Selection strategies and artificial evolution

J. P. Gibson

Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph, Ontario, N1G 2W1, Canada

Summary. Artificial selection results in biolgical changes, creating artificial evolution. When using selection indexes, the artificial evolution depends on the relative economic (or other) weight of traits in the breeding objective, and on the phenotypic and genetic variances and covariances among these traits and the traits recorded in the selection index. As shown here, the selection strategy (in this case, individual selection versus progeny test selection) can also have marked effects on the kind of artificial evolution produced. Thus, where economic weights are uncertain, choice between alternative selection strategies might take into account the different types of animal or plant resulting.

Key words: Selection indexes – Biological responses

Introduction

Artificial selection of plants or animals results in biological change, which may be termed artificial evolution. To obtain genetic improvement, there are often many alternative ways of collecting and utilizing information and subsequently applying selection, known as selection strategies. The present paper examines the extent to which the strategy of selection affects the outcome in artificial evolution.

In the genetic improvement of a given species, there are usually several traits of varying importance which would benefit from improvement. At the suggestion of Ronald Fisher, Smith (1936) proposed the use of a discriminant function (Fisher 1936) of available information, selection on which would maximize improvement in the aggregate genetic value of the traits of interest. The genetic basis for constucting the linear (discriminant) function of available information, the selection index, was described by Hazel (1943). Various modifications to the

basic principle of the selection index have been proposed, such as restricted indexes, desired gains indexes and nonlinear indexes, which have been variously used in animal and plant breeding (Lin 1978; Baker 1986, reviews).

The principal goal of livestock and plant breeding is improvement of economic efficiency of production. Thus, the relative values (weights) ascribed to the traits of interest are economic weights. The outcome and efficiency of index selection is then variously affected by variation in the economic weights (Vandepitte and Hazel 1977; Smith 1983), estimates of the phenotypic and genetic parameters (Harris 1964; Sales and Hill 1976; Hayes and Hill 1980) and variation in selection intensity due to sampling variation (Razungles 1977). These studies concentrated on the effects on total economic gain rather than the changes in individual traits, artificial evolution. However, it is implicit in the method of constructing selection indexes that variation in economic weights and phenotypic and genetic parameters will affect artificial evolution. Such effects on artificial evolution do not appear to have been systematically quantified in the literature.

The strategy of selection refers to the choice of breeding system, which traits are recorded and on which individuals and at what age individuals are selected. Choice among selection strategies usually depends primarily on the rates of economic genetic gain each is expected to produce. However, if there is concern about biological changes in some traits or if economic weights are uncertain, decision among alternative strategies might be influenced by the nature of the artificial evolution they produce. The scope for different selection strategies resulting in different patterns of artificial evolution has not been previously reported. Results presented here indicate that for some combinations of economic weights and population parameters, altering the selection strategy can have marked effects on artificial evolution.

Theory

Following a standard notation (e.g. Cunningham et al. 1970), the objective of selection is taken to be maximum improvement of an aggregate genotype,

$$T = v' y,$$

where y is an $m \cdot 1$ vector of additive genetic values for m traits in the candidate for selection and v is an $m \cdot 1$ vector of economic weights for these traits. The selection index will be,

$$I = b' x,$$

where x is an $n \cdot 1$ vector of information on the candidate for selection and b is an $n \cdot 1$ vector of coefficients. Maximum gain in T will be achieved by selection on that I which has maximum correlation with T. This occurs when,

$$P b = G v,$$

where P is an $n \cdot n$ matrix of covariances among the elements of x, and G is an $n \cdot m$ matrix of covariances between the elements of x and those of y.

When selection is based on the index, I, the regression, S_i, of the genetic change of the ith trait on the index is given by,

$$S_i = \frac{b' G_i}{b' G v}$$

where, G_i is the ith column of G. For any given index, the artificial evolution produced is described by the S_i's for the traits of interest. The ratio of response of trait i to trait j is,

$$R_{ij} = \frac{S_i}{S_j} = \frac{b' G_i}{b' G_j}. \tag{1}$$

Consider the situation where the aggregate genotype (T) remains constant but an alternative selection strategy is applied, so that the new ratio of responses of traits i and j is,

$$R_{ij}^* = \frac{b^{*\prime} G_i^*}{b^{*\prime} G_j^*}. \tag{2}$$

Two alternative selection strategies will then produce the same relative artificial evolution of traits when,

$$R_{ij} = R_{ij}^* \tag{3}$$

for all combinations of i and j. From the nature of Eqs. (1) and (2), it is clear that equality 3 will hold only for rather restricted circumstances. The most general circumstances are when

$$G = c G^* \tag{4}$$

and

$$b = d b^* \tag{5}$$

where c and d are scalars. Otherwise, equality 3 will hold only for some unique conditions, if at all.

Equality 4 will hold whenever the two selection strategies involve the same traits in x and x^* (i.e. the strategies differ only in the number of relatives and records contributing information), but is unlikely to be true otherwise. Since,

$$b = P^{-1} G v,$$

if equality 4 is true, then equality 5 is true if,

$$P = (c/d) P^*. \tag{6}$$

The elements of P (and P^*) are given by more or less complex formulae, which each have the same form in $\sigma_{P_{ij}}$ and $\sigma_{G_{ij}}$ (phenotypic and genetic covariances between traits i and j) if the same number and type of relatives, each with one record, contribute information on each trait. In this situation, if $\sigma_{G_{ij}}/\sigma_{P_{ij}}$ is the same for all ij, then equality 6 will be true. Put another way, equality 6 is true when the heritabilities of all traits are identical and the genetic correlation equals the phenotypic correlation for all pairs of traits. With multiple records, yet more restrictive conditions could be defined for equality 6 to be true.

The conditions defined in the above paragraph are certainly not the only conditions where artificial evolution will be identical for two alternative selection strategies. For many selection strategies, there will be a variety of combinations of economic weights, phenotypic and genetic parameters which result in identical patterns of evolution. Nevertheless, it is clear that under the vast majority of circumstances, two alternative selection strategies will cause different patterns of artificial evolution.

Examination of differences in artificial evolution

The extent to which different selection strategies produce different patterns of artificial evolution was examined for several simple situations.

Two extreme examples of selection strategies were contrasted. The first was selection of individuals based on a single record of their own performance for each trait. The second was sire selection on the basis of an infinite number of half-sib daughters, equivalent to selecting directly on aggregate genotype.

The absolute rates of change of individual traits were not considered, only the relative changes as measured by the relative response, R_{ij}, defined above. The phenotypic variances of each trait were therefore scaled to equal 1.0, so that additive genetic variances were equivalent to heritabilities, and responses were expressed in phenotypic standard deviations (SD).

The absolute values of economic weights do not affect the correlated responses of traits in the aggregate geno-

type. Therefore, in all cases, economic weights were set in the range 1.0 to -1.0, with at least one trait having an economic weight of 1.0.

All selection index calculations were performed with a modified version of the SELIND computer program (Cunningham and Mahon 1977). All parameter values were checked for statistical validity by testing for positive variance of all principal components of a canonical analysis of the $\cdot G$, P and $GP^{-1}G$ matrices (L. R. Schaeffer, personal communication). Heritabilities, genetic and phenotypic correlations were chosen such that the implied environmental correlations, r_e, were $-1 < r_e < 1$.

In the first instance, only two traits were considered. In the examples presented, R_{ij} was calculated such that trait 1, usually the trait expected to make the greatest response, was the denominator [Eqs. (1) and (2)]. To express the difference in artificial evolution between the two selection strategies, D, the difference between R_{ij} for sire selection and R_{ij} for individual selection was calculated. D is an estimate of the difference between the two selection strategies in the response of trait 2, in SD units, when trait 1 changes by one SD unit in each strategy. To display the results, values of D are plotted against $r_g - r_p$, the difference between the genetic and phenotypic correlations for the two traits. Thic scale was chosen in preference to r_g, because, in some circumstances, $D = 0$ when $r_g - r_p = 0$, providing symmetry in presentation.

Before considering specific situations, the processes involved are illustrated with a simple example. Consider two traits of equal economic importance and phenotypic variability, but trait 1 has a high heritability and trait 2 a low heritability. The responses to index selection based on single records for each trait (mass selection) versus 'a progeny test with a large number of progeny are illustrated in Fig. 1. With mass selection (line M), the additive genetic value of the trait with the highest heritability is estimated with the highest accuracy. The index puts most emphasis on trait 1, and more genetic response is made for trait 1 than trait 2. With the progeny test (line P), the additive genetic value of both traits is estimated with high accuracy, so that nearly equal genetic response is made in both traits. Because of the higher accuracy of the progeny test, more overall genetic change is made, so that line P is longer than line M. The direction of evolutionary change for the two traits is given by the slope of the lines, the difference in the two slopes being D. Note that D is a measure of the difference in the direction of evolutionary genetic change between the alternative strategies of selection, and does not take into account differences in the overall amount, or rate, of genetic change.

The potential size of D was first examined for two traits with equal economic weights, but different heritabilities. Figure 2 plots D against $r_g - r_p$ for the case where $h_1^2 = 0.25$, $h_2^2 = 0.25, 0.20, 0.15, 0.10,$ or $0.05, r_p = 0.5$ and r_g ranges from -1.0 to 1.0 at 0.1 intervals. The horizontal
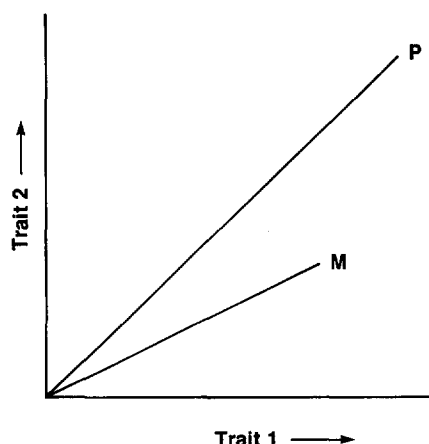

Fig. 1. Response to two-trait index selection for traits of equal value when trait 1 has a high heritability and trait 2 a low heritability, and selection is based on single records (mass selection, line M) or on a progeny test with a large number of progeny (line P)
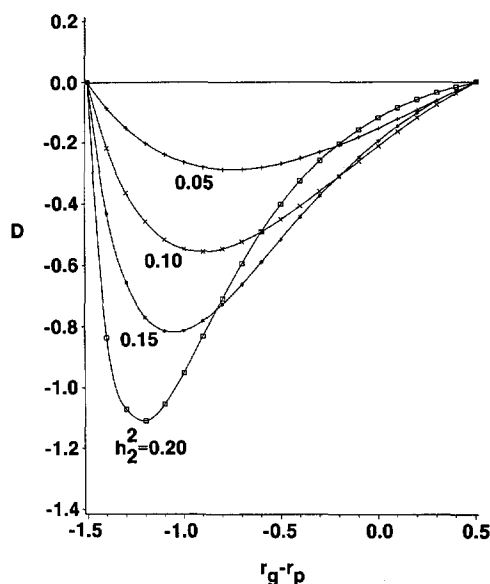

Fig. 2. Plots of D against $r_g - r_p$, for varying heritabilities of the second trait when both traits have the same economic weight in a two-trait selection index. $h_1^2 = 0.25$, $h_2^2 = 0.20, 0.15, 0.10$ and 0.05; $r_p = 0.5$; $v_1 = v_2 = 1.0$

line through $D = 0.0$, corresponds to $h_1^2 = h_2^2 = 0.25$ and also to $h_2^2 = 0.0$. The result that $D = 0.0$, for two traits with equal economic weights and heritabilities, whatever the genetic correlation, is a special case. It applies only when the traits in the index supply equal information about the traits in the aggregate genotype, so that both are selected at equal rates by the index. Had one of the selection strategies involved unequal information in the index, for example by having more relatives contributing
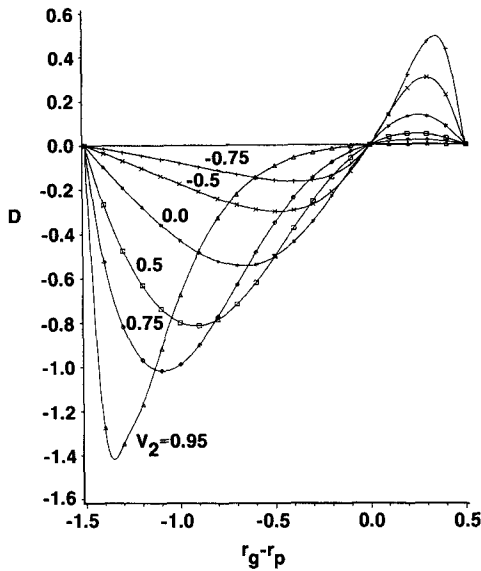
Fig. 3. Plots of D against $r_g - r_p$, for varying economic weights of the second trait, when both traits have the same heritability, in a two-trait selection index. $h_1^2 = h_2^2 = 0.25$; $r_p = 0.5$; $v_1 = 1.0$; $v_2 = 0.95, 0.75, 0.5, 0.0, -0.5$ and -0.75

Fig. 4. Plots of D against $r_g - r_p$, for varying phenotypic correlations, when both traits have the same heritability and the economic weight of the second trait is half that of the first, in a two-trait selection index. $h_1^2 = h_2^2 = 0.25$; $r_p = 0.9, 0.7, 0.5, 0.0, -0.5, -0.7$ and -0.9; $v_1 = 1.0$; $v_2 = 0.5$

information to one trait than to the other, D would have varied with $r_g - r_p$. When $h_2^2 = 0.0$, $D = 0.0$, for all $r_g - r_p$, since no response will occur for a trait with zero heritability. When h_2^2 lies between 0.0 and 0.25, D differs from zero, whenever r_g is not equal ± 1. The greatest values of D occur for large (though not the largest) differences between r_g and r_p, which are probably infrequent in practice. However, even when r_g is close to r_p, D can easily be 0.4 SD or more for modest differences in the heritabilities of the two traits. Similar results (not shown) were found for a wide range of values of h_1^2 and r_p.

The next examples were situations where the two traits had equal heritabilities but different economic weights. Figure 3 plots D against $r_g - r_p$ when $h_1^2 = h_2^2 = 0.25$, $r_p = 0.5$, r_g ranges from -1.0 to 1.0, $v_1 = 1.0$ and $v_2 = 0.95, 0.75, 0.5, 0.0, -0.5$ and -0.75 (v_1 and v_2 being the economic weights of traits 1 and 2). The horizontal line through $D = 0.0$ represents the line for $v_2 = \pm 1.0$. As demonstrated in the theory, when $r_g = r_p$ (i.e. $r_g - r_p = 0.0$), $D = 0.0$. D can be as large as 1.0 SD, but only for large, and biologically unlikely differences between r_g and r_p. However, for modest differences between r_g and r_p, of the order of 0.4, D can be as large as 0.5 SD for a wide range of v_2. Similar results (not shown) were found for a wide range of heritabilities and r_p.

Figure 4 shows D plotted against $r_g - r_p$, for a range of values of r_p, when $h_1^2 = h_2^2 = 0.25$, $v_1 = 1.0$, $v_2 = 0.5$ and r_g ranges from -1.0 to 1.0. The lines differ in length, reflecting the allowable parameter space (values of $r_g - r_p$ outside this space imply r_e lies outside the range -1.0 to 1.0). For all but very high values of r_p, it is possible
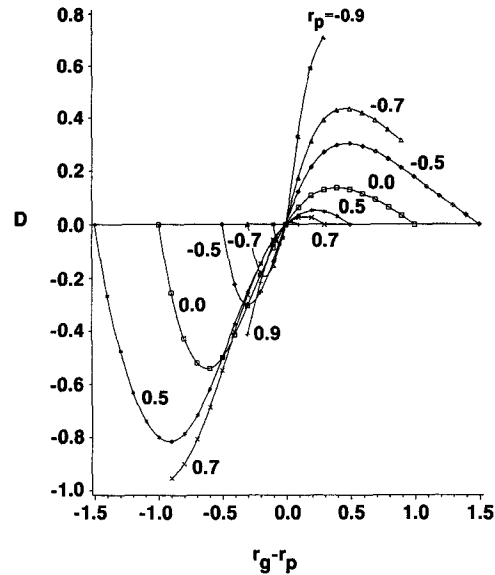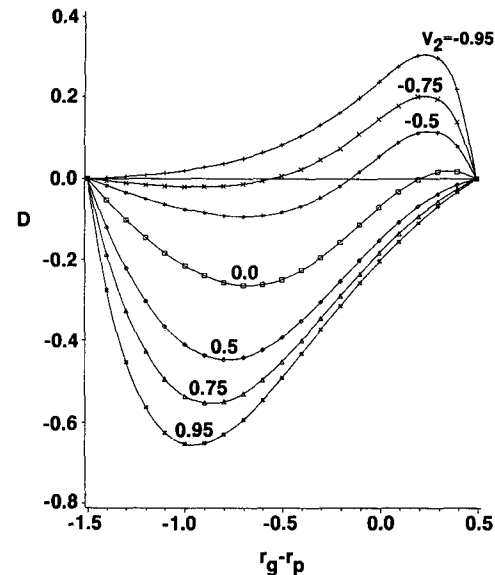
Fig. 5. Plot of D against $r_g - r_p$, for varying economic weights of the second trait, when the heritability of the second trait is half that of the first, in a two-trait selection index. $h_1^2 = 0.25$, $h_2^2 = 0.125$; $r_p = 0.5$; $v_1 = 1.0$; $v_2 = 0.95, 0.75, 0.5, 0.0, -0.5, -0.75$ and -0.95

to get values of D around ± 0.3 SD for biologically plausible differences between r_g and r_p. Again, similar results were found (not shown) for a wide range of heritabilities and differences in economic weights.

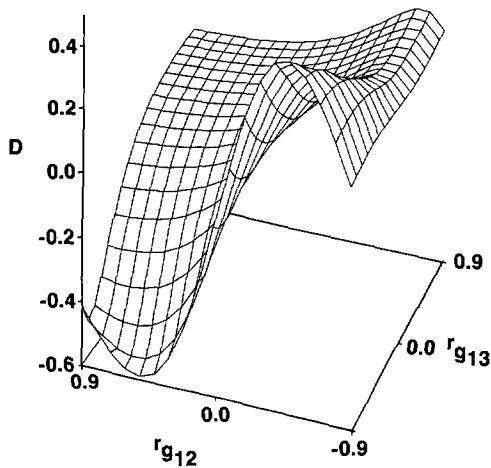Figure 5 shows an example of D plotted against $r_g - r_p$, when both the heritabilities and the economic

Fig. 6. Three dimensional plot of D for trait 3 against r_{g12} and r_{g13}, when all three traits have equal economic weights and heritabilities and all phenotypic correlations are zero, in a three-trait selection index. $h_1^2 = h_2^2 = h_3^2 = 0.25$; r_p's $= 0.0$, $v_1 = v_2 = v_3 = 1.0$

Discussion

The outcome of any selection index is a pattern of evolutionary change which optimizes the increase in the aggregate genotype under selection. Consequently, any alteration of the parameters used in deriving the index can be expected to alter the expected pattern of evolutionary change. A principal assumption in the derivation of an economic selection index is that the economic weights of the traits in the index are correct. This being so, the pattern of change of the traits in the aggregate genotype is of little more than academic interest since it is, by definition, the optimum pattern of change to improve economic merit. Consequently, the extent of this change is rarely emphasized.

There are, however, several arguments for examining artificial evolution a little more closely. In the first place, it is of interest to know where livestock and agricultural plant populations are headed. Predictions of long-term biological changes in our agricultural species could help plan for altered agricultural practices to accommodate such changes. Moreover, in an uncertain future, it may be advantageous to produce biologically diverse breeds, strains or varieties, both to maintain genetic diversity and to allow rapid response to altered conditions (Smith 1985).

In the shorter-term, the accuracy of economic weights can be debated. Uncertainty over economic weights can arise due to fluctuations in true value over time, insufficient data, particularly for product quality traits whose value is difficult to determine objectively, and debate over theoretical frameworks for deriving economic weights (Gibson 1989).

weights of the two traits differ. In the example, $h_1^2 = 0.25$, $h_2^2 = 0.125$, $r_p = 0.5$, r_g ranges from -1.0 to 1.0, $v_1 = 1.0$ and $v_2 = 0.95, 0.75, 0.5, 0.0, -0.5, -0.75$ and -0.95. The largest values of D occur when the second trait has a large positive or negative economic weight. In these cases, values of D of $0.2-0.3$ SD are quite possible. When trait 2 has a lower economic value, D generally takes much smaller values.

The situation where the trait with the highest heritability had the lower economic weight was also examined. In this situation, it is not always clear in advance which trait makes the greatest change under selection. Consequently, R_{ij}, as defined previously, can run to infinity as genetic change in the denominator tends to zero. Graphical display of the results in thus less useful and has been omitted. However, for a range of different parameter values, it was clear that substantial differences in artificial evolution often occurred between the two strategies.

Extension to three or more traits becomes increasingly more complex. An example is provided for a simple situation where three traits have the same heritabilities (0.25) and the same economic weights and all phenotypic correlations are zero. The genetic correlation between traits 1 and 2 (r_{g12}) ranges from -0.9 to 0.9, as does r_{g13}, and r_{g23} equals the product of r_{g12} and r_{g13}. D is calculated separately for traits 2 and 3 (trait 1 is the denominator). In Fig. 6, D for trait 3 is plotted against r_{g12} and r_{g13}. The same general features as for two-trait indexes emerge. The greatest values of D occur for large differences between phenotypic and genetic correlations, with substantial values for D at biologically reasonable differences.

Conclusions

The choice between two selection strategies is usually based on expected rates of genetic gain in economic merit and relative costs. However, account should be taken of the expected artificial evolution if economic weights are uncertain. Also, if two strategies are comparable on economic grounds, choice between the two could be made on the preferred artificial evolution resulting. The preference might be based on uncertainty over the biological consequences of some changes or on perceived long-term adaptability of one biological type over another.

References

Baker RJ (1986) Selection indexes in plant breeding. CRC Press, Boca Raton/FL

Cunningham EP, Mahon GAT (1977) SELIND. A fortran computer program for genetic selection indexes. University of Dublin

Cunningham EP, Moen RA, Gjedrem T (1970) Restriction of selection indexes. Biometrics 26:67–74

Fisher RA (1936) The use of multiple measurements in taxonomic problems. Ann Eugen 7:179–189

Gibson JP (1989) Altering milk composition through genetic selection. J Dairy Sci (in press)

Harris DL (1964) Expected and predicted progress from index selection involving estimates of population parameters. Biometrics 20:46–72

Hayes JF, Hill WG (1980) Modification of estimates of parameters in the construction of genetic selection indices ('bending'). Biometrics 37:483–493

Hazel LN (1943) The genetic basis for constructing selection indexes. Genetics 28:393–399

Lin CY (1978) Index selection for genetic improvement of quantitative characters. Theor Appl Genet 52:49–56

Razungles J (1977) Variation in the phenotypic selection differential. Ann Genet Select Anim 9:105–112

Sales J, Hill WG (1976) Effect of sampling errors on the efficiency of selection indices. 2. Use of information on associated traits for improvement of a single important trait. Anim Prod 23:1–14

Smith C (1983) The effects of changes in economic weights on the efficiency of index selection. J Anim Sci 56:1057–1064

Smith C (1985) The scope for selecting many breeding stocks of possible economic value in the future. Anim Prod 41:403–412

Smith HF (1936) A discriminant function for plant selection. Ann Eugen 7:240–250

Vandepitte WM, Hazel LN (1977) The effect of errors in the economic weights on the accuracy of selection indexes. Ann Genet Sel Anim 9:87–103