

# Comparative simulation studies on the effects of selection for quantitative traits in autogamous crops: early selection versus single seed descent

A. J. VAN OEVEREN & P. STAM

*Department of Genetics, Agricultural University, Dreijenlaan 2, 6703 HA Wageningen, The Netherlands and Centre for Plant Breeding and Reproduction Research CPRO-DLO, P.O. Box 16, 6700 AA Wageningen, The Netherlands*

A comparison is made between two contrasting breeding procedures for self-pollinating crops by means of computer simulation studies. The first is an early cross selection method, based on cross prediction by  $F_3$  line estimates of the cross mean and between-line variance. Subsequently, line selection is performed. Selection is postponed in the second procedure to a more homozygous,  $F_6$  generation, which is obtained by single seed descent. Only then is line selection performed, regardless of the pedigree.

The two procedures are compared for the cases of one and five crosses under selection. If only one cross is concerned, the early selection method reduces to early line selection and it was found to supply better inbred lines than SSD. But when more crosses are involved, the SSD procedure performed just as well as the early selection method and, at low heritability, even better because the early cross prediction was often poor. Dominance appears to be of very little influence on either selection procedure.

**Keywords:**  $F_\infty$  cross prediction, quantitative trait, selection, self-pollinating crops, simulation, single seed descent.

## Introduction

Various methods are used to obtain high performing inbred lines in the breeding of self-pollinating crops. Selection can be applied in different stages of the breeding process and with varying intensities. In order to obtain a better view of the efficiency of different selection procedures, two contrasting selection methods are examined.

The first is a method, based on a quantitative genetic theory, which has been developed to predict the genetic potential of a certain cross in an early breeding generation (e.g. Mather & Jinks, 1971). This technique, which makes use of estimates of the mean ( $m$ ) and additive genetic variance ( $A$ ) of a cross, should, in theory, allow the selection of those crosses most likely to produce superior inbred lines. Then, only the most promising crosses would be retained in the breeding programme. After this stage there is an opportunity to perform pedigree selection in the subsequent generations. This is referred to as 'early selection' (ES). In this

form ES comprises both between- and within-cross selection.

A second, completely different method avoids selection in the early generations and waits until a high degree of both homogeneity and homozygosity has been reached. Genotypic differences between lines are then more pronounced and stable. At this stage selection between lines can be performed, without regard to the pedigree. An example of this is the method of single seed descent (SSD), and for crops which can be forced to a short generation cycle (e.g. spring cereals), SSD results in a quick advancement towards homozygosity.

It was found from experimental studies on ES (Jinks & Pooni, 1980; Van Ooijen, 1989a,b) that using  $F_3$  lines to estimate the additive variance ( $A$ ) is by far the most practical method for self-pollinating species. Estimates based on  $F_3$  lines, however, may be biased by the effects of intergenotypic competition and in small grain crops this bias can be particularly large because plot size is limited by the amount of seed. Van Ooijen (1989a,b) concluded from his studies with mixtures of

spring wheat that the estimates of the genetic parameters for yield are in fact severely biased, leading to unreliable cross predictions. In addition, early selection is very demanding in terms of labour and the trial field area.

The SSD method has proved to be a fast breeding procedure but tests on the advantages compared to other methods are quite inconsistent. Knott & Kumar (1975) found in their field experiment with spring wheat that an early yield test procedure (EYT) produced lines with a significantly higher average than did SSD. The yield level of the 20 per cent best lines did not, however, differ significantly for EYT and SSD. Boerma & Cooper (1975) also compared EYT with SSD and Pedigree Selection (PS) within crosses of soybean. They found no consistent differences in selection results between the three procedures and therefore regarded the rapid SSD method as most efficient. Computer simulation studies showed that, especially with low heritability, SSD performs just as well as pedigree selection (Casali & Tigchelaar, 1975); however, they considered only one population on a strictly additive model and selection was based on individual plant performance.

A field experiment was carried out to compare both selection procedures, using pseudo-lines of spring wheat (Van Oeveren, 1992). It was concluded that ES can easily lead to erroneous cross selection and that SSD is to be preferred for this particular set of crosses.

It is risky to depend solely on field trials when evaluating the effectiveness of the two different breeding strategies. They will only show a sample of what could happen. In this way, occasional random effects can misrepresent the average actual situation. It can therefore be very useful to examine the very same methods by means of computer simulations: the large number of replicate experiments that can be run provides a more secure base for statements about the average performance of both procedures. There are, of course, some disadvantages to this approach; simplifying assumptions must be made. On the other hand, simulation models allow variation of the parameter values, such that a realistic range of input variables can easily be tested; and more importantly, the relative significance of the input variables on the selection results can be studied.

## Materials and methods

A computer program was written in PASCAL which could simulate a breeding programme from the  $F_1$  to the  $F_7$ .

A relatively simple genetic model was assumed; a varying number of loci (up to 100) was supposed to determine a true quantitative trait. The  $F_1$  could be

segregating for 30 of these loci, at most. All loci were assumed to have equal effect and to show no interaction (i.e. no epistasis) and no linkage. The two different breeding strategies are schematically visualized in Fig. 1. Starting with a certain  $F_1$ , which is heterozygous for a number of loci, an  $F_2$  of size  $NF_2$  is created by selfing. Two sets of progeny are then created from this  $F_2$ . All  $F_2$  plants are advanced to  $F_3$  lines for the early selection procedure and for the SSD procedure just a single  $F_3$  plant is derived from each  $F_2$  plant. Every plant is supposed to give an equal number of progeny.

The  $F_2$  plant progenies will be large enough to allow an early yield testing. Each line is grown in the three-row plot in two replications. Based on the  $F_3$  cross mean and the additive variance, which is estimated as twice the between-line variance, the probability for each cross of retaining superior inbred lines in the  $F_\infty$  is calculated (see Van Ooijen, 1989b). Only the most promising crosses are propagated and from these a mild selection (50 per cent) is made among lines. These best lines are again increased to larger  $F_4$  lines which give a more sound yield estimate. Each line is grown in a larger plot in four replicates. The 20 best lines are selected. From each line five plants are randomly chosen and they are each propagated to two three-row  $F_5$  plots. These again offer a rough yield estimate and allow the opportunity for line selection. Again the best 50 per cent are increased to the  $F_6$  which consists of fields of the same size as the  $F_4$ . The final evaluation takes place in this generation. The genotypic mean, maximum and standard deviation of the 10 phenotypic best lines are recorded.

An  $F_5$  generation is derived from the SSD- $F_3$  by two more successive rounds of SSD. This  $F_5$  is space planted and increased to  $NF_2$   $F_6$  lines. Yield estimates are obtained from the  $F_6$  lines and the best 50 per cent are again propagated to large  $F_7$  lines in four replicates. The 10 phenotypically best lines are finally evaluated and compared to those of the ES- $F_6$ . The number of lines in each generation, relative to the number of  $F_2$  plants are given in Table 1.

Phenotypic values are simulated for the two types of plots: the three-row plots (ES- $F_3$ , ES- $F_5$  and SSD- $F_6$ ), which consist of 100 plants and the large plots (ES- $F_4$ , ES- $F_6$  and SSD- $F_7$ ), which consist of 600 plants. The model for the phenotype of plant  $l$  of line  $k$  in replicate  $j$  is as follows.

$$y_{jkl} = \mu + r_j + E_{jk} + ic_{jk} + g_{jkl} + e_{jkl},$$

where  $\mu$  = the overall mean,

$r_j$  = the effect of replicate  $j$

$E_{jk}$  = the between-plot environmental error of plot  $jk$

$$r_j \approx N(0, \sigma_r^2),$$

$$E_{jk} \approx N(0, \sigma_{eh}^2),$$

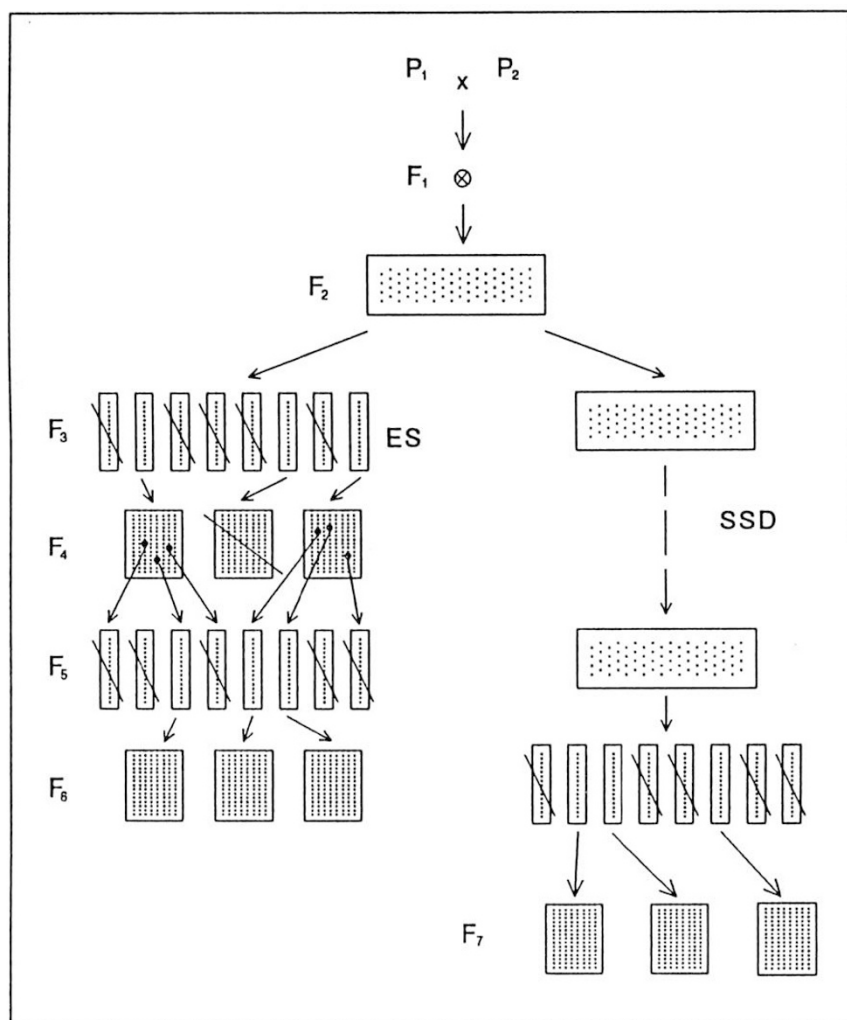


Fig. 1 Scheme of the two breeding procedures in case of one cross. Early Selection (left) and Single Seed Descent (right).

Table 1 Numbers of lines in each generation for both selection procedures as a function of the number of F<sub>2</sub> plants. Except for the ES generations F<sub>4</sub>-F<sub>6</sub> these are numbers of lines per cross

NF <sub>2</sub>	ES				SSD	
	F <sub>3</sub>	F <sub>4</sub>	F <sub>5</sub>	F <sub>6</sub>	F <sub>3</sub> ...F <sub>6</sub>	F <sub>7</sub>
25	25	20	100	50	25	20
50	50	25	100	50	50	25
100	100	50	100	50	100	50
200	200	100	100	50	200	100
400	400	200	100	50	400	100

The parameter  $\mu$  is entered as an input variable, as well as the additive ( $a$ ) and dominance ( $d$ ) effects, which are simply summed over all loci to give the genotypic value  $g$ .

The heritability on a per plant basis can be derived from the input variables  $H^2$  ( $F_3$  between line heritability),  $a$ ,  $d$  and  $n$  (number of plants per  $F_3$  plot).

$$h^2 = \frac{H^2 \cdot n\sqrt{n} - (n-1) \cdot (\sqrt{n}-1)}{n\sqrt{n} - H^2 \cdot (n-1) \cdot (\sqrt{n}-1)}$$

As the genetic variance between  $F_3$  lines equals:  $V_{IF_3} = \frac{1}{2}A + \frac{1}{16}D$  (i) [in the absence of epistasis;  $A$  and  $D$  are the sum of the quadratic effects of additivity and dominance respectively over all segregating loci (Mather & Jinks, 1971)] and (ii)  $h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2)$ , then the environmental error variance can be calculated as:

$$\sigma_e^2 = \frac{1-h^2}{h^2} \cdot \left( \frac{1}{2}A + \frac{1}{16}D \right)$$

$i_{jk}$  = the between-plot competition effect of plot  $jk$

$$i_{jk} \approx N(0, \sigma_{ic}^2)$$

$g_{jkl}$  = the genotype of plant  $jkl$ ,

$e_{jkl}$  = the within-plot environmental error of plant  $jkl$

$$e_{jkl} \approx N(0, \sigma_{ew}^2)$$

This variance can be divided into a between- and a within-plot component. Assuming a soil heterogeneity index of 0.5 (Fairfield Smith, 1938) this amounts to:

$$\sigma_{eb}^2 = \frac{\sigma_e^2}{\sqrt{n}} \text{ (between plots) and}$$

$$\sigma_{ew}^2 = \sigma_e^2 - \sigma_{eb}^2 \text{ (within plots).}$$

The replicate effect is assessed as

$$\sigma_r^2 = \sigma_{eb}^2 \cdot n \cdot m,$$

where  $m$  stands for the number of plots in one replicate.

Finally, an extra random effect was added to the plot sums to cover the effects of intergenotypic competition. From earlier research (A. J. van Oeveren unpublished), simulating both yield and competition, it was found that competition effects within plots were negligible compared to both the between-plot competition effects and the environmental error. A normally distributed error term could well mimic the between-plot competition and the size of the effect was found to be related to both the number of plants per plot and the between-line variance, which varies with the level of heterogeneity of the generation. The following formula was found to perform well.

$$\sigma_{ic}^2 = 2.25 \cdot n \cdot G_F, \text{ where}$$

$$G_F = \begin{cases} 1 & \text{in } F_3\text{-ES,} \\ 1.7 & \text{in } F_5\text{-ES and} \\ 2 & \text{in } F_6\text{-SSD.} \end{cases}$$

## Results

### One cross

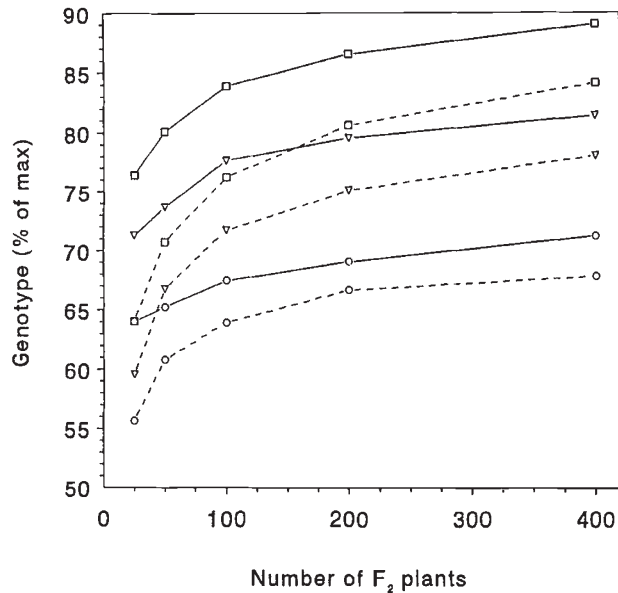
At first, simulations were performed to compare ES with SSD when only one cross was considered. In fact this results in early line selection instead of cross selection. As the possible negative effects of selecting the wrong crosses are completely absent, it is expected that ES will perform better than SSD, especially at high heritability. The quantitative trait of interest is assumed to be determined by 30 loci and the  $F_1$  is heterozygous for 10 of them in the range 16–25. This indicates that 15 loci are positive homozygous and five are negative homozygous. Heritability and the number of  $F_2$  plants ( $NF_2$ ) appear to be the two main sources of influences and results have been obtained in relation to these two variables. The first varies from 0.02 up to 1.0 and  $NF_2$  varies from 25 to 400 plants. The results are presented in Table 2. The genotypic plot totals are given, expressed as a percentage of their maximum possible

**Table 2** Genotypic mean, standard error of that mean and maximum of the 10 phenotypic best lines in both the ES- $F_6$  and the SSD- $F_7$ . Mean and maximum expressed as a percentage of the difference between the theoretical maximum and minimum genotype. Selection in only one cross with 10 segregating loci. The number of runs varies from 150 to 300

$H^2$	$NF_2$	ES			SSD		
		Mean	s.e.	max.	Mean	s.e.	max.
0.02	25	64.0	4.9	81.9	55.7	4.8	78.1
	50	65.2	4.7	82.3	60.8	4.5	80.2
	100	67.4	4.2	84.7	64.0	4.2	83.5
	200	69.0	4.3	86.3	66.6	4.0	85.4
	400	71.2	4.5	87.9	67.8	4.1	86.5
0.05	25	68.5	4.7	83.9	58.0	4.5	78.7
	50	70.4	4.6	86.0	64.6	4.0	82.4
	100	73.3	3.6	88.0	68.9	3.8	84.8
	200	75.0	3.8	89.5	71.9	3.5	87.7
	400	78.0	4.0	90.7	73.7	3.7	90.0
0.1	25	71.3	4.6	84.4	59.5	4.3	78.7
	50	73.7	4.5	87.1	66.7	3.9	82.8
	100	77.7	3.5	90.0	71.7	3.4	86.4
	200	79.6	3.8	91.2	75.1	3.5	88.7
	400	81.4	3.5	92.8	78.0	3.1	91.0
0.25	25	74.2	4.0	85.5	62.2	4.0	79.7
	50	77.7	3.7	88.2	69.2	3.5	82.9
	100	80.8	3.5	90.7	74.3	2.9	86.9
	200	83.3	2.6	93.0	78.1	2.5	89.1
	400	85.7	2.8	94.8	81.8	2.6	92.3
0.5	25	75.7	3.9	85.6	62.8	3.8	79.5
	50	79.4	3.2	88.5	70.3	3.4	83.0
	100	82.6	3.3	91.0	76.0	3.4	87.5
	200	86.0	2.5	94.2	80.3	2.4	89.7
	400	88.1	2.5	95.7	83.4	2.4	91.8
1	25	76.4	3.9	85.3	64.0	4.0	79.5
	50	80.1	3.2	88.7	70.7	3.1	82.9
	100	83.9	2.7	91.3	76.2	2.9	86.9
	200	86.5	2.4	93.8	80.6	2.6	90.2
	400	89.0	2.2	95.3	84.1	2.1	92.5

genotypic value, compared to the minimum value. Thus when all 10 segregating loci are positive homozygous the genotype equals 100 per cent and with all segregating loci negative it equals 0 per cent. The mean of the phenotypic 10 best lines is also shown graphically in Fig. 2.

The influence of the number of segregating loci is also tested. Two other crosses are examined, which are identical to the one described above except for the number of heterozygous loci in the  $F_1$ . Cross two segregates for loci 8–25 (18 loci) and cross three for range 1–25 (25 loci). The results for all three crosses at a



**Fig. 2** The effects of the number of F<sub>2</sub> plants [= number of ES-F<sub>3</sub> and SSD-F<sub>6</sub> lines] on the genotypic mean of the 10 best lines resulting from both selection procedures at three different levels of heritability. (□)  $H^2 = 1.0$ , (▽)  $H^2 = 0.1$ , (○)  $H^2 = 0.02$ . (—) ES, (---) SSD.

heritability of 1.0 (which is the most discriminating) and for a varying number of  $NF_2$  are presented in Table 3; the genotypic means are plotted in Fig. 3.

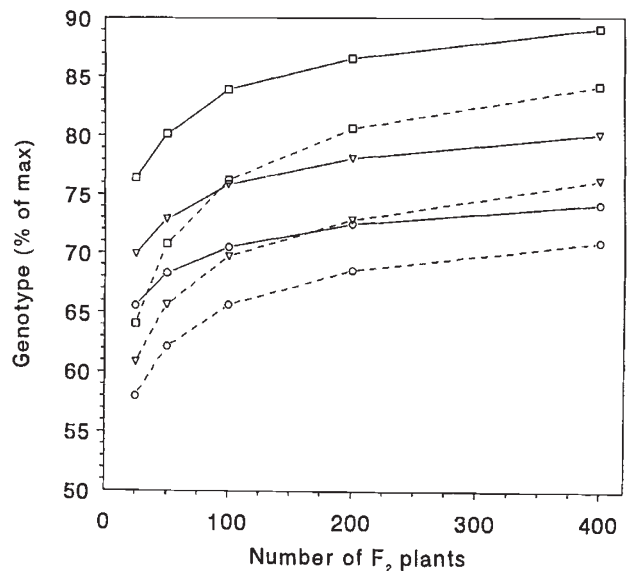
#### Five crosses

As ES comprises both selection between and within crosses, it is not fair to compare ES and SSD by only one cross; therefore, a breeding programme initiated with five crosses is considered. More crosses could be considered (and will be in an actual breeding programme) but our investigations were restricted to a subset of the potentially best crosses, which do not differ widely in population mean and variance. These five crosses have varying numbers and ranges of segregating loci. The characteristics of the crosses are given in Table 4.

The threshold value beyond which a recombinant inbred line is considered to perform well was set at 80 per cent of the best conceivable genotype (with all positive alleles accumulated). The expected probabilities of each cross that exceeded this threshold value are also given in Table 4. It can be seen that cross C has the highest probability with cross E second, although the latter can in fact deliver the highest yielding inbred line. Only the one best cross is selected and propagated with continuing line selection. The results of both breeding methods are given in Tables 5a and b, for various

**Table 3** Genotypic mean, standard error of that mean and maximum of the 10 phenotypic best lines in both the ES-F<sub>6</sub> and the SSD-F<sub>7</sub> at varying numbers of segregating loci. Mean and maximum expressed as a percentage of the difference between the theoretical maximum and minimum genotype.  $H^2$  equals 1. 100 runs

Segregating loci	$NF_2$	ES			SSD		
		Mean	s.e.	max.	Mean	s.e.	max.
25	25	65.5	2.7	71.4	57.9	2.5	67.4
	50	68.3	2.3	73.8	62.1	2.2	70.6
	100	70.5	1.6	76.0	65.6	1.8	72.9
	200	72.5	2.0	77.3	68.5	1.4	75.4
	400	74.1	1.5	79.1	70.9	1.5	76.9
18	25	69.9	2.6	76.7	60.8	2.5	72.5
	50	72.9	2.6	79.0	65.6	2.6	75.1
	100	75.8	2.3	82.0	69.7	1.9	78.2
	200	78.1	2.0	83.7	72.8	2.0	80.1
	400	80.1	1.9	85.4	76.2	1.7	82.6
10	25	76.4	3.9	85.3	64.0	4.0	79.5
	50	80.1	3.2	88.7	70.7	3.1	82.9
	100	83.9	2.7	91.3	76.2	2.9	86.9
	200	86.5	2.4	93.8	80.6	2.6	90.2
	400	89.0	2.2	95.3	84.1	2.1	92.5



**Fig. 3** The effects of the number of F<sub>2</sub> plants and the number of segregating loci on the genotypic mean of the 10 best lines resulting from both selection procedures. Heritability equals 1. (□) 10 loci, (▽) 18 loci, (○) 25 loci. (—) ES, (---) SSD.

heritabilities and two different numbers of F<sub>2</sub> plants; 50 and 100. Means of the 10 best lines for both selection methods in the case of 50 F<sub>2</sub> plants and for both one and five crosses are plotted in Fig. 4.

**Table 4** Genetic construction of the five crosses and the expected probability of retrieving well performing inbred lines in the  $F_{\infty}$  with a given threshold value (see text)

	Segregating loci	Range	<i>P</i>
Cross A	4	11-14	0.008
Cross B	6	9-14	0.003
Cross C	6	10-15	0.025
Cross D	8	8-15	0.008
Cross E	10	7-16	0.016

*Dominance*

In addition, the effect of dominance was tested on the selection results, in both cases of one and five crosses. Only unidirectional dominance was considered as it might produce the largest possible effect. It was applied on an intermediate ( $d = \frac{1}{2}a$ ) and a complete dominance level ( $d = a$ ) and results were compared with the situation where dominance was absent (Table 5a). Results are given in Tables 6a and 6b.

**Table 5a** Genotypic mean and maximum of the 10 phenotypic best lines of both selection methods. Two situations: five crosses and one cross; for the first the frequency of selection of each cross is also given. Mean and maximum expressed as a percentage of the difference between the theoretical maximum and minimum genotype. Fifty  $F_2$  plants. Number of runs = 300

	$H^2$					
	1	0.5	0.25	0.1	0.05	0.02
Five crosses						
Mean						
ES	82.3	79.9	76.8	73.2	69.6	65.4
SSD	80.9	80.0	77.8	74.5	71.2	66.5
cross A	79.4	77.8	75.4	71.8	69.6	66.6
cross B	75.6	72.8	71.7	67.7	63.8	60.6
cross C	83.4	82.7	79.5	76.9	73.5	70.1
cross D	79.7	77.6	74.4	71.3	66.8	63.0
cross E	81.8	80.5	78.0	74.7	70.5	65.0
Max.						
ES	88.1	86.1	84.5	83.5	81.3	79.5
SSD	88.0	88.0	88.0	86.3	85.1	82.5
cross A	80.0	80.0	80.5	79.6	79.1	77.7
cross B	80.0	79.1	79.6	77.9	75.5	74.1
cross C	88.7	88.7	87.5	86.9	84.9	84.5
cross D	85.3	85.6	82.9	82.7	80.4	79.1
cross E	90.0	90.1	88.4	88.1	84.9	83.1
Frequency						
cross A	0.07	0.22	0.25	0.25	0.30	0.31
cross B	0.01	0.07	0.10	0.11	0.14	0.17
cross C	0.54	0.37	0.37	0.28	0.26	0.21
cross D	0.10	0.12	0.14	0.17	0.13	0.19
cross E	0.28	0.23	0.15	0.19	0.17	0.13
One cross						
Mean						
ES	80.1	79.4	77.7	73.7	70.4	65.2
SSD	70.7	70.3	69.2	66.7	64.6	60.8
Max.						
ES	88.7	88.5	88.2	87.1	86.0	82.3
SSD	82.9	83.0	82.9	82.8	82.4	80.2

**Table 5b** Genotypic mean and maximum of the 10 phenotypic best lines of both selection methods. Two situations: five crosses and one cross; for the first the frequency of selection of each cross is also given. Mean and maximum expressed as a percentage of the difference between the theoretical maximum and minimum genotype. One hundred  $F_2$  plants. Number of runs = 200

	$H^2$					
	1	0.5	0.25	0.1	0.05	0.02
Five crosses						
Mean						
ES	86.8	83.6	79.9	76.1	72.6	68.2
SSD	82.8	82.2	79.9	75.6	72.2	67.3
cross A	80.0	79.2	77.0	74.4	72.6	68.7
cross B	78.7	77.0	73.7	69.5	67.3	63.5
cross C	87.9	85.9	83.6	80.0	77.2	72.8
cross D	83.1	80.5	79.0	75.0	70.9	65.1
cross E	87.8	86.5	82.6	79.4	74.7	70.3
Max.						
ES	89.7	88.1	86.0	84.8	82.6	81.3
SSD	89.6	90.4	89.0	86.8	85.9	82.7
cross A	80.0	80.0	80.0	79.8	79.6	78.2
cross B	80.0	80.0	79.6	77.6	77.4	76.3
cross C	89.9	89.7	89.3	88.4	87.3	85.2
cross D	88.6	87.0	86.7	85.1	82.6	80.4
cross E	93.9	94.4	91.3	91.3	88.4	87.8
Frequency						
cross A	0.11	0.20	0.20	0.25	0.27	0.27
cross B	0.01	0.07	0.13	0.14	0.22	0.16
cross C	0.62	0.37	0.29	0.23	0.22	0.21
cross D	0.04	0.10	0.19	0.17	0.13	0.22
cross E	0.23	0.27	0.18	0.21	0.17	0.15
One cross						
Mean						
ES	83.9	82.6	80.8	77.7	73.3	67.4
SSD	76.2	76.0	74.3	71.7	68.9	64.0
Max.						
ES	91.3	91.0	90.7	90.0	88.0	84.7
SSD	86.9	87.5	86.9	86.4	84.8	83.5

## Conclusions and discussion

It appears from Table 2 and Fig. 2 that better lines are retained when heritability is high, as was expected. Likewise, results are higher when more  $F_2$  plants (and therefore  $F_3$  and  $F_6$  lines) are taken. The advantage of ES over SSD is largest when the heritability is high and the number of  $F_2$  plants is low. Obviously the ES procedure is relatively inefficient when the  $F_3$  line estimates are influenced by a large environmental error. On the other hand, the SSD programme leads to poor results when less than 100  $F_2$  plants are taken.

The number of segregating loci that characterize the trait appears to have quite a large impact on the selec-

tion response. A much better line can be retained if only 10 loci are involved (90 per cent of the maximum versus 70 per cent with 25 loci). This is not surprising because the chance of retrieving a genotype with all positive alleles accumulated will be much smaller when 25 instead of 10 loci are segregating. Accordingly, the absolute differences between ES and SSD are slightly smaller when more loci are involved. The general trend is the same, however: if only one cross is considered, ES performs better than SSD, at least, with equal numbers of  $F_3/F_6$  lines.

These findings do agree with those of Casali & Tigchelaar (1975) concerning different procedures of plant selection in one population. They also concluded

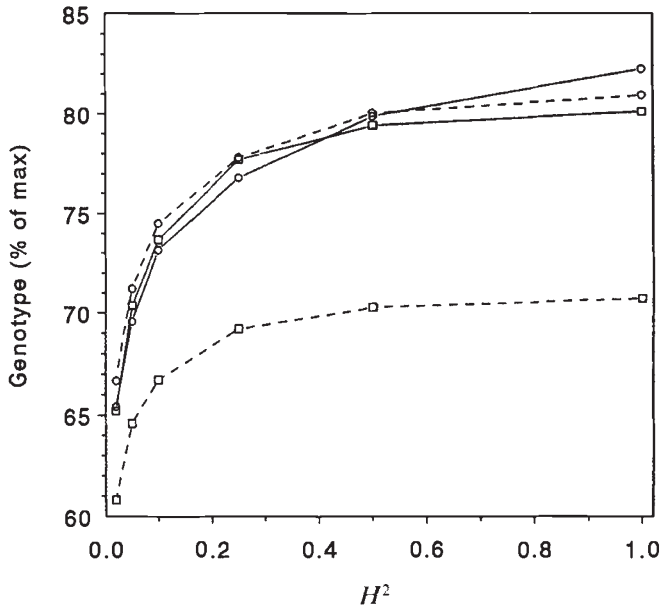


Fig. 4 The difference in selection results between ES (—) and SSD (---) in the case of selection within one cross ( $\square$ ) and between and within five crosses ( $\circ$ ), in dependence of the heritability. Fifty  $F_2$  plants.

that an (early) pedigree selection was more efficient than SSD at heritabilities varying from 0.1 to 1.0.

It is evident from Tables 5a and b that the advantage of ES over SSD in the case of one cross is completely absent in the more realistic situation of several crosses under selection. In the case of the one best line, the SSD method is even superior to ES at low and moderate heritability. As can be seen from the frequencies with which the different crosses are selected, the cross selection is close to random when heritability is low. When heritability increases the better crosses (C and E) are selected more frequently and the worst crosses (B and D) less frequently. It can be concluded that at low heritabilities the advantage of ES is completely lost due to erroneous cross selection. This was also found from field trials concerning ES (Van Ooijen, 1989b; Van Oeveren, 1992).

The effect of dominance is small. At moderate and high heritabilities ( $H^2 > 0.1$ ) genotypic  $F_6$  and  $F_7$  values are slightly higher compared to the corresponding situation where dominance is absent. This effect increases with the number of segregating loci. Because the level of heterozygosity in the  $F_6/F_7$  is not very high, the

Table 6a Genotypic mean and maximum of the 10 phenotypic best lines in both the ES- $F_6$  and the SSD- $F_7$  and the selection frequencies of each cross. Mean and maximum expressed as a percentage of the difference between the theoretical maximum and minimum genotype. Dominance  $d = \frac{1}{2}a$ ,  $NF_2 = 50$ , 200 runs

	$H^2$					
	1	0.5	0.25	0.1	0.05	0.02
Mean						
ES	83.2	80.1	77.6	73.7	70.1	65.4
SSD	81.0	80.1	78.4	74.6	71.6	67.1
cross A	79.3	77.7	75.5	72.8	70.3	66.8
cross B	—	72.7	70.9	66.3	65.3	59.3
cross C	84.1	82.8	80.3	76.8	73.5	70.1
cross D	78.8	76.8	75.3	71.4	67.6	63.2
cross E	82.5	80.3	79.4	75.5	71.0	66.1
Max.						
ES	89.3	86.7	85.4	84.2	82.2	79.7
SSD	87.8	87.7	88.1	87.0	84.9	82.8
cross A	80.0	80.0	79.9	79.7	79.6	78.0
cross B	—	80.0	78.2	76.8	75.7	73.2
cross C	89.1	89.0	88.2	86.7	85.4	83.3
cross D	85.8	84.3	83.6	83.1	80.9	79.6
cross E	90.6	89.3	89.7	89.2	87.2	85.6
Frequency						
cross A	0.02	0.17	0.19	0.20	0.24	0.25
cross B	0.00	0.04	0.06	0.11	0.15	0.21
cross C	0.59	0.37	0.32	0.30	0.26	0.23
cross D	0.05	0.11	0.22	0.18	0.16	0.15
cross E	0.34	0.32	0.22	0.23	0.20	0.15



**Table 6b** Genotypic mean and maximum of the 10 phenotypic best lines in both the ES-F<sub>6</sub> and the SSD-F<sub>7</sub> and the selection frequencies of each cross. Mean and maximum expressed as a percentage of the difference between the theoretical maximum and minimum genotype. Dominance  $d = a$ ,  $NF_2 = 50,200$  runs

	$H^2$					
	1	0.5	0.25	0.1	0.05	0.02
<b>Mean</b>						
ES	83.3	80.9	77.3	73.6	70.4	66.1
SSD	81.4	80.1	78.2	74.5	71.1	66.5
cross A	80.0	77.9	75.7	72.3	70.0	67.8
cross B	—	72.6	71.7	66.4	65.1	60.4
cross C	84.5	82.7	80.2	77.3	74.1	69.6
cross D	79.4	78.4	76.2	71.4	67.6	64.4
cross E	82.7	81.5	78.8	74.8	72.5	66.2
<b>Max.</b>						
ES	89.6	88.0	85.4	84.6	82.6	79.6
SSD	88.4	88.3	88.2	87.0	85.1	82.1
cross A	80.0	79.9	79.8	79.6	79.6	78.0
cross B	—	78.8	79.1	77.5	76.6	74.1
cross C	89.4	88.9	88.1	87.4	85.7	83.1
cross D	85.4	87.1	85.8	84.5	81.1	78.8
cross E	90.6	90.5	88.9	88.7	87.5	84.6
<b>Frequency</b>						
cross A	0.02	0.11	0.18	0.21	0.22	0.27
cross B	0.00	0.03	0.12	0.09	0.16	0.18
cross C	0.44	0.39	0.29	0.29	0.26	0.23
cross D	0.05	0.13	0.18	0.20	0.16	0.18
cross E	0.50	0.36	0.24	0.22	0.20	0.15

genotypic line values do not differ much from those in the completely additive situation. The effect of dominance on the SSD procedure is therefore small. It could have an effect on the ES programme if the cross prediction changes. As can be seen from the frequencies of selected crosses, there is a tendency towards a higher selection frequency of the crosses with the largest variances. Because they are, in this case, also the most promising crosses, the selection results could be higher than results in the completely additive situation. The ES-F<sub>7</sub> line genotypes are indeed slightly higher at high and moderate heritabilities but it is not clear to what extent this is due to the selection of a better cross or to the fact that dominance increases the value of heterozygotes. If the crosses with the largest variance are not the most promising, dominance could have a negative effect on the final selection result. In any case, the overall effect is small.

It is interesting to examine the total trial field area needed for both programmes. The size of a large yield plot (F<sub>7</sub>-type) will be about four times the size of a small three-row plot (F<sub>3</sub>-type). In addition, the F<sub>7</sub>-type lines

are grown in four replicates instead of two, so they occupy eight times as much space as the F<sub>3</sub>-type lines. The total number of lines used for the F<sub>3</sub> to the F<sub>6</sub>/F<sub>7</sub> generation is given in Table 7, together with the total area needed, expressed as the number of F<sub>3</sub>-type plots. It can be seen that, with a single cross, ES occupies much more space than SSD. When equal trial sizes are engaged for both procedures, the advantage of ES over SSD will be less obvious and maybe even absent. In the case of five crosses, the ES trial is larger than the SSD trial except for  $NF_2 = 50$ . In most cases this implies an even larger advantage of SSD over ES when equal trial sizes are used. Of course greenhouse area is also necessary for the early SSD generations. This is partly compensated in the above comparison by the fact that the last SSD generation is one ahead of the last ES generation. When a true economically based comparison is made, other factors have to be regarded, such as the amount of labour and materials.

Some other aspects have not been considered. For instance, a genotype-environment (-year) interaction can have a large potential bias on selection results. It

**Table 7** Total number of  $F_3$ -type ( $m_1$ ) and  $F_6$ -type ( $m_2$ ) lines for both selection procedures, in relation to the number of  $F_2$  plants ( $NF_2$ ), in the case of One and five crosses. Calculated from these are the total field trial area (t.a.) occupied by them, expressed as the number of  $F_3$ -type plots, and the according ratio

	$NF_2$				
	25	50	100	200	400
<b>One cross</b>					
<b>ES</b>					
$m_1$	125	150	200	300	500
$m_2$	70	75	100	150	250
t.a.	1370	1500	2000	3000	5000
<b>SSD</b>					
$m_1$	25	50	100	200	400
$m_2$	20	25	50	100	100
t.a.	370	500	1000	2000	2400
ES/SSD	3.7	3	2	1.5	2.1
<b>Five crosses</b>					
<b>ES</b>					
$m_1$	225	350	600	1100	2100
$m_2$	70	75	100	150	250
t.a.	1570	1900	2800	4600	8200
<b>SSD</b>					
$m_1$	125	250	500	1000	2000
$m_2$	62	100	100	100	100
t.a.	1242	2100	2600	3600	5600
ES/SSD	1.3	0.9	1.1	1.3	1.5

would probably most influence the ES procedure because the chance of discarding potentially good crosses or lines due to a bad performance in an occasional selection environment would be high. This

would increase the advantage of SSD. Another aspect is the number of selected crosses. The results of ES are likely to improve when a second or third cross is kept.

## References

- BOERMA, H. R. AND COOPER, R. L. 1975. Comparison of three selection procedures for yield in soybeans. *Crop Sci.*, **15**, 225-229.
- CASALI, V. W. D. AND TIGCHELAAR, E. C. 1975. Computer simulation studies comparing pedigree, bulk and single seed descent selection in self pollinated populations. *J. Am. Soc. Hort. Sci.*, **100**, 364-367.
- FAIRFIELD SMITH, H. 1938. An empirical law describing heterogeneity in the yields of agricultural crops *J. Agric. Sci.*, **28**, 1-23.
- JINKS, J. L. AND POONI, H. S. 1980. Comparing predictions of mean performance and environmental sensitivity of recombinant inbred lines based upon  $F_3$  and triple test cross families. *Heredity*, **45**, 305-312.
- KNOTT, D. R. AND KUMAR, J. 1975. Comparison of early generation yield testing and a single seed descent procedure in wheat breeding. *Crop Sci.*, **15**, 295-299.
- MATHER, K. AND JINKS, J. L. 1971. *Biometrical Genetics*. Chapman and Hall, London, 382 pp.
- OEVEREN, A. J. VAN, 1992. A comparison between single seed descent and early cross selection in wheat breeding. *Euphytica*, **58**, 275-287.
- OIJEN, J. W. VAN, 1989a. The predictive value of quantitative genetic parameters in autogamous crops: bias caused by intergenotypic competition. 1.  $F_\infty$ -mean. *Euphytica*, **44**, 81-94.
- OIJEN, J. W. VAN, 1989b. The predictive value of quantitative genetic parameters in autogamous crops: bias caused by intergenotypic competition. 2.  $F_\infty$ -variance. *Euphytica*, **44**, 95-108.