Theor Appl Genet (1991) 82:457–465

THEORETICAL AND APPLIED GENETICS

An approximate method for optimum independent culling level selection for n stages of selection with explicit solutions *

W. M. Muir ** and Shizhong Xu

Department of Animal Sciences, Purdue University, West Lafayette, IN 47907, USA

Received April 23, 1990; Accepted January 23, 1991
Communicated by D. Van Vleck

Summary. An approximate method with explicit solutions to apply independent culling levels for multiple traits in n-stages of selection was developed. An approximate solution was found for sequentially selected traits. Two assumptions were necessary. The first was to assume that subsequent selection would not appreciably change the mean of traits already selected, and the second was to approximate the variance of a correlated trait in a selected population with an upward biased projection. The procedure was shown to give near optimal results regardless of selection intensity or genetic correlations if phenotypic correlations among traits were low. The procedure gave poor results only for certain sequences of selection when phenotypic correlations were high. However, in those cases good results were obtained using a different sequence of selection. With high correlations, the procedure is recommended only after comparing solutions and expected genetic gain for all sequences of selection. If the expected aggregate gain for the sequence of selection desired is less than that of another order, culling points associated with the optimal ordering must be determined. Genetic gain from use of culling points is independent of order of selection. The procedure is recommended for use with computer programs that attempt to find optimal culling points to reduce computational time and to check results.

Key words: Artificial selection – Independent culling – Multiple traits – Optimum – Selection methods

Introduction

Economic merit of an animal usually depends on many traits. These traits should be considered simultaneously when a selection program is conducted. Three methods of selection are available for the improvement of multiple traits: index selection, independent culling levels, and tandem selection. Theoretically, index selection is the optimal method for multitrait improvement, with maximum superiority when traits are equally important (Young 1961).

In practice, independent culling levels may be preferred because the independent culling method is more intuitive and convenient to use and also allows animals to be culled as traits become available without waiting until all traits are measured (Hazel and Lush 1942). Independent culling may be applied either simultaneously or sequentially to each trait by use of either truncation points or proportions selected.

If truncation points are used, the procedure is the same for either simultaneous or sequential selection. If proportions are used, the percent saved for each trait with simultaneous and sequential selection refers respectively to the unselected population and the fraction of the population remaining after each culling. Proportions for simultaneous selection can be obtained from tables if the truncation points are known and the traits have a multivariate normal distribution (Falconer 1981). Proportions for sequential selection, however, are difficult to determine since the values depend on correlations among traits. If truncation points and other genetic parameters are known, and the traits are multivariate normal, the CULLTEST program of Saxton (1989) will compute these proportions by numerical integration. The procedure outlined in this paper also gives sequential proportions.

---

* Journal Paper No. 12448 of the Purdue University Agricultural Experiment Station
** To whom correspondence should be addressed

The practical problems with independent culling are (1) how to choose the proper truncation point for each trait to obtained a desired genetic gain, or (2) for a given fixed total proportion selected, how to allocate the selection intensity to each trait to maximize gain in economic merit. Hazel and Lush (1942) used a table which lists different combinations of proportions selected for each trait to find the optimum combination for maximum economic gain. Young and Weiler (1961) and Williams and Weiler (1964) developed charts useful for such intuitive schemes. Smith and Quaas (1982) derived an iterative method to obtain optimum truncation points. Brown (1967), Namkoong (1970) and Cotterill and James (1981) also investigated two-stage independent culling selection for n-traits and developed some tabular solutions. Recently, Saxton (1989) developed a computer program to find optimum culling levels for up to seven stages.

With the advent of Saxton's (1989) program there is little need for tabular procedures for two-stage selection, provided computer availability, which may not always be the case, especially in developing countries. For more than two stages, Saxton (1989) acknowledges that his program requires extensive central processing time even on large mainframe computers, with costs that may be prohibitive. Also, the program is based on a nonlinear optimization procedure that may converge to a local maximum rather than the global maximum.

An approximate method of independent culling level selection for any number of traits, with individuals selected in any number of stages, is described in this paper. This procedure may be used in conjunction with Saxton's (1989) program for either checking results for convergence to a local maximum or could be incorporated into the program to provide a starting point and thus eliminate the need for a grid search.

Theory and notation

Assume that k correlated traits, following a multivariate normal distribution, are to be selected by independent culling levels in k stages. If p_j is the proportion selected based on the j^{th} trait, then the total proportion selected is the product of the proportions $p = p_1 p_2 p_3 \ldots p_k$. Let X_j and μ_j be respectively means of the selected and overall populations for the j^{th} trait. The problem is how best to obtain the proportions to satisfy a breeder's objective.

Define $\Delta \boldsymbol{g}^*$ as the unstandardized genetic gain vector, **A** as the additive genetic variance-covariance matrix, **P** as the phenotypic variance-covariance matrix, and $\boldsymbol{c}$ as the simultaneous selection intensity vector. Let corresponding elements from each matrix be denoted as Δg_i^*, a_{ij}, p_{ij}, and c_i. Also, let **S** be a diagonal matrix with elements $s_{ii} = (p_{ii})^{1/2}$. The following relationships can be expressed as products of these matrices.

(a) The standardized genetic variance-covariance matrix:

$$\mathbf{H} = \mathbf{S}^{-1}\,\mathbf{A}\mathbf{S}^{-1}, \text{ with elements } h_{ij}. \qquad (1)$$

(b) The phenotypic correlation matrix:

$$\mathbf{R} = \mathbf{S}^{-1}\,\mathbf{P}\mathbf{S}^{-1}, \text{ which elements } r_{ij}. \qquad (2)$$

(c) The genetic gain vector in standardized units:

$$\Delta \boldsymbol{g} = \mathbf{S}^{-1}\,\Delta \boldsymbol{g}^*, \text{ with elements } \Delta g_i. \qquad (3)$$

If selection is based on trait m only, then

$$c_j = r_{jm}\, c_m. \quad \text{(Turner and Young 1969)} \qquad (4)$$

Using this relationship, it can be shown that for mass selection based on individual phenotypes, the genetic gain vector in standardized units may be expressed as:

$$\Delta \boldsymbol{g} = \mathbf{H}\mathbf{R}^{-1}\,\boldsymbol{c} \qquad (5)$$

which is an extension of Young and Weiler's (1961) equations (11) and (15). Thus, if the desired genetic gain is given, the simultaneous selection intensity vector can be found by rearrangement of equation (5):

$$\boldsymbol{c} = (\mathbf{H}\mathbf{R}^{-1})^{-1}\,\Delta \boldsymbol{g}. \qquad (5')$$

Unfortunately this equation is not very useful since the simultaneous proportions, and hence truncation points, cannot be easily determined from selection intensities. The tabular association between selection intensity and percent saved given by Falconer (1981)

$$c_j = (X_j - \mu_j)/\sigma_j$$

is based on assumptions that the population from which the observations were sampled was infinite in size and the trait was normally distributed with a mean of μ_j and standard deviation σ_j; thus c_j will have a unit normal distribution with a mean of zero. In contrast, the simultaneous selection intensity for the j^{th} trait is

$$c_{j(k-1)} = [x_{j(k-1)} - \mu_j]/\sigma_j$$

where $x_{j(k-1)}$ is the mean of the j^{th} trait given that $k-1$ other traits were also selected. The selected population has neither a mean of μ_j nor a standard deviation of σ_j; therefore $c_{j(k-1)}$ will have neither a mean of zero nor unit variance.

An approximate solution can be found by assuming that subsequent selection will not appreciably change the mean or variance of traits already selected. This assumption is valid if correlations among traits are small, but will be increasingly transgressed as phenotypic correlations among traits become strong.

If traits are selected in the order 1, 2, ..., k, the progressive standardized selection differentials are defined as the sequential selection intensities. Assuming that subsequent selection will not appreciably change the mean of traits already selected, the sequential selection intensity vector is

$$\boldsymbol{d}=\begin{bmatrix} d_1 \\ d_{2(1)} \\ \vdots \\ d_{k(k-1)} \end{bmatrix}=\begin{bmatrix} (x_1-\mu_1)/\sigma_1 \\ (x_{2(1)}-\mu_{2(1)})/\sigma_{2(1)} \\ \vdots \\ (x_{k(k-1)}-\mu_{k(k-1)})/\sigma_{k(k-1)} \end{bmatrix} \qquad (6)$$

where $\mu_{k(k-1)}$ and $\sigma^2_{k(k-1)}$ are the conditional expectation and variance of trait k given that the previous $k-1$ traits have been selected. The utility of $\boldsymbol{d}$ is that it is in a form for which tabular solutions can easily be found since the $x_{k(k-1)}$ are sampled from a population with a mean of $\mu_{k(k-1)}$ and standard deviation $\sigma_{k(k-1)}$; therefore $d_{k(k-1)}$ has a unit normal distribution with mean zero. However, in order for this result to be useful, the relation between $\boldsymbol{c}$ and $\boldsymbol{d}$ must be determined.

Using the previous assumptions, the simultaneous selection intensity vector reduces to:

$$\boldsymbol{c}=\begin{bmatrix} c_{1(k-1)} \\ c_{2(k-1)} \\ \vdots \\ c_{k(k-1)} \end{bmatrix}=\begin{bmatrix} (x_1-\mu_1)/\sigma_1 \\ (x_{2(1)}-\mu_2)/\sigma_2 \\ \vdots \\ (x_{k(k-1)}-\mu_k)/\sigma_k \end{bmatrix} \qquad (7)$$

The approximate linear relationship between $\boldsymbol{c}$ and $\boldsymbol{d}$ is:

$$\boldsymbol{c} \cong \mathbf{T}'\boldsymbol{d} \qquad (8)$$

where $\mathbf{T}$, with elements t_{ij}, is the upper triangular Cholesky decomposition of matrix $\mathbf{R}$ such that $\mathbf{T}'\mathbf{T}=\mathbf{R}$, i.e.,

$$t_{ji}=\left(r_{ij}-\sum_{k=1}^{j-1} t_{ik}\,t_{jk}\right)\Big/t_{jj} \qquad (j=1,\ldots,i-1)$$

$$t_{ii}=\left(r_{ii}-\sum_{k=1}^{i-1} t^2_{ik}\right)^{1/2}. \qquad \text{(Martin et al. 1965)}$$

The basis for this result is as follows (derivation for more than two traits is tedious, so only the case of two traits is presented, but without loss of generality).

Under the assumption of a multivariate normal distribution

$$\mu_{2(1)}=\mu_2+r_{12}\,\sigma_2(x_1-\mu_1)/\sigma_1 \qquad (9)$$

$$\sigma^2_{2(1)}=\sigma^2_2(1-r^2_{12}\,v) \qquad (10)$$

where $v=c_1(c_1-t)$ and t is the point on the abscissa of the standard normal curve above which a fraction p_1 of the area lies (Cochran 1951; Cotterill and James 1981). Note that v is a positive number between 0 and 1. Thus, the exact relationship between $\sigma^2_{2(1)}$ and σ^2_2 is rather complex, but the following inequality can be assured:

$$\sigma^2_{2(1)} \geq \sigma^2_2(1-r^2_{12}). \qquad (11)$$

Substituting equations (9) and (10) into the corresponding elements of equation (6) gives

$$d_{2(1)}=\frac{x_{2(1)}-[\mu_2+r_{12}\,\sigma_2(x_1-\mu_1)/\sigma_1]}{\sigma_2(1-r^2_{12}\,v)^{1/2}}.$$

From inequality (11)

$$d_{2(1)} \leq \frac{x_{2(1)}-[\mu_2+r_{12}\,\sigma_2(x_1-\mu_1)/\sigma_1]}{\sigma_2(1-r^2_{12})^{1/2}}$$

$$d_{2(1)} \leq \frac{x_{2(1)}-\mu_2}{\sigma_2(1-r^2_{12})^{1/2}}-\frac{r_{12}(x_1-\mu_1)/\sigma_1}{(1-r^2_{12})^{1/2}}$$

or upon substituting elements of equation (7)

$$d_{2(1)} \leq \frac{c_{2(1)}-r_{12}\,c_{1(2)}}{(1-r^2_{12})^{1/2}}. \qquad (12)$$

If c_1 is large, v tends to unity and inequality (12) comes close to equality, otherwise $d_{2(1)}$ will underestimate its true value. In matrix notation, the approximate equation is expressed as

$$\begin{bmatrix} d_1 \\ d_{2(1)} \end{bmatrix} \cong \begin{bmatrix} 1 & 0 \\ -r_{12}/(1-r^2_{12})^{1/2} & 1/(1-r^2_{12})^{1/2} \end{bmatrix}\begin{bmatrix} c_{1(2)} \\ c_{2(1)} \end{bmatrix}$$

Therefore,

$$\begin{bmatrix} c_{1(2)} \\ c_{2(1)} \end{bmatrix} \cong \begin{bmatrix} 1 & 0 \\ -r_{12}/(1-r^2_{12})^{1/2} & 1/(1-r^2_{12})^{1/2} \end{bmatrix}^{-1}\begin{bmatrix} d_1 \\ d_{2(1)} \end{bmatrix}$$

or

$$\boldsymbol{c} \cong \begin{bmatrix} 1 & 0 \\ r_{12} & (1-r^2_{12})^{1/2} \end{bmatrix}\boldsymbol{d} \cong \mathbf{T}'\boldsymbol{d}.$$

where $\mathbf{T}$ is the upper triangular Cholesky decomposition of $\mathbf{R}$.

A. Determining sequential selection intensities for a desired genetic gain vector

Combining (5) and (8)

$$\Delta\boldsymbol{g} \cong \mathbf{H}\mathbf{R}^{-1}\,\mathbf{T}'\boldsymbol{d}. \qquad (13)$$

If $\Delta\boldsymbol{g}$ (a desired genetic gain vector) is given, then

$$\boldsymbol{d} \cong [\mathbf{H}\mathbf{R}^{-1}\mathbf{T}']^{-1}\,\Delta\boldsymbol{g}. \qquad (14)$$

The sequential proportion, $p_{j(j-1)}$, for each trait, given that $j-1$ other traits have already been selected, can then be found for corresponding values of d_j by use of the table given by Falconer (p 316, 1981), or an expanded table can be generated by using the following relationships:

$t = \text{PROBIT}(1-p)$

$y=[1/\sqrt{2\pi}]\exp(-t^2/2)$

$i = y/p$

where p is the percent saved, PROBIT is the inverse of the standard normal distribution function (available in several computer software packages such as SAS 1985), which returns t the point of truncation, y is the ordinate of the standard normal distribution at t, and i is the standardized selection differential.

Consider the following numerical example:

$$\text{Let } \mathbf{S}=\begin{bmatrix}3 & 0 & 0\\ 0 & 5 & 0\\ 0 & 0 & 2.5\end{bmatrix} \qquad \mathbf{R}=\begin{bmatrix}1.0 & 0.4 & 0.5\\ 0.4 & 1.0 & 0.6\\ 0.5 & 0.6 & 1.0\end{bmatrix}$$

$$\mathbf{A}=\begin{bmatrix}5 & 2 & 2\\ 2 & 4 & 3\\ 2 & 3 & 3\end{bmatrix} \quad \text{and} \quad \Delta\boldsymbol{g}=\begin{bmatrix}0.40\\ 0.20\\ 0.40\end{bmatrix}$$

then,

$$\mathbf{H}=\mathbf{S}^{-1}\,\mathbf{A}\mathbf{S}^{-1}=\begin{bmatrix}0.5556 & 0.1333 & 0.2667\\ 0.1333 & 0.1600 & 0.2400\\ 0.2667 & 0.2600 & 0.4800\end{bmatrix}$$

The sequential selection intensity vector, ***d***, is

$$\boldsymbol{d}\cong[\mathbf{H}\mathbf{R}^{-1}\,\mathbf{T}']^{-1}\,\Delta\boldsymbol{g}=\begin{bmatrix}0.7318\\ 0.2579\\ 0.4420\end{bmatrix}.$$

The corresponding sequential proportions are:

$$\begin{bmatrix}p_1\\ p_{2(1)}\\ p_{3(2)}\end{bmatrix}=\begin{bmatrix}.543\\ .860\\ .737\end{bmatrix}$$

and the total proportion selected is

$$p^*=p_1\,p_{2(1)}\,p_{3(2)}=.543\times.860\times.737=.344.$$

B. Determining the sequential selection intensities for a fixed total proportion selected to maximize economic merit

Economic gain is linear combination of genetic gains

$$\Delta\mathrm{E}=\boldsymbol{w}'\,\Delta\boldsymbol{g} \tag{15}$$

where ***w*** is the vector of net economic values. The selection index method is a well-known procedure for solving this problem. Given a fixed proportion selected p, which defines i, the overall selection intensity, the genetic gain vector which gives maximum economic merit gain in standardized units is

$$\Delta\boldsymbol{g}_\mathrm{I}=[\mathrm{i}/(\boldsymbol{b}'\,\mathbf{P}\,\boldsymbol{b})^{1/2}]\,\mathbf{S}^{-1}\,\mathbf{A}\,\boldsymbol{b} \tag{16}$$

where

$$\boldsymbol{b}=\mathbf{P}^{-1}\,\mathbf{A}\,\boldsymbol{w}. \tag{17}$$

Now, $\Delta\boldsymbol{g}_\mathrm{I}$ can be used as the vector of desired genetic gains to solve approximate equation (14). Consequently, a set of sequential proportions $p_1, p_{2(1)}, \ldots, p_{k(k-1)}$ may be obtained from tables generated using the standard normal distribution as previously described. However, p*, the product $p_1\,p_{2(1)}\cdots p_{k(k-1)}$, will be smaller than the desired p because the independent culling procedure is theoretically less efficient than index selection. However, if the relative proportions of genetic gains from independent culling and selection index are fairly constant, then for any value of p the optimum selection intensity vector for independent culling will be some constant times the ***d*** vector computed for index selection. An approximate method of computing this constant is to multiply each sequential proportion $p_{j(j-1)}$ by $\alpha=(p/p^*)^{1/k}$, thus,

$$p^*_{j(j-1)}=\alpha\,p_{j(j-1)}, \text{ which makes}$$

$$p^*_1\,p^*_{2(1)}\cdots p^*_{k(k-1)}=\alpha^k\,p_1\,p_{2(1)}\cdots p_{k(k-1)}=p.$$

The adjusted $p^*_{j(j-1)}$ values will be used for independent culling level selection in the remainder of this paper. This approximation is reasonably good when the $p_{j(j-1)}$'s are between 0.30 and 1.00, since $c_{j(j-1)}$ and $p_{j(j-1)}$ have a nearly linear relationship in that range (Hazel and Lush 1942). Because the adjustment is a constant multiple for each trait, this procedure is not srictly valid and should be viewed as an approximation.

Actual genetic gain expected with this procedure must be determined by numerical integration using the proportions given. However, approximate genetic gain from this procedure ($\Delta\hat{\boldsymbol{g}}$) can be determined by finding the adjusted sequential selection intensity vector (***d****) corresponding to the adjusted proportions and substituting this vector for ***d*** in equation (13):

$$\Delta\hat{\boldsymbol{g}}=\mathbf{H}\mathbf{R}^{-1}\,\mathbf{T}'\,\boldsymbol{d}^*. \tag{18}$$

Equation (5) is a special case where the number of traits measured equals the number of traits to be improved and stages of selection. However, at each stage multiple traits can be selected using an index. If an index is used at any stage, the index is simply treated as a trait.

Consider the following numerical example. For the same set of parameters given in the previous example, we want to maximize ΔE with

$$\boldsymbol{w}=\begin{bmatrix}1\\ 2\\ 1\end{bmatrix}$$

given an overall proportion selected of p = .30, i.e., i = 1.15897.

First, predicting $\Delta\boldsymbol{g}_\mathrm{I}$ by equation (16)

$$\Delta\boldsymbol{g}_\mathrm{I}=[\mathrm{i}/(\boldsymbol{b}'\,\mathbf{P}\,\boldsymbol{b})^{1/2}]\,\mathbf{S}^{-1}\,\mathbf{A}\,\boldsymbol{b}$$

$$=\begin{bmatrix}0.4981\\ 0.2642\\ 0.5357\end{bmatrix}$$

then

$$\boldsymbol{d}=[\mathbf{H}\,\mathbf{R}^{-1}\,\mathbf{T}']^{-1}\,\Delta \boldsymbol{g}_{\mathrm{I}}$$

$$=\begin{bmatrix}0.8996\\0.3034\\0.6648\end{bmatrix}.$$

The corresponding sequential proportions are

$$\begin{bmatrix}\mathrm{p}_1\\\mathrm{p}_{2(1)}\\\mathrm{p}_{3(2)}\end{bmatrix}=\begin{bmatrix}.440\\.830\\.585\end{bmatrix}.$$

Therefore,

$$\mathrm{p}^*=\mathrm{p}_1\,\mathrm{p}_{2(1)}\,\mathrm{p}_{3(2)}=.214,$$

which is smaller than the desired overall proportion of $\mathrm{p}=.30$, so

$$\alpha=(\mathrm{p}/\mathrm{p}^*)^{1/3}=(.30/.216)^{1/3}=1.1192$$

thereby

$$\begin{bmatrix}\mathrm{p}_1^*\\\mathrm{p}_{2(1)}^*\\\mathrm{p}_{3(2)}^*\end{bmatrix}=\alpha\begin{bmatrix}\mathrm{p}_1\\\mathrm{p}_{2(1)}\\\mathrm{p}_{3(2)}\end{bmatrix}=\begin{bmatrix}.4924\\.9289\\.6567\end{bmatrix}.$$

Now, $\mathrm{p}^*=\mathrm{P}_1^*\,\mathrm{p}_{2(1)}^*\,\mathrm{p}_{3(2)}^*=.30=\mathrm{p}$. From tables generated using the standard normal distribution we find

$$\boldsymbol{d}^*=\begin{bmatrix}0.810\\0.145\\0.562\end{bmatrix}.$$

The approximate gain vector $\Delta\hat{\boldsymbol{g}}$ estimated by equation (18) is:

$$\Delta\hat{\boldsymbol{g}}=\begin{bmatrix}0.4594\\0.2169\\0.4499\end{bmatrix}.$$

Evaluation of the approximation

The adequacy of the approximate method was evaluated by numerical multiple integration using the CULLTEST and INDCULL programs of Saxton (1989). The CULLTEST program gives the sequential proportions saved, expected genetic gains, and total economic return for any given set of truncation points. By trial and error, truncation points were established such that the sequential proportions saved for each trait were the same as that saved by the approximate procedure. Genetic gains ($\Delta \boldsymbol{g}$) and aggregate economic improvement (ΔE) from the approximate procedure were compared to those obtained with optimum culling points, determined by use of the INDCULL program with default parameters set by the program. The accuracy of predicting aggregate change in economic value from the approximate method ($\Delta\hat{\mathrm{E}}$), using Equations 15 and 18, was also examined.

From assumptions made when deriving the method, accuracy of the approximation is expected to be influenced by selection intensity, magnitude of correlations, direction of selection, and sequence in which traits are selected. Therefore, for two-stage selection, comparisons were computed for several different cases representing a wide range of genetic and phenotypic correlations, directions of selection, heritabilities, economic weights, order of selection, and total proportion selected.

For three- and four-stage selection, comparisons were limited to the case of equal heritabilities of 0.5 and equal phenotypic and genetic correlations. For three-trait selection, all possible combinations of orders were examined. Not all combinations and selection intensities were examined with four-trait selection since the computer time required was 18 CPU min per combination on an IBM 3090. Five-trait selection was not examined due to the tremendous amount of computer CPU time required per combination (in excess of 10 CPU h). Comparisons were based on expected change in aggregate economic value (ΔE).

Expected selection responses for two-stage selection using the approximate and optimum methods are in Table 1. In some cases the computed relative efficiency of the approximation was greater than 1. Obviously this result is not possible. Saxton (1989) notes in comments sent with his program that the approach is based on a Newton-Raphson maximization procedure and that results are dependent on a good starting point. Although the program evaluated a grid of possible starting points, it sometimes failed to converge to the global maximum. Thus, the solution given by the program was not always optimal. The approximation gave excellent results in all cases where the phenotypic correlation was weak.

With a strong phenotypic correlation, the merit of the approximation was dependent on the order of selection, in combination with the heritabilities and economic weights. The approximation was generally poor in those cases in which the first trait selected had both a high economic weight and high heritability.

Expected selection responses for three-trait selection using the approximate and optimum methods are given in Table 2. Again, some cases were observed in which the relative efficiency of the approximation was estimated to be greater than 1. Generally, the approximation gave near optimum results for weak correlations. From the examination of two-stage selection and the assumptions from which the theory was based, these results are also expected for strong genetic correlations if the phenotypic correlations are low.

Table 1. Expected selection response (Δg) and total economic gain (ΔE) for two traits using approximate and exact independent culling procedures for differing orders of selection, proportions selected (p), economic weights (w), genetic (r_g), and phenotypic (r_p) correlations

Case	p	r_p	r_g	Stage	h^2	w	Independent culling[a]						Relative efficiency[b] ($100 \times \Delta E^a / \Delta E^0$)
							Approximate			Optimum			
							$p_{j(j-1)}$	t	Δg	$p_{j(j-1)}$	t	Δg	
1	.01	.2	.2	1	.2	.2	85.7	−1.07	.26	100.0	−4.00	.24	
				2	1	1	1.2	2.31	2.64	1.0	2.32	2.66	
									2.69			2.72	98.9
2	.01	.2	.2	2	1	1	1.1	2.30	2.63	1.0	2.32	2.66	
				1	.2	.2	94.8	−.94	.26	100.0	−1.99	.24	
									2.68			2.71	98.9
3	.01	.2	.2	1	.2	1	13.0	1.11	.43	15.8	1.00	.42	
				2	1	.2	7.7	1.74	2.04	6.3	1.81	2.10	
									.84			.84	100.0
4	.01	.2	.2	2	1	.2	3.3	1.84	2.13	2.2	2.00	2.29	
				1	.2	1	30.0	.94	.42	45.4	.62	.38	
									.84			.83	101.1
5	.01	.2	.8	1	.2	.2	48.2	.04	.98	84.1	−1.00	.96	
				2	1	1	2.1	2.17	2.63	1.2	2.30	2.67	
									2.82			2.72	103.7
6	.01	.2	.8	2	1	1	1.1	2.30	2.67	1.0	2.32	2.66	
				1	.2	.2	79.4	−.94	.96	100.0	−1.99	.95	
									2.86			2.85	103.3
7	.01	.2	.8	1	.2	1	33.0	.44	.98	50.0	0.00	.98	
				2	1	.2	3.0	2.07	2.57	2.0	2.19	2.63	
									1.49			1.50	99.3
8	.01	.2	.8	2	1	.2	1.6	2.14	2.61	1.0	2.32	2.66	
				1	.2	1	61.5	.20	.98	100.0	−1.99	.95	
									1.50			1.48	101.4
9	.50	.2	.2	1	.2	.2	97.2	−1.91	.08	100.0	−5.60	.07	
				2	1	1	51.4	−.02	.78	50.0	0.00	.80	
									.80			.81	98.9
10	.50	.2	.2	2	1	1	51.2	−.03	.78	50.0	0.00	.80	
				1	.2	.2	97.7	−1.81	.08	100.0	−5.56	.07	
									.79			.81	97.5
11	.50	.2	.2	1	.2	1	73.3	−.63	.13	82.5	−.94	.11	
				2	1	.2	67.8	−.37	.57	60.6	−.20	.66	
									.24			.24	100.0
12	.50	.2	.2	2	1	.2	60.5	−.27	.63	50.0	0.00	.80	
				1	.2	1	82.6	−.80	.12	100.0	−5.61	.07	
									.24			.23	104.4
13	.50	.2	.8	1	.2	1	83.7	−.98	.28	97.9	−2.02	.29	
				2	1	.2	59.7	−.19	.74	51.1	−.02	.80	
									.43			.44	97.7
14	.50	.2	.8	2	1	.2	54.3	−.11	.77	50.0	0.00	.80	
				1	.2	1	92.1	−1.25	.29	100.0	−2.38	.28	
									.44			.44	100.0
15	.50	.2	.8	1	.2	.2	88.6	−1.21	.29	100.0	−3.19	.28	
				2	1	1	56.5	.12	.76	50.0	0.00	.80	
									.82			.85	96.5
16	.50	.2	.8	2	1	1	52.1	−.05	.79	50.0	0.00	.80	
				1	.2	.2	95.9	−1.57	.29	100.0	−3.19	.28	
									.84			.86	97.7

Table 1. (Continued)

Case	p	r_p	r_g	Stage	h^2	w	Independent culling[a]						Relative efficiency[b] (100 × $\Delta E^a/\Delta E^0$)
							Approximate			Optimum			
							$p_{j(j-1)}$	t	Δg	$p_{j(j-1)}$	t	Δg	
17	.01	.8	.2	1	.2	.2	93.9	−1.55	.24	100.0	−4.00	.24	
				2	1	1	1.1	2.66	2.66	1.0	2.32	2.66	
									2.71			2.71	100.0
18	.01	.8	.2	2	1	1	12.6	1.15	.44	2.6	2.32	2.65	
				1	.2	.2	7.9	2.30	.52	38.5	.50	.24	
									.54			2.70	20.0
19	.01	.8	.2	1	.2	1	7.6	1.43	.30	100.0	−4.00	.24	
				2	1	.2	13.1	2.25	2.30	1.0	2.32	2.66	
									.76			.77	98.7
20	.01	.8	.2	2	1	.2	1.1	2.30	2.49	2.6	2.32	2.65	
				1	.2	1	94.4	1.15	.27	38.5	.50	.24	
									.76			.77	98.7
21	.01	.8	.8	1	.2	.2	59.5	−.24	.95	100.0	−4.00	.95	
				2	1	1	1.7	2.33	2.66	1.0	2.32	2.66	
									2.86			2.86	100.0
22	.01	.8	.8	2	1	1	5.0	1.63	1.43	2.6	2.32	2.65	
				1	.2	.2	19.9	2.21	.66	38.5	.50	.95	
									1.56			2.84	54.9
23	.01	.8	.8	1	.2	1	38.0	.31	.95	100.0	−4.00	.95	
				2	1	.2	2.6	2.32	2.66	1.0	2.32	2.66	
									1.48			1.48	100.0
24	.01	.8	.8	2	1	.2	2.9	1.89	1.72	2.6	2.32	2.65	
				1	.2	1	34.1	2.08	.73	38.5	.50	.95	
									1.08			1.48	73.0
25	.50	.8	.2	1	.2	.2	98.4	−2.14	.07	100.0	−4.00	.07	
				2	1	1	50.8	−0.01	.80	50.0	0.00	.80	
									.81			.81	100.0
26	.50	.8	.2	2	1	1	73.3	−.63	.21	50.0	0.00	.80	
				1	.2	.2	68.4	−.06	.14	100.0	−3.55	.07	
									.24			.81	29.6
27	.50	.8	.2	1	.2	1	68.1	−.47	.10	100.0	−4.00	.07	
				2	1	.2	73.4	−.11	.61	50.0	0.00	.80	
									.22			.23	95.6
28	.50	.8	.2	2	1	.2	50.8	−.02	.75	50.0	0.00	.80	
				1	.2	1	98.5	−.92	.08	100.0	−3.16	.07	
									.23			.23	100.0
29	.50	.8	.8	1	.2	.2	91.2	−1.35	.28	100.0	−4.00	.28	
				2	1	1	54.8	0.00	.79	50.0	0.00	.80	
									.86			.86	100.0
30	.50	.8	.8	2	1	1	64.3	−.37	.49	50.0	0.00	.80	
				1	.2	.2	77.8	−.16	.21	100.0	−3.44	.28	
									.54			.86	62.8
31	.50	.8	.8	1	.2	1	85.6	−1.06	.28	100.0	−4.00	.28	
				2	1	.2	58.4	−.01	.78	50.0	0.00	.80	
									.44			.44	100.0
32	.50	.8	.8	2	1	.2	59.5	−.24	.57	50.0	0.00	.80	
				1	.2	1	84.0	−.26	.23	100.0	−3.31	.28	
									.34			.44	77.3

---

[a] $p_{j(j-1)}$, Sequential proportion; t, truncation point

[b] ΔE^a and ΔE^0 are the expected economic gains from truncation points set respectively by the approximate and INDCULL procedures; $\Delta E = w'\Delta g$,

Table 2. Expected total economic gain (ΔE) for three traits using approximate and exact independent culling procedures for differing proportions selected (p), genetic (phenotypic) correlations (r), and economic weights (w)

Case	r	p	Economic weight			Culling procedure			Relative efficiency ($100 \times \Delta E^a/\Delta E^0$)
			Stage			Approximate		Optimum	
			1	2	3	ΔE^a	$\Delta \hat{E}$	ΔE^0	
1	.8	.01	1	1	1	3.64	3.50	3.61	100.7
2	.8	.01	1	1	−1	1.34	1.44	1.39	96.6
3	.8	.01	1	−1	1	1.38	1.44	1.39	99.4
4	.8	.01	−1	1	1	1.13	1.40	1.39	81.4
5	.8	.10	1	1	1	2.38	2.27	2.37	100.4
6	.8	.10	1	1	−1	.81	.92	.87	93.1
7	.8	.50	1	1	1	1.08	1.00	1.05	102.9
8	.8	.50	1	1	−1	.31	.40	.39	74.3
9	.8	.50	1	−1	1	.36	.40	.39	92.3
10	.8	.50	−1	1	1	.24	.38	.39	61.5
11	.2	.01	1	1	1	2.48	2.42	2.50	99.2
12	.2	.01	1	1	−1	1.87	1.90	1.87	99.8
13	.2	.01	1	−1	1	1.87	1.90	1.87	100.0
14	.2	.01	−1	1	1	1.87	1.92	1.87	99.8
15	.2	.10	1	1	1	1.60	1.53	1.61	99.8
16	.2	.10	1	1	−1	1.16	1.20	1.17	99.4
17	.2	.50	1	1	1	.70	.64	.70	99.7
18	.2	.50	1	1	−1	.48	.50	.49	97.9
19	.2	.50	1	−1	1	.48	.50	.49	98.7
20	.2	.50	−1	1	1	.49	.52	.49	99.9

[a] $\Delta \hat{E}$, Approximate gain estimated using Eq. 18; ΔE^a and ΔE^0 are the expected economic gains from truncation points set respectively by the approximate and INDCULL procedures

Table 3. Expected total economic gain (ΔE) for four traits using approximate and exact independent culling procedures for differing proportions selected (p), genetic (phenotypic) correlations (r), and economic weights (w)

Case	r	p	Economic weight				Culling procedure			Relative efficiency ($100 \times \Delta E^a/\Delta E^0$)
			Trait				Approximate		Optimum	
			1	2	3	4	ΔE^a	$\Delta \hat{E}$	ΔE^0	
1	.8	.1	1	1	1	1	3.13	2.96	3.12	100.3
2	.8	.1	1	1	1	−1	1.53	1.57	1.66	91.9
3	.8	.1	1	1	−1	1	1.59	1.57	1.66	95.7
4	.8	.1	1	−1	1	1	1.64	1.59	1.66	98.9
5	.8	.1	−1	1	1	1	.88	1.47	1.66	52.8
6	.8	.1	1	1	−1	−1	.32	.69	.43	75.0
7	.8	.1	1	−1	1	−1	.39	.65	.43	90.3
8	.8	.1	−1	1	1	−1	.40	.65	.43	92.7
9	.8	.1	1	−1	−1	1	.40	.65	.43	92.7
10	.8	.1	−1	1	−1	1	.39	.65	.43	90.3
11	.8	.1	−1	−1	1	1	.32	.69	.43	75.0
12	.8	.5	1	1	1	1	1.41	1.31	1.36	104.2
13	.8	.5	1	1	1	−1	.65	.69	.75	86.9
14	.8	.5	1	1	−1	−1	.09	.30	.13	72.0
15	.2	.1	1	1	−1	−1	1.19	1.27	1.20	99.0
16	.2	.5	1	1	−1	−1	.47	.53	.48	99.5

[a] $\Delta \hat{E}$, Approximate gain estimated using Eq. 18; ΔE^a and ΔE^0 are the expected economic gains from truncation points set respectively by the approximate and INDCULL procedures

The worst efficiencies, Cases 4, 8, and 10, occurred with combinations of strong correlations and certain orders of antagonistic selection. With the same genetic parameters and other orders of selection, near optimal results were obtained.

Since the approximation appeared to work well for two- and three-trait selection with low correlations, further comparisons using four traits were mainly restricted to cases of high correlations. However, some combinations which gave poor relative efficiencies with high correlations were further examined with low correlations. Expected selection responses for four-trait selection using the approximate and exact methods are in Table 3. Two cases with computed relative efficiencies greater than 1 occurred with high correlations and non-antagonistic selection. In general, the same results as with two- and three-stage selection were observed. With weak correlations, the approximation gives excellent results; with high correlations, the order of selection is important. However, for a given set of genetic parameters, an order could usually be found which gave near optimal results.

Also, equation (18) overestimated expected genetic gain if selection was antagonistic and underestimated it otherwise. Relative bias increased as the magnitude of correlations increased or as selection intensity decreased.

From these results some general conclusions can be made. With weak phenotypic correlations among traits, the approximate method gives excellent results for any number of stages, selection intensities and economic weights. With strong phenotypic correlations, the relative efficiency of the approximation is dependent on order of selection. Thus, with high correlations, the procedure is only recommended after comparing the solutions and expected genetic gain for all orders of selection. If the solution for the order of selection desired is less than that of another order, culling points associated with the optimal ordering must be determined. Genetic gain from the use of culling points is independent of order of selection.

Discussion

The main advantage of the method presented is that it provides a general formula for any number of stages of selection with explicit solutions. A general formula for more than two-stages of selection was not previously available. If a computer is not available, this approximation can be used with confidence for all cases in which the phenotypic correlations among traits are small.

If a computer is available, the approximation may be used to establish starting points for Saxton's (1989) program. His program uses a grid search of all possible independent culling truncations, with the best of these used to start the search for the combination giving optimum gain. Saxton notes that the grid search generally is responsible for most of the computation time. The approximation can determine a good starting point and thereby eliminate the grid search. Also, Saxton's program gives expected gain from index selection as an upper limit so that obviously invalid results can be excluded. The approximation can also establish a useful lower bound. There were cases in Tables 1, 2, and 3 in which the results produced by Saxton's program appeared valid when compared to response expected with selection index. However, those cases were determined to be local maxima when compared to response expected with the approximation.

References

Brown GH (1967) The use of correlated variables for preliminary culling. Biometrics 23:551–562

Cochran WG (1951) Improvement by means of selection. In: Neyman J (ed) Proc 2nd Berkeley Symp on Math, Stat and Probability. Berkeley, Calif., pp 449–470

Cotterill PP, James JW (1981) Optimizing two-stage independent culling selection in tree and animal breeding. Theor Appl Genet 59:67–72

Falconer DS (1981) Introduction to quantitative genetics, 2nd ed. Longman, London

Hazel LN, Lush JL (1942) The efficiency of three methods of selection. J Hered 33:393–399

Martin RS, Peters G, Wilkinson JH (1965) Symmetric decomposition of a positive definite matrix. Num Math 7:362–383

Namkoong G (1970) Optimum allocation of selection intensity in two stages of truncation selection. Biometrics 26:465–476

SAS Institute Inc (1985) SAS user's guide: basics, version 5 edn. SAS Institute Inc, Cary, N.C.

Saxton AM (1989) INDCULL version 3.0: Independent culling for two or more traits. J Hered 80:166–167

Smith SP, Quaas RL (1982) Optimal truncation points for independent culling-level selection involving two traits. Biometrics 38:975–980

Turner HN, Young SSY (1969) Quantitative genetics in sheep breeding. Cornell University Press, Ithaca, N.Y.

Williams JM, Weiler H (1964) Further charts for the means of truncated normal bivariate distributions. Aust J Stat 6(3), 117–129.

Young SSY (1961) A further examination of the relative efficiency of three methods of selection for genetic gains under less restricted conditions. Genet Res 2:106–121

Young SSY, Weiler H (1961) Selection for two correlated traits by independent culling levels. J Genet 57:329–338