals, and a value of 1 is commonly assumed for the ratio N_e/N. The purpose of this paper is to show the relation between effective population size and breeding herd size in livestock species. With overlapping generations, N_e can be predicted knowing the number of individuals entering the population per generation and the variance of family size, the latter being directly related to the survival pattern (or replacement policy) in the breeding herd. Assuming an ideal survivorship leading to a geometric age distribution, it can be shown that the number of breeding animals tends to overestimate effective size, particularly in early-maturing species. The ratio of annual effective size to the number of breeding animals is shown to be equal to $[1+(a-1)(1-s)]^2/(1-s^2)$, where a is the age at first offspring and s is the survival rate (including culling) of the parents between successive births. This expression shows to what extent inbreeding may be determined by demography or culling policy independently of the actual herd size. In many situations a fast replacement or an early culling will increase annual effective size. Consequences for the management of small populations are discussed.

1. Introduction

Effective population size is an important parameter in determining the rate of inbreeding and genetic drift in small populations. This concept, initially due to Wright (1931) and first applied to populations with discrete generations, was later extended to populations with overlapping generations, as reviewed by Hill (1972*a*, 1979). In farm animal populations, effective population size may be predicted from the number of breeding animals of both sexes, i.e. breeding herd size. The purpose of this paper is to investigate the relation between effective population size and breeding herd size, and how this relation may be affected by various demographic parameters. This should allow better pedigree-free predictions of inbreeding in the various

farm animal species, and also permit designs for maximizing the genetic variability maintained in small populations.

2. Effective size with overlapping generations

The rule to apply in evaluating effective size (N_e) with overlapping generations, as given by Hill (1972*a*, 1979), is to define N_e as equal to the effective size of a population with discrete generations having the same variance of *lifetime family size* and the same number of *individuals entering the population per generation*. N_e is then expressed as a function of the number of males (M) and females (F) reaching breeding age per unit of time, generation length (L), and variances-covariances of *mm*, *mf*, *fm* and *ff*, which are the numbers of offspring on the paths male to male, male to female, etc. The general

* Corresponding author. Tel: +33 1 34652190. Fax: +33 1 34652210. e-mail: Louis.Ollivier@dga.jouy.inra.fr

expression of N_e is:

$$\frac{1}{N_e} = \frac{1}{16ML}\left[2 + V_{mm} + 2\left(\frac{M}{F}\right)\text{cov}(mm, mf)\right.$$
$$+ \left(\frac{M}{F}\right)^2 V_{mf}\right] + \frac{1}{16FL}\left[2 + \left(\frac{F}{M}\right)^2 V_{fm}\right.$$
$$+ 2\left(\frac{F}{M}\right)\text{cov}(fm, ff) + V_{ff}\right].$$

If family sizes follow a Poisson distribution, the variance of family size equals the mean and the above expression yields $N_e = 4\,MFL/(M+F)$. With overlapping generations, however, the variance of family size exceeds the mean if, as usual, parents are used for different numbers of breeding seasons, as pointed out by James (1982). The effective size depends on the variances and means of the numbers of times (l) the sires and the dams are bred in the herd. Consider *mm*. It has mean of 1 and, assuming equal fecundity at all ages, the mean per breeding season is $1/\bar{l}$ where $\bar{l}$ is the mean number of breeding seasons a sire is used. If l_j is the number of times the *j*th sire is used, its expected number of offspring is $l_j/\bar{l}$. The variance for a given sire, assuming Poisson distributions, equals the mean, and averaging over sires this equals 1. But there is also a between-sire component of $\text{var}(l_j/\bar{l}) = V_l/\bar{l}^2 = v_m$ so that $V_{mm} = 1 + v_m$. By a similar argument $V_{mf} = (F/M) + (F/M)^2 v_m$ and $\text{cov}(mm, mf) = (F/M)v_m$. Similar results apply for the terms in the other square brackets, and when these values are inserted, the following equation results (James, 1982, p. 288):

$$N_e = 4\,MFL/[M(1 + v_f) + F(1 + v_m)], \tag{1}$$

where the *v*s are the standardized variances – or squared coefficients of variation – of *l* in each sex.

Here it is assumed that all animals of the same sex in the herd at any time have the same expected number of progeny. Charlesworth (1980, p. 101) pointed out that his equation (2.39) and the equation of Hill (1972*a*) used above to derive equation (1) are equivalent for the assumed conditions. Equation (1) may thus be seen as a particular case of Charlesworth's equation (2.39) which uses more general demographic parameters and does not depend on the family size assumptions made here.

The values of *v* can easily be obtained knowing the age distribution of the parents. They actually reflect the replacement policy applied, as shown by the numerical example of James (1982, pp. 288–289). If the replacement policies are the same in both sexes, (1) becomes $N_e = 4\,MFL/(M+F)(1+v)$. If *F* is large compared with *M*, (1) may be approximated by $N_e = 4\,ML/(1+v_m)$.

Annual effective size, $N_a = N_e L$, can be derived from the above expressions and values predicted for various livestock species, based on known herd size and replacement policy, as we shall now see.

3. Annual effective size predicted in farm livestock species

Let us assume an ideal replacement rate of the breeding herd leading to a geometric (or Pascal) distribution of the ages, a constant proportion being kept (or surviving) after each breeding season. In such a situation it can be shown that $v = s$, *s* being the survival rate (see for example Subrahmaniam, 1990). Moreover, *L*, *M* and *F* in (1) can be expressed as functions of *s*, *a*, the age at first offspring, and the total (census) number of sires (N_m) and dams (N_f) (see for example Ollivier, 2002). Assuming equal *s* and *a* in both sexes:

$$L = a - 1 + (1-s)^{-1}, \tag{2a}$$

$$M = N_m(1-s), \tag{2b}$$

$$F = N_f(1-s), \tag{2c}$$

and

$$v = s. \tag{2d}$$

The equations (2b) and (2c) are obtained by noting that $N_m = M(1 + s + s^2 + \cdots) = M/(1-s)$ and similarly for N_f.

Hence (1) becomes:

$$N_e = [4\,N_m N_f/(N_m + N_f)]L(1-s)/(1+s).$$

The ratios of effective sizes per generation and per year to herd size, defined as $N = 4\,N_m N_f/(N_m + N_f)$, are then obtained by replacing *L* by its value (2a):

$$N_e/N = [1 + (a-1)(1-s)]/(1+s), \tag{3}$$

and

$$N_a/N = [1 + (a-1)(1-s)]^2/(1-s^2). \tag{4}$$

It should be noted that (4) assumes that both *a* and *s* are expressed on an annual basis. In species reproducing *m* times a year, the time unit will usually be the interval between successive births and equation (4) will have to be divided by *m*.

Table 1 summarizes the values obtained for livestock species in an increasing order of generation length, and for a range of survival rate (*s*) extending from 0·6 to 0·9. It can be seen that effective size (N_e) decreases relatively to herd size as survival rate

Table 1. *Generation* (N_e) *and annual* (N_a) *effective sizes of livestock species relative to herd size* (N), *according to age at first offspring* (a) *and survival* (s) *between successive births. Equal* a *and* s *are assumed in both sexes*

Species	Time unit (for *a* and *s*)	Age at first offspring (*a*)	Generation length[a], L (years)	N_e/N[a]	N_a/N Range[a]	Minimum
Rabbit	2 months	3	0·75–2·00	1·13–0·63	0·84–1·26	0·83
Pig	6 months	2	1·75–5·50	0·88–0·58	1·53–3·48	1·50
Sheep	Year	2	3·50–11·00	0·88–0·58	3·06–6·37	3·00
Cattle	Year	3	4·50–12·00	1·13–0·63	5·06–7·58	5·00
Horse	Year	5	6·50–14·00	1·63–0·74	10·56–10·00	9·00

[a] Range given for *s* varying from 0·6 to 0·9.

increases. A longer generation time is more than offset by a more variable family size. One can also see that with a larger survival rate a given herd size is attained with a smaller annual number of replacements as shown in equations (2b) and (2c).

Equation (3) shows that N_e exceeds N as long as $s < (a-1)/a$. Equation (4) shows that annual effective size (N_a) relative to herd size goes through a minimum when $s = (a-1)/a$, this minimum being $(2a-1)/m$. This also shows that when $a > 1$, annual effective size with overlapping generations always exceeds the value $N_a = aN/m$ prevailing with discrete generations and inspection of equation (4) shows this is also true when $a = 1$.

4. Annual effective size as affected by herd management

(i) *Effect of age at first offspring*

In the previous section we have assumed a fixed age (*a*) at first offspring. In practice, age at first offspring may be manipulated by the breeder. Equation (4) readily shows that annual effective size is an increasing function of *a*. Then, for example in species reproducing annually, a 1 year delay in first offspring would increase the minimum N_a/N by 2 years.

(ii) *Effect of the culling policy*

We have so far assumed a fixed age distribution, determined by a given survival rate, thus excluding voluntary culling. The breeder may change age distribution by applying either *progressive* or *1-stage* culling (James, 1982), among other possible culling policies. The consequences of progressive culling, where a constant proportion *s* of the breeding animals is culled after each breeding season, can be evaluated by using equations (3) and (4), since a geometric age distribution again applies. As mentioned before, equation (4) shows that the 'intermediate' culling rate $1-s = 1/a$ yields the minimum N_a/N. This ratio can be increased either by increasing the culling rate or by doing no culling at all. It should be noted that the

option of maximizing culling rate leads to a limit of $N_a/N = a^2$ (as $s \to 0$). Provided fecundity is sufficient for ensuring such a fast replacement, this option may surpass the alternative of natural replacement if $a \geqslant 3$. In rabbits, for instance, a culling rate of 95% yields a ratio of 1·4, well above the theoretical maximum of 1·26 (assuming $s = 0.9$: see Table 1).

In a 1-stage culling system all breeding animals are discarded after a given number of breeding seasons. The consequences of such a system, under the assumption of a natural geometric age distribution with survival *s*, depend on the properties of the truncated geometric distribution. Equations (2) can be modified to take into account the number of times (*l*) the individuals are bred before culling. If equal *a*, *s* and *l* in both sexes are assumed:

$$L = a - 1 + (1-s)^{-1} - ls^l/(1-s^l), \qquad (5a)$$

$$M = N_m(1-s)/(1-s^l), \qquad (5b)$$

$$F = N_f(1-s)/(1-s^l), \qquad (5c)$$

$$v = [s(1-s)^{-2} - l^2 s^l/(1-s^l)^2]/[(1-s)^{-1} - ls^l/(1-s^l)]^2. \qquad (5d)$$

The ratios (3) and (4) are then more generally expressed as:

$$N_e/N = L(1-s)/(1-s^l)(1+v), \qquad (6)$$

and

$$N_a/N = LN_e/N = L^2(1-s)/(1-s^l)(1+v). \qquad (7)$$

In the ranges of *a* and *s* considered in Table 1, and excluding $l = 1$, which corresponds to discrete generations, the effective size (N_e/N) is a continuously decreasing function of *l*. Annual effective size (N_a/N) follows a different pattern of evolution with increasing *l*, as shown in Fig. 1. In the range $1 < l \leqslant 10$ and excluding the case $a = 2$, the maximum N_a/N is obtained for $l = 2$. The earliest possible 1-stage culling
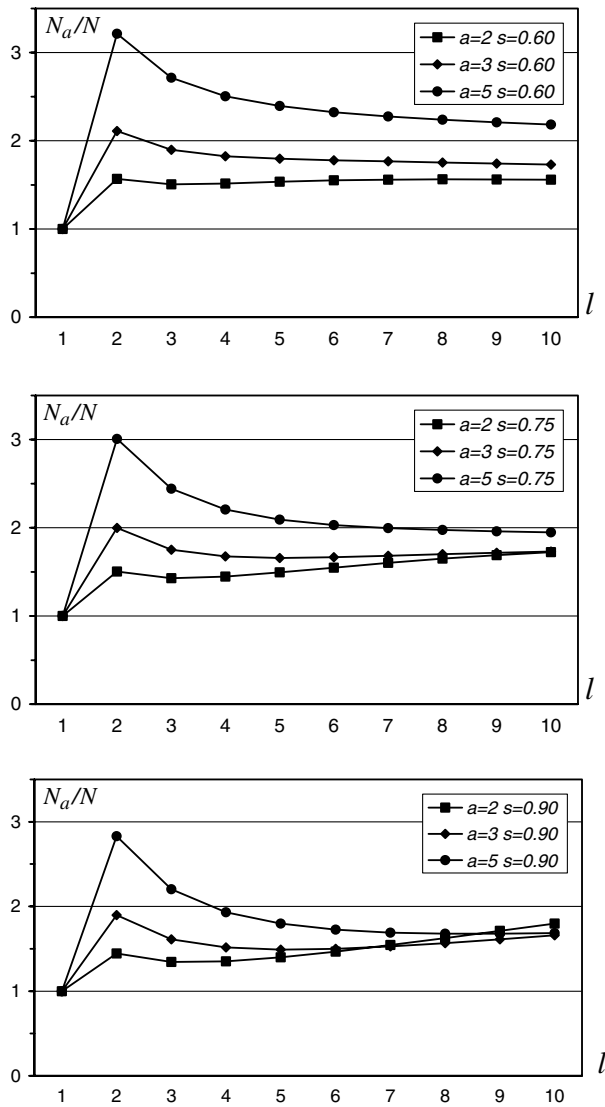
Fig. 1. Ratio of annual effective size to herd size (N_a/N) as a function of the number of times (l) males and females are bred before culling, for various values of age at first offspring (a) and survival (s) between successive breeding seasons, and taking a time unit of 1 year for a and s. Equal a, s and l are assumed in both sexes. N_a/N is expressed in units of a. l varies from 1, corresponding to discrete generations (in which case $N_a/N = a$), to 10.

can therefore usually be recommended. As l increases, N_a/N goes through a minimum occurring for $l \geqslant 3$ and then increases up to a limiting value given by (4). It should be noted, however, that with high values of s the approach to this limit is slow and irregular. In fact the upper limits given in Table 1 imply unrealistically high values of l which will not be reached under practical farm conditions.

5. Discussion and conclusions

The concept of effective size is widely applied in quantitative as well as population genetics. It serves

essentially to evaluate the potential of a population in terms of genetic improvement and to predict the rate of inbreeding ΔF, given the relation $\Delta F = 1/2N_e$. Effective size is also central to conservation genetics, since it is the criterion generally used for defining degree of endangerment of wild as well as domestic animal species or breeds. A very comprehensive review of the methods available for predicting N_e in a variety of situations has been provided by Caballero (1994).

As our results are an extension of the general approach of Hill (1972*a*), the same model applies. Let us recall the main assumptions being made, namely discrete breeding periods, constant number chosen for breeding at each time period, constant parental age distribution, and sampling of genes assumed to be neutral with respect to fitness and any artificial selection pressure. Our results therefore apply only to situations of strict random mating without any selection, as for instance control lines, *in situ* conservation programmes with mild selection, or wild captive species farming.

In natural populations, the number of adults (N) is usually known and N_e can be estimated on the basis of an assumed ratio N_e/N, usually found to be in the range 0·25–0·75 (Nunney, 2000). In farm animals, apart from using pedigree or genetic marker information, N_e can similarly be estimated from the number N of breeding animals. A value of 1 is, for example, implicitly assumed for the ratio N_e/N when the degree of endangerment of livestock breeds is evaluated on the basis of the inbreeding expected over a given period of time (e.g. Simon & Buchenauer, 1993). Our results show that the situation in farm animals differs from that of natural populations, since N_e is generally expected to be closer to the number of adults (Table 1). The higher N_e/N expected in farm populations can be explained by the variations in both life span and fecundity which have to be taken into account in natural populations (Nunney, 1993), whereas only life span variation has to be considered in farm animals. It is indeed worth noting that the general expression of effective size of Nunney (1993), namely equation (5) p. 1331, includes male and female fecundity variations not likely to be important in farm animals. As stated in the derivation of equation (1), our model assumes equal expected number of progeny for animals of the same sex and therefore excludes variation of male seasonal fecundity, connected in particular to mating success, as well as breeding failure of a proportion of females, both factors included (as I_{bm} and α_f respectively) in Nunney's equation (5). It can be seen that when these two parameters are excluded, and $a = 1$ assumed as in Nunney's formula, our equation (1) is obtained.

Table 1, however, shows that large variations around the value $N_e/N = 1$ are to be expected,

depending on age at first offspring (a) and survival rate (s) of the breeding animals. Values close to 0·5 may even be predicted in the extreme case $a = 1$, which is, however, outside the range of species considered in Table 1. Differences in s between species and also between breeds and between sexes are known to occur. The range 0·6–0·9 considered in Table 1 is expected to cover most survivorships commonly encountered in farm animals. Under the assumption of a geometric age distribution, the survival rate s, and consequently the ratios (3) and (4), may alternatively be expressed as functions of a and L using equation (2a). A value of $N_e/N = 0.875$ can then for instance be derived from $a = 5$ and $L = 10.5$ years estimated by Langlois (1976) for thoroughbreds.

The effective sizes of Table 1 should, however, generally be considered as upper limits, because deviations from the simple model assumed will tend to reduce N_e. Selection is a particularly important factor for effective size reduction in farm animals through mechanisms reviewed by Caballero (1994). The joint management of genetic gain and inbreeding in artificial selection programmes has received considerable attention in the literature starting from the basic paper of Robertson (1961). The theory of genetic contributions more recently developed can be used to predict the rate of inbreeding in populations undergoing selection (Woolliams & Thompson, 1994). Using this theory, Bijma *et al.* (2001) showed how the annual rate of inbreeding could be predicted in specific livestock improvement schemes. For given numbers and age classes of the breeding animals, the inbreeding predicted was increased by selection, depending on the selection method, the selection intensity and the heritability of the selected trait. As an example, the breeding scheme of table 2 in Bijma *et al.* (2001) with $M = 20$, $F = 60$, $a = 1$ and 2 age classes, i.e. $l = 2$, is approximately equivalent to $s = 0.7$ and $L = 1.4$ in the least intensive and lowest heritability scheme. In such a scheme our equation (6) predicts $N_e/N = 0.74$ compared with 0·47 with selection.

Our investigation also shows how effective size may be manipulated by varying the herd demography parameters which are controlled by the farmer, namely age at first breeding, survival rate and culling policy, and excluding any pedigree information. As an example, because of the dominating effect of male age distribution in equation (1), N_e/N is expected to exceed 1 when male turnover is high. In the particular situation of a large number of females compared with males and males used only once (i.e. $v_m = 0$), it can readily be seen that N_e/N and N_a/N are approximated by L and L^2 respectively. The general tendency is contrary to the expectation that annual effective size would increase with increasing life span or with increasing culling stage. In fact, the balance between the positive effect of high survival rate on generation length and its negative effect on genetic drift plays in favour of shortening the breeding life span. In many situations, and whenever possible, an early culling may in fact considerably increase annual effective size.

Pedigree information, not considered in this paper, is known to be useful for maintaining constant family sizes and thus decreasing the rate of inbreeding, as investigated for instance in the design of control populations (Hill, 1972b). Smith (1976) showed that even in a balanced breeding structure aiming at equal family sizes, the effects of mortality and fertility are important, and in such situations the annual effective size will be reduced as the survival rate falls, since here a low survival acts against equalizing family size. Mating schemes intended to avoid inbreeding have been intensively investigated, but usually require discrete generations for efficient implementation (see Sanchez *et al.*, 2003 and Fernandez *et al.*, 2003 for various schemes). The genetic contribution theory previously mentioned also allows constraining the rate of inbreeding to a predefined value while optimizing the breeding scheme (Avendano *et al.*, 2004). Finally, it should be recalled that all predictions of effective size assume the populations to be closed. In practice, farm animal populations are usually managed as open breeding systems, which automatically induce increases in effective size, as shown by James (1978).

References

Avendano, S., Woolliams, J. A. & Villanueva, B. (2004). Mendelian sampling terms as a selective advantage in optimum breeding schemes with restriction on the rate of inbreeding. *Genetical Research* **83**, 55–64.

Bijma, P., Van Arendonk, J. A. M. & Woolliams, J. A. (2001). Predicting rates of inbreeding for livestock improvement schemes. *Journal of Animal Science* **79**, 840–853.

Caballero, A. (1994). Developments in the prediction of effective population size. *Heredity* **73**, 674–679.

Charlesworth, B. (1980). *Evolution in Age-Structured Populations*. Cambridge: Cambridge University Press.

Fernandez, J., Toro, M. A. & Caballero, A. (2003). Fixed contributions design vs minimization of global coancestry to control inbreeding in small populations. *Genetics* **165**, 885–894.

Hill, W. G. (1972a). Effective size of populations with overlapping generations. *Theoretical Population Biology* **3**, 278–289.

Hill, W. G. (1972b). Estimation of genetic change. I. General theory and design of control populations. *Animal Breeding Abstracts* **40**, 1–15.

Hill, W. G. (1979). A note on effective population size with overlapping generations. *Genetics* **97**, 317–322.

James, J. W. (1978). Effective population size in open nucleus breeding schemes. *Acta Agriculturae Scandinavica* **28**, 387–392.

James, J. W. (1982). Influences of population structure and size on genetic progress. In *Proceedings of the World Congress on Sheep and Beef Cattle Breeding*, vol. 1: *Technical* (ed. R. A. Barton & W. C. Smith), pp. 283–290. Palmerston North, New Zealand: Dunmore Press.

Langlois, B. (1976). Estimation des quelques paramètres démographiques du Pur Sang Anglais en France. *Annales de Génétique et Sélection animale* **8**, 315–329.

Nunney, L. (1993). Influence of mating system and overlapping generation on effective population size. *Evolution* **47**, 1329–1341.

Nunney, L. (2000). The limit to knowledge in conservation genetics. The value of effective population size. *Evolutionary Biology* **32**, 178–194.

Ollivier, L. (2002). *Eléments de génétique quantitative*, 2nd edn. Paris: INRA.

Robertson, A. (1961). Inbreeding in artificial selection programmes. *Genetical Research* **2**, 189–194.

Sanchez, L., Bijma, P. & Woolliams, J. A. (2003). Minimizing inbreeding by managing genetic contributions across generations. *Genetics* **164**, 1598–1595.

Simon, D. L. & Buchenauer, D. (1993). *Genetic Diversity of European Livestock Breeds*. EAAP publication no. 66. Wageningen: Wageningen Pers.

Smith, C. (1976). Alternative forms of genetic controls. *Animal Production* **23**, 403–412.

Subrahmaniam, K. (1990). *A Primer in Probability*, 2nd edn. New York: Marcel Dekker.

Woolliams, J. A. & Thompson, R. (1994). A theory of genetic contributions. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production* **19**, 127–134.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.