

Selection for improved yield in inter-specific mixtures or intercrops

A. J. Wright*

Plant Breeding Institute, Maris Lane, Trumpington, Cambridge CB2 2LQ, England

Received May 30, 1984; Accepted July 21, 1984

Communicated by P. M. A. Tigerstedt

Summary. A simple recurrent selection scheme using randomly constructed intercrop or mixture treatments for the mutual improvement of two or more species is described. Several possible selection criteria are available for each species, including its own yield, those of each of its associates, and combinations of these. Statistical expressions are developed to describe the expected gain in economic yield in each species given selection for one or more species on the basis of any common criterion, and it is shown that the gains from all responding species and applied selection criteria are additive. Negative correlations between the direct and associate effects of genotypes favour the selection of whole mixtures. Selection indices of several species yields can be applied either on an individual species basis or to whole mixtures, but only in the latter case can the statistics required for the calculation of the optimum index be estimated from the data provided by the trial. Some general properties and possible long term effects of different selection methods are discussed.

Key words: Competition – Inter-specific mixture – Intercrop – Selection

1 Introduction

In spite of the widespread use of multiple cropping and other mixed culture systems throughout the world, research in this area is far less advanced than comparable work in pure crop culture. This is due in part to the

wide use of pure cultures in the developed world, in part to the relative lack of resources in the developing world, but not least to the complexity of problems involved. Interactions between species add another dimension to the agronomic problems of optimisation of culture and management techniques. Not surprisingly, relatively little work has been done on the design of breeding systems for the production of varieties for use in specific multiple cropping situations, still less to test or routinely use such systems (Francis 1981). There is accumulating evidence in many species that the higher yielding genotypes in sole crop culture do not give the highest yields when intercropped (see Francis 1981 for a review), and that selection in the two regimes has different results (Eagles 1983). It can be predicted that the efficiency of utilisation of common environmental resources in an intercrop can only be optimised by a system of breeding designed to achieve a high degree of mutual adaptation of the component species. In those systems involving intimate association of the components and consequent inter-specific competition for resources, breeding methods appropriate for the development of components must therefore be based on the assessment and selection of potential material under intercrop conditions. This paper suggests a simple scheme answering these requirements, and examines its properties.

2 The selection system

A simple system for the mutual improvement of an arbitrary number of species for use in a single intercrop would involve the random choice of one member from the available population of units of each species to make up a single mixture or intercrop, as many such

* Present address: Department of Statistics, North Carolina State University, Box 8203, Raleigh, NC 27695-8203, USA

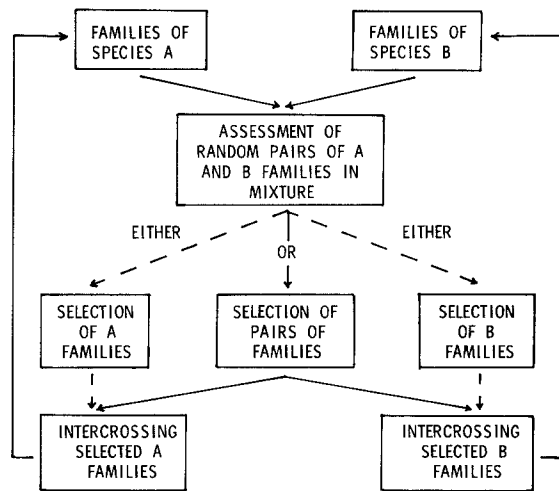


Fig. 1. A diagrammatic representation of the selection scheme

mixtures being constructed as are allowed by the populations and the experimental resources available, and grown in an appropriate replicated trial. Some units from each species would then be selected either on the basis of the yield of the whole mixture in which they occur, their own individual yield, or some other criterion. Selected units from each species would be intercrossed to produce a new set of units for assessment in a new cycle. Each unit would therefore typically be a family, the precise type of family being chosen for genetic or logistic reasons, or may be prescribed by the mating system of the species. So, although the selection system could be used to choose intercrop components from a range of available varieties in a single cycle, it is essentially conceived for recurrent application over the long term. The alternation of phases of selection and recombination parallel those in systems of breeding for pure culture, and so involve no greater expenditure of time or other resources. This system is summarised in Fig. 1.

This system, at least for two species, is analogous to the method of full-sib reciprocal recurrent selection used for the improvement of a hybrid between populations of a single species, in which the genetic value of each member of one population is assessed by means of a full-sib family produced by crossing it with a random member of the other (Hallauer 1967). One important difference is that whereas a single family has to serve as a means of assessment for both its parents, the species in an intercrop can often be separately measured, thus allowing a choice among several selection criteria. This facility will be particularly useful when the products of the species, loosely referred to here as yield, are of a different order of magnitude, or even of a different type with different economic value. This

will not be the case with more intimately grown mixtures, such as grass swards, whose total production has to be measured.

The following analysis deals first with systems of two species, as these are both simpler and likely to be of greater importance in practice than the more general systems which follow. It is assumed throughout that sowing rates and all other cultural and management procedures are equivalent to those to be applied in the conditions of use for which the improved mixture or intercrop is intended.

3 Systems of two species

(a) Statistical models

Because different intercrop components may give products which differ in their economic value, analysis is based on their economic yield, defined as the observed yield multiplied by a coefficient representing the economic value of a unit of yield. For intercrops or mixtures whose components give products of equal economic value which can be simply added to give the total production, then no weighting is necessary. The economic yield of any selection unit of species 1, say the r th, when grown in association with a randomly chosen unit from species 2, the s th, can be described by the model

$$x_{1r} = u_1 + v_{1r} + a_{2s} + (va)_{1r2s} + e_{1r}, \quad (1)$$

where u_1 is the mean of all units of species 1 grown, v_{1r} is the direct effect of the r th unit of 1 in mixture with species 2, a_{2s} is the associate effect of the s th unit of 2 on species 1, and $(va)_{1r2s}$ is their interaction. These parameters are functions of both the inter- and intra-specific competition effects experienced by plants in the particular system of culture. The term e_{1r} represents an error in the estimation of x_{1r} which might be achieved as the mean of several replicate plots. Similarly

$$x_{2s} = u_2 + v_{2s} + a_{1r} + (va)_{2s1r} + e_{2s},$$

while the total economic yield of the mixture can be written as

$$y_{rs} = u_{12} + g_{1r} + g_{2s} + d_{1r2s} + e_{rs}, \quad (2)$$

where $u_{12} = u_1 + u_2$ is the mean of all mixture totals, $g_{1r} = v_{1r} + a_{1r}$ is the general mixture effect of r and $d_{1r2s} = d_{2s1r} = (va)_{1r2s} + (va)_{2s1r}$ is the interaction of r and s in total mixture yields.

For model (1), the parameters have the usual properties:

$$\sum_r v_{1r} = \sum_r a_{1r} = \sum_s v_{2s} = \sum_s a_{2s} = \sum_r (va)_{1r2s} = \sum_s (va)_{1r2s} = 0,$$

$$\begin{aligned} \text{and } E(v_{1r}^2) &= \sigma^2 v_1, E(a_{1r}^2) = \sigma^2 a_1, E(v_{2s}^2) = \sigma^2 v_2, \\ E(a_{2s}^2) &= \sigma^2 a_2, E((va)_{1r2s}^2) = \sigma^2 va_{12}, E((va)_{2s1r}^2) = \sigma^2 va_{21}, \\ E(e_{1r}^2) &= \sigma^2 e_1, E(e_{2s}^2) = \sigma^2 e_2, E(v_{1r} a_{1r}) = \sigma v_1 a_1, \\ E(v_{2s} a_{2s}) &= \sigma v_2 a_2, E((va)_{1r2s} (va)_{2s1r}) = \sigma va_{12} va_{21}, \\ \text{and } E(e_{1r} e_{2s}) &= \sigma e_1 e_2, \end{aligned}$$

all other cross products having zero expectations. Hence each species is characterised by two variances and a covariance, as is the species pair.

For model (2) similarly,

$$\begin{aligned} \sum_r g_{1r} &= \sum_s g_{2s} = \sum_r d_{1r2s} = \sum_s d_{2s1r} = 0, \\ E(g_{1r}^2) &= \sigma^2 g_1, E(g_{2s}^2) = \sigma^2 g_2, \text{ and } E(d_{1r2s}^2) = \sigma^2 d_{12}. \end{aligned}$$

(b) Population statistics

The expectations of various estimable statistics characterising the population of mixtures can now be written. The variance of economic yields of species 1 and 2 are

$$\begin{aligned} \text{var } x_1 &= \sigma^2 v_1 + \sigma^2 a_2 + \sigma^2 va_{12} + \sigma^2 e_1 = \text{var}(g) x_1 + \sigma^2 e_1 \\ \text{var } x_2 &= \sigma^2 v_2 + \sigma^2 a_1 + \sigma^2 va_{21} + \sigma^2 e_2 = \text{var}(g) x_2 + \sigma^2 e_2 \end{aligned}$$

and their covariance is

$$\begin{aligned} \text{cov } x_1 x_2 &= \sigma v_1 a_1 + \sigma v_2 a_2 + \sigma va_{12} va_{21} + \sigma e_1 e_2 \\ &= \text{cov}(g) x_1 x_2 + \sigma e_1 e_2. \end{aligned}$$

It may be noted that the first two terms in the expectation of $\text{cov } x_1 x_2$ are likely to be negative, whereas the error covariance $\sigma e_1 e_2$ will normally be positive. The suffix (g) is used to denote a variance or covariance of true, error-free, values. The covariance of each species with the mixture totals can then be found as $\text{cov } x_1 y = \text{var } x_1 + \text{cov } x_1 x_2$, and the variance of the mixture totals is

$$\begin{aligned} \text{var } y &= \text{cov } x_1 y + \text{cov } x_2 y = \sigma^2 v_1 + 2\sigma v_1 a_1 + \sigma^2 a_1 \\ &\quad + \sigma^2 v_2 + 2\sigma v_2 a_2 + \sigma^2 a_2 + \sigma^2 va_{12} + 2\sigma va_{12} va_{21} \\ &\quad + \sigma^2 va_{21} + \sigma^2 e \end{aligned}$$

which can be written more easily in terms of model (2) as

$$\text{var } y = \sigma^2 g_1 + \sigma^2 g_2 + \sigma^2 d_{12} + \sigma^2 e$$

since $\sigma^2 g_1 = \sigma^2 v_1 + 2\sigma v_1 a_1 + \sigma^2 a_1$, $\sigma^2 g_2 = \sigma^2 v_2 + 2\sigma v_2 a_2 + \sigma^2 a_2$, $\sigma^2 d_{12} = \sigma^2 va_{12} + 2\sigma va_{12} va_{21} + \sigma^2 va_{21}$, and $\sigma^2 e = \sigma^2 e_1 + 2\sigma e_1 e_2 + \sigma^2 e_2$.

(c) Selection

Three simple selection criteria are available for each species in an intercrop in which they can be separately harvested, namely x_1 , x_2 and y . The expected gain in each case can be expressed by expanding the usual formula for linear change in a normally distributed

population under truncation selection in the form

$$G(R; S; C) = i_C \text{cov}(R; S; C) / (\text{var } C)^{1/2}$$

(Falconer 1960), where i_C is the intensity of selection applied to the selection criterion C , $\text{var } C$ is the variance of C and $\text{cov}(R; S; C)$ is a covariance which depends on the species on which selection is practised (S), the species in which response is measured (R), and the selection criterion. The expectations of $\text{var } C$ for each case have already been given in the previous section, and can always be estimated directly from the available data. The properties of the covariance $\text{cov}(R; S; C)$ will now be examined in more detail for specific types of selection.

(i) Selection for either or both species on the basis of their own economic yields (i.e. using x_1 as criterion for species 1 and x_2 for species 2). Selection for species 1 can operate only on those terms in the selection criterion x_1 which are specific properties of species 1, namely v_1 effects. For the effect of this selection on the yield obtained from species 1, we have

$$\text{cov}(1; 1; 1) = \text{cov } v_1 x_1 = \sigma^2 v_1$$

and the effect on species 2 is

$$\text{cov}(2; 1; 1) = \text{cov } v_1 x_2 = \sigma v_1 a_1.$$

The total gain in the system following this selection in species 1 can be derived from a covariance in which the element R is written as the list 1, 2, so that

$$\text{cov}(1, 2; 1; 1) = \text{cov}(1; 1; 1) + \text{cov}(2; 1; 1), \text{ and}$$

$$G(1, 2; 1; 1) = G(1; 1; 1) + G(2; 1; 1)$$

$$\begin{aligned} &= \frac{i_1}{\sigma x_1} (\text{cov}(1; 1; 1) + \text{cov}(2; 1; 1)) \\ &= \frac{i_1}{\sigma x_1} (\sigma^2 v_1 + \sigma v_1 a_1). \end{aligned}$$

The gain from selection in species 2 can be derived similarly. When selection is carried out simultaneously in both species, then the total gain can be subdivided into separate responses from each species, separate selection effects of both species, or into the four components due to the effect of each selection on each species. So the total gain in yield of species 1 is

$$G(1; 1; 1) + G(1; 2; 2) = \frac{i_1}{\sigma x_1} \sigma^2 v_1 + \frac{i_2}{\sigma x_2} \sigma v_2 a_2,$$

and that for species 2 is

$$G(2; 1; 1) + G(2; 2; 2) = \frac{i_1}{\sigma x_1} \sigma v_1 a_1 + \frac{i_2}{\sigma x_2} \sigma^2 v_2.$$

(ii) Selection for either or both species on the basis of the yield of the associate (i.e. using x_2 as a criterion for species 1 and x_1 for species 2). Selection in either

case can now operate only on the associate (a) effects, so

$$\text{cov}(1; 1; 2) = \text{cov } a_1 x_1 = \sigma v_1 a_1 \quad \text{and}$$

$$\text{cov}(2; 1; 2) = \text{cov } a_1 x_2 = \sigma^2 a_1$$

follows selection in species 1.

The total gains in the two species following simultaneous selection for both are then

$$G(1; 1; 2) + G(1; 2; 1) = \frac{i_2}{\sigma x_2} \sigma v_1 a_1 + \frac{i_1}{\sigma x_1} \sigma^2 a_2 \quad \text{and}$$

$$G(2; 1; 2) + G(2; 2; 1) = \frac{i_2}{\sigma x_2} \sigma^2 a_1 + \frac{i_1}{\sigma x_1} \sigma v_2 a_2$$

respectively.

(iii) Selection for both species on the basis of the yields of one, say x_1 irrespective of that of the second, x_2 . A new principle is introduced in this case, because selection can now pick out specific pairs in mixture and operate on terms which are either a specific property of one component or a joint property of both. So, for the selected list 1, 2:

$$\text{cov}(1; 1, 2; 1) = \text{cov}(v_1 + a_2 + (va)_{12}) x_1 = \sigma^2 v_1 + \sigma^2 a_2 + \sigma^2 va_{12} = \text{var}(g) x_1$$

$$\begin{aligned} \text{cov}(2; 1, 2; 1) &= \text{cov}(v_1 + a_2 + (va)_{12}) x_2 \\ &= \sigma v_1 a_1 + \sigma v_2 a_2 + \sigma va_{12} va_{21} \\ &= \text{cov}(g) x_1 x_2 \end{aligned}$$

and the total gain is

$$\begin{aligned} G(1, 2; 1, 2; 1) &= \frac{i_1}{\sigma x_1} (\text{var}(g) x_1 + \text{cov}(g) x_1 x_2) \\ &= \frac{i_1}{\sigma x_1} \text{cov}(g) x_1 y. \end{aligned}$$

Here the use of a common selection criterion for both species allows the utilisation of the variance due to their interaction, whereas with separate criteria these interactions are lost. Note that, as before, the covariance has been defined for the effect of each selection criterion (now only one) on each responding species.

(iv) Selection of both species can be carried out on the basis of total mixture yields (y). Again a common criterion allows interaction terms to be selected for, so that all terms which are properties of either or both species in y contribute, and for species 1 the covariance is that between x_1 and the true values of y , so

$$\text{cov}(1; 1, 2; 1, 2) = \text{cov}(g) x_1 y = \text{var}(g) x_1 + \text{cov}(g) x_1 x_2,$$

and for species 2

$$\text{cov}(2; 1, 2; 1, 2) = \text{cov}(g) x_2 y = \text{cov}(g) x_1 x_2 + \text{var}(g) x_2,$$

so that the total gain is

$$G(1, 2; 1, 2; 1, 2) = G(1; 1, 2; 1, 2) + G(2; 1, 2; 1, 2)$$

$$= \frac{i_y}{\sigma y} \text{var}(g) y.$$

The gain in species 1 is less than that in (iii) because the new contribution from $\text{cov}(g) x_1 x_2$ is likely to be negative whereas the gain in species 2 is increased by the addition of $\text{var}(g) x_2$.

4 General systems of m species

(a) Statistical models

For systems of m species, where $m > 2$, each species contributes an associate effect to more than one other species. Each such effect is characteristic of the two species involved and so has to be defined as a_{ijs} for the average effect of the s th member of the j th species on all members of the i th, with $(va)_{irjs}$ remaining as the specific effect on the r th member. The yield of the r th unit of species 1 when grown with randomly chosen units of all other species (notated here for simplicity as the s th throughout) can be written as

$$x_{ir} = u_i + v_{ir} + \sum_{j \neq i} a_{ijs} + \sum_{j \neq i} (va)_{irjs} + e_{ir} \quad (3)$$

with summation over all m species. It should be noted that all the parameters are specific to culture in mixtures with the specific species involved, and so those with species 1, say, need bear no relationship with those for species 1 in formula (1). All other things being equal, the proportion of the total competition experienced by any plant which is inter-specific increases with m , and with random mixing it is $(m-1)/m$. The a_{ij} parameters may also differ considerably for different i , as they are a property of the economic yields of these different species. The model deals only with interactions up to the first order, whereas interactions up to the $(m-1)$ th order are possible. Some influences of higher order effects will be discussed later. A model for mixture totals is now

$$y = \sum_i u_i + \sum_i g_{ir} + \sum_i \sum_{j > i} d_{irjs} + e. \quad (4)$$

Where $g_{ir} = v_{ir} + \sum_{j \neq i} a_{jir}$ and $d_{irjs} = (va)_{irjs} + (va)_{jsir}$. For model (3) the parameters are constrained so that $\sum_r v_{ir} = \sum_s a_{ijs}$ $= \sum_r (va)_{irjs} = \sum_s (va)_{irjs} = 0$, with $E(v_{ir}^2) = \sigma^2 v_i$, $E(a_{ijs}^2) = \sigma^2 a_{ij}$, $E(a_{ijs} a_{kjs}) = \sigma a_{ij} a_{kj}$, $E(v_{ir} a_{jir}) = \sigma v_i a_{ij}$, $E((va)_{irjs}^2) = \sigma^2 va_{ij}$, $E((va)_{jsir}^2) = \sigma^2 va_{ji}$, and $E((va)_{irjs} (va)_{jsir}) = \sigma va_{ij} va_{ji}$, each defined for any values of i and j . For model (4) similarly, $E(g_{ir}^2) = \sigma^2 g_i$ and $E(d_{irjs}^2) = \sigma^2 d_{ij}$ for all values of i and j .

(b) Population statistics

The variance of yields of species i now has the expectation

$$\text{var } x_i = \sigma^2 v_i + \sum_{j \neq i} \sigma^2 a_{ij} + \sum_{j \neq i} \sigma^2 va_{ij} + \sigma^2 e_i$$

and the covariance of yields of species i and j is

$$\text{cov } x_i x_j = \sum_{k \neq ij} \sigma a_{ik} a_{jk} + \sigma v_i a_{ji} + \sigma v_j a_{ij} + \sigma va_{ij} va_{ji} + \sigma e_i e_j$$

The covariance of yields of species *i* with mixture totals is

$$\begin{aligned} \text{cov } x_i y &= \text{var } x_i + \sum_{j \neq i} \text{cov } x_i x_j = \sigma^2 v_i + \sum_{j \neq i} \sigma^2 a_{ij} \\ &+ \sum_{j \neq i} \sum_{k \neq i, j} \sigma a_{ik} a_{jk} + \sum_{j \neq i} \sigma v_i a_{ji} + \sum_{j \neq i} \sigma v_j a_{ij} \\ &+ \sum_{j \neq i} \sigma^2 v a_{ij} + \sum_{j \neq i} \sigma v a_{ij} v a_{ji} + \sigma^2 e_i + \sum_{j \neq i} \sigma e_i e_j \end{aligned}$$

so that the variance of mixture totals can be written as

$$\begin{aligned} \text{var } y &= \sum_i \text{cov } x_i y = \sum_i \sigma^2 v_i + \sum_i \sum_{j \neq i} \sigma^2 a_{ij} + \sum_i \sum_{j \neq i} \sum_{k \neq i, j} \sigma a_{ik} a_{jk} \\ &+ 2 \sum_i \sum_{j \neq i} \sigma v_i a_{ji} + \sum_i \sum_{j \neq i} \sigma^2 v a_{ij} + \sum_i \sum_{j \neq i} \sigma v a_{ij} v a_{ji} \\ &+ \sum_i \sigma^2 e_i + \sum_i \sum_{j \neq i} \sigma e_i e_j \end{aligned}$$

or much more simply in terms of model (4) as

$$\text{var } y = \sum_i \sigma^2 g_i + \sum_i \sum_{j > i} \sigma^2 d_{ij} + \sigma^2 e$$

since $\sigma^2 g_i = \sigma^2 v_i + 2 \sum_{j \neq i} \sigma v_i a_{ji} + \sum_{j \neq i} \sigma^2 a_{ji} + \sum_{j \neq i} \sum_{k \neq i, j} \sigma a_{ik} a_{jk}$,
 $\sigma^2 d_{ij} = \sigma^2 v a_{ij} + 2 \sigma v a_{ij} v a_{ji} + \sigma^2 v a_{ij}$, and $\sigma^2 e = \sum_i \sigma^2 e_i + \sum_i \sum_{j \neq i} \sigma e_i e_j$.

(c) Selection

The possible systems of selection for mixtures or intercrops of more than two species are more numerous and potentially more complex. However, certain of the principles developed for that simple case apply in general and can be used to guide the development of more general formulae. It was shown that the important covariance cov (R; S; C) is a function of three parameters: the species in which gain is measured (R), the species under selection (S), and the species included in the criterion used for assessment and selection (C). It was also seen that the total gain in the system is always the sum of the gains of the individual species. Similarly, the gains from the application of separate selection criteria are also independent and additive provided that each species is selected only on the basis of one criterion, even if this criterion is based on the yields of more than one species. Thus it is necessary only to develop a formulation sufficiently general to express the covariance appropriate to the response of a single species following the selection of an arbitrary number of species on the basis of a single common criterion. In the notation of such a covariance it is therefore necessary to introduce a list only with respect to the second parameter, S.

The selection criterion can in general be the weighted sum of the yields of several species, as would be the case for a selection index. In this case, selection is based on the weighted effects

$$C = \sum_i b_i x_i = \sum_i b_i v_i + \sum_i b_i \sum_{j \neq i} a_{ij} + \sum_i b_i \sum_{j \neq i} (va)_{ij}.$$

The individual coefficients b_i applied to each x_i may take any value, including zero. It has been noted that selection can operate on effects which are individual or joint properties of the species under selection. The above expression can be converted into the effective selection criterion by the inclusion of an indicator function ϕ_i which takes the value 1 when species *i* is under selection and 0 otherwise. The criterion is then

$$SC = \sum_i b_i \phi_i v_i + \sum_i b_i \phi_i \sum_{j \neq i} a_{ij} + \sum_i b_i \phi_i \sum_{j \neq i} \phi_j (va)_{ij}.$$

Response is to be measured in species *k*, so

$$x_k = \sum_i \theta_i v_i + \sum_{j \neq i} \theta_i a_{ij} + \sum_i \theta_i \sum_{j \neq i} (va)_{ij},$$

where θ_i takes the value 1 when $k=i$, but 0 otherwise. The covariance of the selection criterion with x_k can now be expanded as

$$\begin{aligned} \text{cov} &= \sum_i b_i \phi_i \theta_i \sigma^2 v_i + \sum_i b_i \phi_i (1 - \theta_i) \sigma v_i a_{ki} \\ &+ \sum_i b_i \theta_i \sum_{j \neq i} \phi_j \sigma^2 a_{ij} + \sum_i b_i (1 - \theta_i) \sum_{j \neq i} \phi_j \sigma a_{ij} a_{ki} \\ &+ \sum_j b_j \sum_{i \neq j} \phi_i \theta_i \sigma v_i a_{ji} + \sum_i b_i \phi_i \theta_i \sum_{j \neq i} \phi_j \sigma^2 v a_{ij} \\ &+ \sum_i b_i \phi_i \sum_{j \neq i} \phi_j \theta_j \sigma v a_{ij} v a_{ji}. \end{aligned} \tag{5}$$

This expression includes only the terms up to the first order of interaction that were included in model (3). Contributions from second order terms include variances of associate \times associate interactions ((aa)_{*i,j*}) of the *j* th and *i* th species onto the *i* th, variances of direct \times associate \times associate interactions (e.g. (vaa)_{*ijl*}) and several covariances among these parameter sets. With higher order interactions the situation rapidly becomes very complex, with a large number of statistics with coefficients written in terms of the indicator functions already defined.

Particular types of selection can be derived from formula (5) by the omission of redundant indicator functions when these are all unity, or by taking only terms whose indicator products are non-zero otherwise. The following general properties can be recognised.

(i) When selection is carried out on a single species (the *i* th) on the basis of a selection criterion of one species yields, then the possible covariances are cov (*i*; *i*; *i*) = $\sigma^2 v_i$, cov (*j*; *i*; *i*) = cov (*i*; *j*; *j*) = $\sigma v_i a_{ji}$, cov (*j*; *i*; *j*) = $\sigma^2 a_{ji}$ and cov (*k*; *i*; *j*) = $\sigma a_{ji} a_{ki}$, according to the identities of the selected, responding and criteria species.

(ii) When selection is applied to all species simultaneously on the basis of the yields of one, then

$$\begin{aligned} \text{cov} (i; 1 \dots m; i) &= \text{var} (g) x_i, \quad \text{and} \\ \text{cov} (i; 1 \dots m; j) &= \text{cov} (j; 1 \dots m; i) = \text{cov} (g) x_i x_j, \end{aligned}$$

according to whether the responding and criterion species are identical or not. It is only in this case that the covariance reduces to a simple statistic which can be readily estimated from the available data, although the omission of higher order terms from the formulae means that this has not been explicitly shown.

In either of the above cases, the covariance due to including several species (say 1... *t*) in the selection criterion is obtained simply by summing individual covariances over these species, irrespective of the single species or group on which response is measured (R) and of the single species or group under selection (S), so that

$$\text{cov} (R; S; 1 \dots t) = \sum_i \text{cov} (R; S; i).$$

Similarly covariances for all *r* responding species can be summed so that

$$\text{cov} (1 \dots r; S; C) = \sum_i \text{cov} (i; S; C).$$

The covariances for separate selection criteria cannot be summed in this way, as the gain from each depends on its standard deviation and selection intensity, and any summation

must therefore be carried out at the level of the gains themselves:

$$G(R; 1 \dots s; C) = \sum_i^s G(R; i; C).$$

5 Some specific properties of selection methods

Consider for simplicity a two species system in which yield is controlled only by direct effects. The gain from direct selection expected in species 1 is then $i\sigma^2v_1/(\sigma^2v_1 + \sigma^2e_1)^{1/2}$ whereas that resulting from selection among whole mixtures is $i\sigma^2v_1/(\sigma^2v_1 + \sigma^2v_2 + \sigma^2e)^{1/2}$. Thus direct selection is the more efficient provided that the error variance σ^2e is not markedly smaller than σ^2e_1 , which is unlikely. The inferiority of mixture selection is due to the necessity of selecting a randomly associated pair of species candidates as a single unit, and would therefore be even more marked in systems with more than two species. When both direct and associate effects are present, then provided that these are only weakly correlated, selection for either direct or associate effects may still be more effective than mixture selection. However, negative correlations between direct and associate effects can be expected to be rule rather than the exception, particularly for crops in which yield is itself a major determinant of aggressiveness, such as the forages. The gain in yield in each species under direct selection now consists of a positive term due to its increased yield potential following its own selection ($i_1\sigma^2v_1/\sigma x_1$ for species 1) and a negative term due to increased competition following similar selection in its associate ($i_2\sigma v_2a_2/\sigma x_2$). This reduction in gain from individual species selection results from the presence of negative covariances in the numerators of the response formulae which are not present in the denominators (σx_1 , or σx_2), and is therefore more severe than in the case of whole mixture selection for which both the numerator and denominator (σy) are affected. Negative correlations between direct and associate effects therefore favour the use of whole mixture selection. It is worth emphasising at this point that the absence of interactions, as assumed here, does not indicate that monoculture or any other form of testing need be an efficient substitute for assessment in mixture or intercrop. Intercrop and monoculture performance remain two distinct attributes whose correlation can only be determined empirically.

In spite of the above arguments, the presence of important interactions between direct and associate effects is likely to be the most compelling reason for the preference of mixture selection. Although the higher order interactions possible in systems of several species have not been treated in any detail here, the principles governing two species systems can be generalised to some extent. Simultaneous selection of a set of

species on the basis of a common criterion will allow utilisation of multiple interactions involving all of them, but this grouping of selection units will reduce the gain from all lower order effects and interactions. There is therefore little purpose in the joint selection of groups larger than that of the highest level of interaction present. In practice, interactions of more than two species may be relatively unimportant, so that pairwise groupings would be adequate. However, even in a three species system, conflicts would arise if species A interacts with B and with C but other interactions are absent.

6 Selection indices

A selection index is a selection criterion derived as a weighted sum of a series of available measurements relating to the units under selection. In the case of intercrop assessment, each of the m species provides a yield variate which can be included in such an index. The use of characters other than yield will not be considered here, although it can be treated in exactly the same way. The coefficients to be applied to each variate in the case of an index designed to give maximum total gain are found by solution of the system of equations described in matrix form as

$$b = P^{-1} G a \quad (\text{Smith 1936})$$

Here P is the $(m \times m)$ variance-covariance matrix of the yield variates, with its ij th element equal to $\text{cov } x_i x_j$. The $(m \times 1)$ vector a is used to apply differential economic values to the yields of each species, but since it has been assumed that all yields used have already been measured in economically standardised units, then its elements may be set to unity. Alternatively, indices may be calculated from untransformed yields and the elements of the vector used to represent the relative economic values of unit yields in each species.

The elements of the $(m \times m)$ matrix G are members of the $\text{cov}(R: S: C)$ family already discussed, and depend in each case on the species on which the index is to be used, as follows.

(a) Separate indices for each species

Selection for each species may be based on its own index. For the k th species, the appropriate G matrix has as its k th row

$$\sigma v_k a_{k1}, \dots, \sigma v_k a_{k(k-1)}, \sigma^2 v_k, \sigma v_k a_{k(k+1)}, \dots, \sigma v_k a_{km}$$

while any other row, the i th say, is

$$\sigma a_{ki} a_{k1}, \dots, \sigma a_{ki} a_{k(i-1)}, \sigma^2 a_{ki},$$

$$\sigma a_{ki} a_{k(i+1)}, \dots, \sigma a_{ki} a_{km}.$$

The use of such individual indices for one species is expected to maximise the utilisation of the species v and a effects. Simultaneous applications of specific indices to each species will make good use of all v and a effects, but because each index is calculated independently of the others and yet employs the same yield variates, it is doubtful if they jointly possess any property of maximisation of gain. They can make fortuitous use of interactions only when indices for different species happen to result in the selection of units from the same intercrop. With large populations this is unlikely, particularly for small values of m . The most serious problem facing the user of such indices however is the difficulty of estimation of the elements of the matrix G , as these cannot be achieved from the data provided by the trial itself.

(b) Common index for all species

A common index to be applied to all species, and therefore resulting in the choice or rejection of all components in a particular treatment, is calculated in the same way as those above, but using a G matrix whose ij th element is $\text{cov}(i; 1 \dots m; j)$, or simply $\text{cov}(g) x_i x_j$. There is therefore little problem in estimating the elements of G from a replicated trial. The index will be expected to make good use of interaction effects, but be less efficient than separate indices with respect to direct and associate effects. It may be that an index or set of indices for all species exists which has an overall optimal property, but this must remain an unsolved problem for the present.

(c) A specific use of a selection index

A situation which might commonly occur in practice is the mixed cropping of two species of which one is very sensitive to the competition exerted by the other, as when species 1 is much smaller in stature and grows in the shade of species 2. A possible selection system would be to select both species on the basis of the yield of the weaker, so that the whole of the selective effort would be aimed at increasing yields of species 1, both through its own yield potential and through reduced competition from species 2. With negative correlations of direct and associate effects for either species, this system would necessarily be expected to reduce yields of species 2 however, and this would be unacceptable if its produce was of importance. This would also be the expected result if a fixed representative of species 2 was used in the testing of species 1 units as is sometimes the practice (Willey 1979). A compromise which could be made would be the use of a selection index designed to hold the yield of species 2 constant while improving species 1. A common index for the two

species would then have the form

$$I = x_1 + b x_2, \quad \text{where } b = -\text{cov}(g) x_1 x_2 / \text{var}(g) x_2.$$

But this index would not necessarily have the property of maximising gain in species 1. A unique pair of indices, one for each species, must exist with the joint properties of giving the maximum gain in species 1 while holding that of species 2 constant, but this has yet to be solved.

7 Discussion

It has been shown that no predictions of the absolute or relative gains to be expected from selection based on any criterion other than whole intercrop yields can be derived from data provided by a trial involving randomly constituted intercrops. Estimates of the variances and covariances of direct, associate and interaction effects of all species which are necessary for the choice of individual selection criteria or calculation of indices can be obtained only from an experimental design in which each species unit is tested in association with at least two units from each other species, and preferably more. The most comprehensive design would include all possible combinations of selection units from the different species.

Such a system was suggested for the selection of intercrop treatments of two species by Hamblin, Rowell and Redden (1976). However, since this design generates n^m treatments from n units for each of m species, it could only be applied to a sample of the available breeding material as a preliminary to the selection scheme.

The gains predicted by the above methods are the genotypic gains resulting from the selection phase, and the proportion transmitted through the sexual phase which follows depends on the inheritance of the various effects. Strictly speaking, all the statistics used in prediction formulae should be replaced by their additive genetic counterparts, the exact form of these depending on the families used as selection units and the mating system of the species to which they apply. Any differences in heritability of the direct and various associate effects of any species (i.e. v_i and all a_{ji} effects for the i th species) will therefore reduce the accuracy of predictions based only on information about genotypic and phenotypic performance. In principle, additive genetic parameters for outpollinating species could be estimated from an experiment in which the units are full-sib families which belong to defined half-sib groups, and if the half-sib groups of different species coincided in mixtures this also allow the estimation of the additive genetic fraction of interaction variances and covariances. For inbreeding systems, segregating families in a pedigree system or homozygotes following single seed would be used. Differences in breeding

systems between the species involved would certainly cause problems, both for the estimation of parameters and the efficiency of selection, because varying levels of inbreeding are likely to influence the relative vigour of components.

In comparison with pure culture testing, with which it is directly comparable in terms of experimental resource utilisation, the system of intercrop or mixture selection proposed here has both advantages and disadvantages. First, it makes a direct assessment of mixture effects, whereas pure culture selection relies on the correlation of pure and mixture values. Estimates of this important statistic vary among intercropping systems (Francis 1981), although the correlation would be expected to be lower for randomly sown mixtures than for intercrops grown in rows. Second, only mixture selection can make use of component interactions, although the importance of these remains unknown. Third, the major disadvantage of the mixture selection scheme is that the random allocation of the components to mixtures reduces the gain due to their main effects by a factor of about $m^{1/2}$ for systems of m species in comparison with pure culture selection. It was this factor which earlier led to the conclusion that in many circumstances pure culture selection would be more efficient than random mixtures for the improvement of inter-genotypic mixture yield (Wright 1983).

However, these arguments are all derived from the formula for gain which can only make predictions with respect to single selection cycles. The concept of recurrent application embodies the assumption that behaviour over the long term involves phenomena which cannot be predicted in the short term, and so the ultimate potentials of pure culture and different types of mixed culture selection can only be guessed at. This is because, unlike the reciprocal recurrent selection method to which analogy has been made and which is underpinned by genetic laws (Comstock et al. 1949), there is no underlying causative model on which to base any predictions.

Interactions among components provide one mechanism by which the long term behaviour of different selection methods can differ. When selection is carried out for just one component species, then the formulae given earlier are exact, but when two or more are simultaneously selected there is an additional second order effect due to the fact that the population with which each species associates after selection differs from that against which it was tested prior to selection. Consider for simplicity a two species system in which selection for each is based on its own yields. When only one species is selected, the array of interaction terms $\sum_r \sum_s d_{1r2s}$ is reduced in only one dimension, and its expected value (zero) is not changed. When selection is carried out on both species however, then this array is

restricted in both dimensions, and its mean value need no longer be zero. This potential contribution to selection gain by interaction terms is distinct from that of the variances $\sigma^2 v_{a_{12}}$ or $\sigma^2 d_{12}$ and can occur even when selection in the two species is based on independent criteria. In fact it is due to a correlated response of the d_{12} terms to joint changes in v_1 and v_2 , and depends on the complex cross product $\sum_r \sum_s v_{1r} v_{2s} d_{1r2s}$.

This type of phenomenon was earlier examined for the case of within population selection of single plants (Wright 1977). Analysis of diallel sets of binary mixtures suggested that the interactions in groups of plants with high individual forage yields are negative and so retard selection gains. One effect this pattern of interactions has is to induce a curvature into the regression of associate onto direct effects so that it steepens as selection proceeds. When the gradient falls to -1 , then the increased yield potential of selected plants is exactly balanced by their increased mutual aggression, no gain results and a selection plateau is reached. Whether similar patterns of interaction are to be expected in intercrop combinations is unknown. The question also arises as to whether cross products of the form $\sum_r \sum_s a_{1r} a_{2s} d_{1r2s}$ arising from selection based on associate yields are of the same magnitude and sign as those from direct selection shown above, or indeed parallel effects following pure culture assessment. Any difference would be a basis for a difference in long term potential for the methods due to the selection of quite different plant types. In any case, any possibility of this deleterious drift of interactions is best counteracted by the use of whole mixture selection which actively promotes favourable interactions, whether or not it is predicted to give the best rate of gain in the short term.

It has been assumed that all cultural and management practices applied to intercrops in the selection programme are identical to those in the agricultural context for which they are intended. Clearly this is a difficult requirement to fulfill, particularly since optimal sowing dates and densities and fertiliser and pesticide treatment are themselves likely to depend on season and location (Willey 1979; Francis 1981). Interactions of genotypes with managements, locations and seasons are also likely and will reduce the effectiveness of a programme conducted in one location and with one type of management. It is also recognised that the transformation of the measured yields of the intercrop components simply by weighting them according to some assumed economic value is a simplification of a complex problem discussed at greater length elsewhere (Willey 1979; Mead and Riley 1981), as this can depend on whether the crop products are required for subsistence or cash, and on fluctuations in market demand.

References

- Comstock RE, Robinson HF, Harvey PH (1949) A breeding procedure designed to make maximum use of both general and specific combining ability. *Agron J* 41:360–367
- Eagles CF (1983) Relationship between competitive ability and yielding ability in mixtures and monocultures of *Dactylis glomerata* (L). *Grass Forage Sci* 38:21–24
- Falconer DA (1960) Introduction to quantitative genetics. Oliver and Boyd, London
- Francis CA (1981) Development of plant genotypes for multiple cropping systems. Plant breeding II. Iowa State University Press, Ames Iowa
- Hallauer AR (1967) Development of single-cross hybrids from two-eared maize populations. *Crop Sci* 7:192–195
- Hamblin J, Rowell JG, Redden R (1976) Selection for mixed cropping. *Euphytica* 25:97–106
- Mead R, Riley J (1981) A review of statistical ideas relevant to intercropping research. *J R Stat Soc* 144:462–509
- Smith HF (1936) A discriminant function for plant selection. *Ann Eugenics* 7:240–250
- Willey RW (1979) Intercropping – its importance and research needs. 2. Agronomy and research approaches. *Field Crops Abstr* 32:73–85
- Wright AJ (1977) Predictions of non-linear responses to selection for forage yield under competition. *Theor Appl Genet* 49:201–207
- Wright AJ (1983) The expected efficiencies of some methods of selection of components for intergenotypic mixtures. *Theor Appl Genet* 67:45–52